

Titre: Identification de biomarqueurs non invasifs et multi-omiques pour la stratification pronostique des métastases hépatiques colorectales et du cancer pulmonaire basée sur l'apprentissage automatique
Title:

Auteur: Ralph Saber
Author:

Date: 2025

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Saber, R. (2025). Identification de biomarqueurs non invasifs et multi-omiques pour la stratification pronostique des métastases hépatiques colorectales et du cancer pulmonaire basée sur l'apprentissage automatique [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. <https://publications.polymtl.ca/65810/>
Citation:

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/65810/>
PolyPublie URL:

Directeurs de recherche: Samuel Kadoury, & Simon Turcotte
Advisors:

Programme: Génie biomédical
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Identification de biomarqueurs non invasifs et multi-omiques pour la
stratification pronostique des métastases hépatiques colorectales et du cancer
pulmonaire basée sur l'apprentissage automatique**

RALPH SABER

Institut de génie biomédical

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie biomédical

Mai 2025

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

**Identification de biomarqueurs non invasifs et multi-omiques pour la
stratification pronostique des métastases hépatiques colorectales et du cancer
pulmonaire basée sur l'apprentissage automatique**

présentée par **Ralph SABER**

en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Guy CLOUTIER, président

Samuel KADOURY, membre et directeur de recherche

Simon TURCOTTE, membre et codirecteur de recherche

Quoc-Huy Vincent TRINH, membre

Martin VALLIÈRES, membre externe

DÉDICACE

À ceux qui m'ont donné tout ce que j'ai . . .

REMERCIEMENTS

Je tiens tout d’abord à exprimer ma profonde gratitude à mon directeur de thèse, le Professeur Samuel Kadoury, pour son encadrement, ses conseils et son soutien tout au long de ce parcours. Votre expertise et votre accompagnement ont été des piliers dans l’aboutissement de ce travail. Je remercie également mon co-directeur, le Docteur Simon Turcotte, pour ses idées éclairantes, son accompagnement attentif et son engagement indéfectible. Votre contribution a été inestimable.

Je souhaite également remercier les membres du laboratoire MedICAL pour les moments agréables et les discussions enrichissantes qui ont marqué ces années.

Un immense merci à ma famille, dont le soutien inconditionnel et l’amour m’ont porté dans les moments les plus difficiles. Vous avez toujours cru en moi, même lorsque le doute s’installait, et je vous dois tout.

Je tiens aussi à exprimer ma reconnaissance au Fonds de recherche du Québec – Nature et technologies (FRQNT) pour le financement qui a rendu possible la réalisation de cette thèse.

Enfin, je remercie du fond du cœur tous mes amis qui m’ont soutenu tout au long de ce chemin. Vos encouragements, votre écoute et votre présence ont été une source inestimable de motivation et de réconfort.

À vous tous, merci infiniment.

RÉSUMÉ

Le cancer colorectal (CCR) et le cancer du poumon figurent parmi les quatre types de cancer les plus fréquemment diagnostiqués au Canada. En 2024, le cancer du poumon était la première cause de décès liés au cancer au Canada et le CCR se classait au deuxième rang chez les hommes canadiens et au troisième rang chez les femmes canadiennes. Les métastases hépatiques colorectales (MHCR) se développent comme le site le plus courant de progression du cancer colorectal. Actuellement, la résection chirurgicale complète des MHCR, associée à une chimiothérapie systémique, est la principale voie vers une guérison potentielle. Néanmoins, plus de 70 % des patients présentent une récurrence de la maladie après l'opération. Le cancer du poumon non à petites cellules (CPNPC) représente 85 % des cas de cancer du poumon, ce qui en fait le type le plus courant. L'introduction des inhibiteurs de point de contrôle immunitaire (ICI) a révolutionné la prise en charge du CPNPC, entraînant des améliorations substantielles des taux de survie des patients depuis l'avènement de l'immunothérapie. Pourtant, le taux de réponse objective du CPNPC avancé aux ICI est inférieur à 43 %.

Cette incertitude pronostique causée par les résultats oncologiques hétérogènes de ces deux tumeurs malignes souligne un défi majeur en oncologie : le décalage entre des thérapies de plus en plus sophistiquées et la capacité à prédire de manière fiable la réponse des patients à des traitements spécifiques. Il existe un besoin critique d'approches de médecine personnalisée qui peuvent prédire avec précision le pronostic des patients et guider des décisions thérapeutiques adaptées. Actuellement, la médecine personnalisée repose sur des biomarqueurs qui aident à décrypter la biologie tumorale impliquée et les caractéristiques du microenvironnement immunitaire. Les biomarqueurs conventionnels sont collectés à partir d'échantillons de tissus, traditionnellement obtenus par biopsie ou évalués après une intervention chirurgicale par immunohistochimie (IHC). Néanmoins, ces méthodes souffrent d'hétérogénéité spatiale, de risques procéduraux et de limitations techniques, notamment des différences dans la qualité des échantillons et les protocoles de coloration. De plus, les biopsies répétitives sont impraticables, ce qui rend difficile le suivi longitudinal des changements de biomarqueurs pendant le traitement.

Des images de tomographie par densitométrie (TDM) sont régulièrement acquises pour le diagnostic initial et le suivi continu des MHCR et du CPNPC. L'intégration des caractéristiques de TDM avec des outils d'intelligence artificielle pourrait aider à saisir les caractéristiques du microenvironnement immunitaire de la tumeur et à en tirer des informations pronostiques.

En outre, les approches multi-omiques intégrant des données médicales provenant de sources multiples pourraient prédire de manière fiable la biologie des tumeurs et la progression du cancer grâce à des outils d'apprentissage profond de pointe.

Cette thèse vise à combler le décalage entre la biologie des tumeurs et la prise de décision clinique en développant des pipelines informatiques pour l'identification de biomarqueurs non invasifs du cancer ainsi que des approches multi-omiques pour améliorer la sélection des traitements et le pronostic du cancer. Cette thèse comprend quatre études interconnectées, chacune abordant un objectif spécifique et contribuant à l'objectif d'amélioration du pronostic. Le premier objectif est de concevoir, mettre en œuvre et valider un pipeline informatique pour prédire des substituts non invasifs et pertinents sur le plan pronostique des biomarqueurs immunitaires dans le cancer du foie, en particulier l'expression de l'ecto-5'-nucléotidase, ou CD73, en utilisant des images de TDM préopératoires. Une étude de suivi vise à classer le profil immunologique CD3/CMH-I dans le MHCR à l'aide d'un scanner préopératoire. Le troisième objectif est de développer et d'évaluer un outil permettant de déduire, à partir d'un scanner pré-traitement, le niveau d'expression du ligand de mort programmé 1 (PD-L1) dans le CPNPC avancé et limité traité par ICI seul ou en combinaison avec une chimiothérapie. Le quatrième objectif est de proposer une nouvelle approche multi-omique pour la prédiction du risque de récurrence du MHCR à deux ans après résection chirurgicale.

La première étude vise à prédire de manière non invasive l'expression de CD73 dans les métastases hépatiques du cancer du poumon. CD73 est une enzyme qui catalyse la conversion de l'AMP en adénosine immunosuppressive. Elle joue donc un rôle pronostique dans le microenvironnement des MHCR et constitue une cible thérapeutique potentielle. Nous avons entraîné un modèle d'apprentissage tabulaire interprétable et attentif (TabNet) pour prédire les niveaux d'expression stratifiés de CD73 à l'aide de caractéristiques radiomiques extraites d'images de TDM préopératoires. L'ensemble de données comprenait 160 MHCR segmentés obtenus à partir de 122 patients avec des données de microréseaux de tissu correspondantes utilisées comme vérité terrain. TabNet a montré une performance prédictive élevée et le score dérivé du modèle, appelé rad-CD73, a été corrélé avec l'expression histologique de CD73. Les patients présentant un rad-CD73 élevé ont eu une récurrence plus précoce et une survie plus courte, ce qui suggère que ce biomarqueur non invasif peut avoir une pertinence pronostique.

La deuxième étude s'est concentrée sur l'identification non invasive des patients atteints de MHCR présentant un profil immunologique caractérisé par une faible densité de cellules T CD3+ associée à une forte expression du CMH-I ($CD3^{Low}CMH^{High}$). Les lymphocytes infiltrant les tumeurs (TIL), en particulier les cellules T CD3+, et l'expression du CMH-I sont des déterminants essentiels de la réponse immunitaire ciblant les cellules tumorales.

Le profil $CD3^{Low}CMH^{High}$ est également associé à un pronostic et une réponse plus défavorables à la thérapie systémique. En proposant un pipeline d'apprentissage d'ensemble, appelé Multi-TabNet, l'étude a obtenu une précision prédictive élevée et a surpassé d'autres modèles d'apprentissage automatique. Le score immunitaire proposé, basé sur $CD3/CMH-I$, a également été associé aux résultats des patients. L'étude montre que les caractéristiques d'imagerie pourraient être associées aux densités d'infiltration immunitaire.

La troisième étude visait à prédire de manière non invasive l'expression de PD-L1 dans le CPNPC susceptible de répondre aux ICI. PD-L1 inhibe les réponses immunitaires antitumorales. Pour les patients atteints d'un CPNPC avancé sans mutations conductrices, c'est le biomarqueur le plus établi pour les ICI et il joue un rôle central dans l'orientation des stratégies de traitement. Cette étude a proposé un pipeline d'auto-apprentissage intégrant le modèle Feature Tokenizer Transformer pour prédire l'expression de PD-L1 à l'aide d'images de TDM de prétraitement. Le pipeline a été entraîné et validé en interne sur un ensemble de données multi-institutionnelles comprenant des patients atteints de CPNPC avancé provenant de trois centres médicaux. Il a obtenu de solides performances prédictives lors de la validation interne ($AUC = 0,75$). Il a également fait preuve d'une bonne généralisabilité sur une cohorte indépendante de patients atteints de CPNPC à un stade limité traités par ICI néoadjuvant et chimiothérapie, la norme de soins actuelle pour les patients éligibles atteints d'une maladie résécable. En outre, les patients présentant des scores radiomiques PD-L1 élevés ont bénéficié d'une survie sans progression plus longue, ce qui souligne l'utilité clinique de ce biomarqueur non invasif. Cette méthode répond à la demande cruciale de biomarqueurs fiables et non invasifs pour éclairer les choix de traitement par ICI chez les personnes atteintes d'un cancer du poumon.

La quatrième étude a proposé un pipeline multi-omique pour prédire le risque de récurrence des MHCR dans les deux ans suivant la résection chirurgicale. Compte tenu de la complexité de la biologie tumorale, cette étude a utilisé collectivement des images de TDM préopératoires, des données de séquençage de l'ARN et des caractéristiques cliniques dans un pipeline pronostique complet. De plus, nous avons abordé le problème des transcrits de gènes d'ARN non détectés en utilisant un mécanisme d'imputation antagoniste générative. Nous avons également abordé la dimensionnalité élevée des ensembles de données d'expression génique en synthétisant des instances artificielles par apprentissage génératif. Le module TDM comprend un modèle de fondation pour extraire des caractéristiques qui complètent les caractéristiques radiomiques. Notre pipeline multi-omique a atteint une AUC de 0,75 sur un ensemble de données monocentrique maintenu de manière prospective de 129 patients atteints de MHCR. Il a surpassé les approches unimodale et bimodale et a montré une amélioration de 7,14 % par rapport au score de risque clinique, un outil de pronostic du MHCR établi cliniquement. Il a

également fourni une stratification significative des risques, comme le reflètent les courbes de Kaplan-Meier. L'approche proposée pourrait être exploitée pour stratifier les patients à risque élevé qui seraient susceptibles de bénéficier d'un traitement adjuvant et d'une surveillance intensifiée pour la détection et la gestion précoces des récidives. Elle pourrait également guider le développement d'essais cliniques évaluant des thérapies adjuvantes innovantes.

Collectivement, ces études ouvrent la voie à une prise en charge plus personnalisée des MHCR et des CPNPC. Un rad-CD73 validé de manière prospective pourrait servir de biomarqueur non invasif des agents thérapeutiques ciblant le CD73 ou les récepteurs de l'adénosine, tandis que le profil d'imagerie $CD3^{Low}CMH^{High}$ pourrait permettre une évaluation immunitaire non invasive des MHCR. Ils peuvent également être utilisés pour orienter de manière dynamique les thérapies adjuvantes et la fréquence des suivis après l'opération. Le score radiomique PD-L1 pourrait être utilisé dans des études longitudinales pour évaluer la résistance initiale et acquise aux ICI. L'approche multiomique démontre le potentiel de l'intégration de l'imagerie, de la transcriptomique et des données cliniques grâce à des techniques avancées d'apprentissage automatique pour améliorer la précision de la pronostication des MHCR. Cette thèse vise à révolutionner le pronostic du cancer en passant d'approches réactives à des stratégies proactives qui exploitent les données multiomiques intégrées pour éclairer la prise de décision thérapeutique.

ABSTRACT

Colorectal cancer (CRC) and lung cancer are among the top four most diagnosed cancer types in Canada. In 2024, lung cancer was the first cause of cancer-related Canadian deaths and CRC ranked second in Canadian men and third in Canadian women. Colorectal liver metastases (CRLM) develop as the most common site of colorectal cancer progression. Currently, complete surgical resection of CRLM combined with systemic chemotherapy, is the main path to a potential cure. Nevertheless, more than 70% of patients experience disease recurrence after surgery. Non-small cell lung cancer (NSCLC) accounts for 85% of lung cancer cases, making it the most common type. The introduction of immune checkpoint inhibitors (ICI) has revolutionized the management of NSCLC, resulting in substantial improvements in patient survival rates since the advent of immunotherapy. Yet, the objective response rate of advanced NSCLC to ICI is lower than 43%.

This prognostic uncertainty caused by the heterogeneous oncological outcomes of these two malignancies underlines a major challenge in oncology: the disconnect between increasingly sophisticated therapies and the ability to reliably predict patients' response to specific treatments. There is a critical need for personalized medicine approaches that can accurately predict patient prognosis and guide tailored therapeutic decisions. Currently, personalized medicine relies on biomarkers that help decipher the involved tumor biology and the characteristics of the immune microenvironment. Conventional biomarkers are collected using tissue samples, traditionally obtained through biopsies or assessed after surgery using immunohistochemistry (IHC). Nevertheless, these methods suffer from spatial heterogeneity, procedural risks, technical limitations including differences in sample quality and staining protocols. Moreover, repetitive biopsies are impracticable making it difficult to longitudinally track biomarker changes during treatment.

Computed tomography (CT) images are routinely acquired for initial diagnosis and ongoing monitoring of both CRLM and NSCLC. The integration of CT features with artificial intelligence tools could help capture the characteristics of the tumor immune microenvironment and derive prognostic insights. Moreover, multi-omics approaches integrating medical data from multiple sources could reliably predict tumor biology and cancer progression through state-of-the-art deep learning tools.

This thesis aims to bridge the gap between tumor biology and clinical decision-making by developing computational pipelines for the identification of non-invasive cancer biomarkers along with multiomics approaches to improve treatment selection and cancer prognosis.

This thesis comprises four interconnected studies, each addressing a specific objective and contributing to the goal of improved prognostication. The first objective is to design, implement, and validate a computational pipeline for predicting prognostically relevant, non-invasive surrogates of immune biomarkers in CRLM, particularly the expression of ecto-5'-nucleotidase, or CD73, utilizing preoperative CT images. A follow-up study aims to classify the CD3/MHC-I immunological profile in CRLM using presurgical CT. The third objective is to develop and evaluate a tool to infer, from pre-treatment CT, the expression level of PD-L1 in advanced and limited NSCLC treated with ICI alone or in combination with chemotherapy. The fourth objective is to propose a novel multi-omics approach for the prediction of the two-year CRLM recurrence risk following surgical resection.

The first study seeks to non-invasively predict CD73 expression in CRLM. CD73, an enzyme that catalyzes the conversion of AMP to immunosuppressive adenosine. Thus, it plays a prognostic role within CRLM microenvironment and is a potential therapeutic target. We trained an Attentive Interpretable Tabular Learning (TabNet) model to predict stratified CD73 expression levels using radiomic features extracted from preoperative CT images. The dataset included 160 segmented CRLM obtained from 122 patients with corresponding microarray data used as ground truth. TabNet exhibited high predictive performance and the model-derived score, termed rad-CD73, was correlated with histological CD73 expression. Patients with high rad-CD73 had an earlier recurrence and a shorter survival, suggesting that this non-invasive biomarker can hold prognostic relevance.

The second study focused on the non-invasive identification of CRLM patients exhibiting an immunological profile characterized by low CD3+ T cell density concurrent with a high MHC-I expression ($CD3^{Low}MHC^{High}$). Tumor-infiltrating lymphocytes (TILs), particularly CD3+ T cells, and MHC-I expression are critical determinants of immune response targeting tumor cells. The $CD3^{Low}MHC^{High}$ profile is also associated with worse prognosis and response to systemic therapy. By proposing an ensemble learning pipeline, termed Multi-TabNet, the study achieved a strong predictive accuracy and outperformed other machine learning models. The proposed immune scoring based on CD3/MHC-I was also associated with patients' outcomes. The study shows that imaging features could be associated with immune infiltration densities.

The third study aimed to non-invasively predict the Programmed Cell Death Ligand 1 (PD-L1) expression in NSCLC amenable to ICI. PD-L1 inhibits anti-tumor immune responses. For patients with advanced NSCLC without driver mutations, it is the most established biomarker for ICI and it plays a pivotal role in guiding treatment strategies. This study proposed a self-training pipeline incorporating the Feature Tokenizer Transformer model to predict PD-L1

expression using pretreatment CT images. The pipeline was trained and internally validated on a multi-institutional dataset comprising advanced NSCLC patients from three medical centers. It achieved strong predictive performance in internal validation (AUC=0.75). It also exhibited a good generalizability on an independent cohort of patients with limited-stage NSCLC treated with neoadjuvant ICI and chemotherapy, the current standard of care for eligible patients with resectable disease. Furthermore, patients with high radiomic PD-L1 scores were associated with longer progression-free survival, highlighting the clinical utility of this non-invasive biomarker. This method tackles the crucial demand for reliable, non-invasive biomarkers to inform ICI treatment choices for individuals with lung cancer.

The fourth study proposed a multi-omics pipeline to predict the risk of CRLM recurrence within two years after surgical resection. Taking into account the complexity of tumor biology, this study collectively used preoperative CT images, RNA sequencing data, and clinical features into a comprehensive prognostic pipeline. Moreover, we tackled the problem of undetected RNA gene transcripts by leveraging a generative adversarial imputation mechanism. We also addressed the high dimensionality of gene expression datasets by synthesizing artificial instances through generative learning. The CT module includes a foundation model to extract features that complement radiomics features. Our multi-omics pipeline achieved an AUC of 0.75 on a single-center prospectively maintained dataset of 129 CRLM patients. It outperformed its single and dual-modal approaches and exhibited a 7.14% improvement over the Clinical Risk Score, a clinically established CRLM prognostication tool. It also provided significant risk stratification as mirrored by the Kaplan-Meier curves. The proposed approach could be leveraged to stratify patients with elevated risk who would likely benefit from adjuvant therapy and intensified surveillance for early recurrence detection and management. It could also guide the development of clinical trials evaluating innovative adjuvant therapies.

Collectively, these studies pave the way for a more personalized CRLM and NSCLC management. A prospectively validated rad-CD73 could serve as a non-invasive biomarker of therapeutic agents targeting CD73 or adenosine receptors while the imaging $CD3^{Low}MHC^{High}$ profile could enable a non-invasive immune scoring of CRLM. They can also be leveraged to dynamically guide adjuvant therapies and the frequency of follow-up after surgery. The radiomic PD-L1 score could be used in longitudinal studies to assess initial and acquired resistance to ICI. The multiomics approach demonstrates the potential of integrating imaging, transcriptomics, and clinical data through advanced machine learning techniques to enhance prognostic accuracy in CRLM. This thesis seeks to revolutionize cancer prognostication by shifting from reactive approaches to proactive strategies that leverage integrated multiomics data to inform therapeutic decision-making.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	ix
LISTE DES TABLEAUX	xvii
LISTE DES FIGURES	xviii
LISTE DES SIGLES ET ABRÉVIATIONS	xxvii
LISTE DES ANNEXES	xxix
CHAPITRE 1 INTRODUCTION	1
1.1 Contributions	2
1.2 Structure de la thèse	4
CHAPITRE 2 CONTEXTE	6
2.1 Les métastases hépatiques colorectales	6
2.1.1 Le score de risque clinique	6
2.2 Le cancer pulmonaire non à petites cellules	6
2.3 Outils diagnostiques et pronostiques	7
2.3.1 Imagerie médicale	7
2.3.2 Histopathologie	8
2.3.3 Séquençage d'ARN	9
2.4 Biomarqueurs immunitaires du cancer	11
2.4.1 Infiltrats immunitaires et immunoscore	11
2.4.2 Cibles thérapeutiques	11
2.5 Prise en charge des métastases hépatiques colorectales	14
2.6 Les normes de traitement du CPNPC	14
2.7 La radiomique	16
2.8 Apprentissage automatique	18
CHAPITRE 3 REVUE DE LITTÉRATURE	20

3.1	Apprentissage automatique en imagerie médicale du cancer	20
3.1.1	Pronostication basée sur l'imagerie du cancer	21
3.1.2	Prédiction de biomarqueurs pronostiques à l'aide de l'imagerie du cancer	24
3.2	Apprentissage automatique en génomique du cancer	27
3.2.1	Pronostication du cancer basée sur la transcriptomique	28
3.3	Pronostication du cancer basé sur les caractéristiques cliniques	29
3.4	Approches multimodales d'apprentissage profond	30
3.5	Techniques d'apprentissage profond pour la recherche sur le cancer	35
3.5.1	Apprentissage profond pour les données tabulaires	37
3.5.2	Modèles génératifs	38
3.6	Paradigmes d'apprentissage avancés	40
3.6.1	Apprentissage auto-supervisé et semi-supervisé	40
3.6.2	Modèles de fondation dans le domaine biomédical	41
3.7	Résumé	42
CHAPITRE 4 MÉTHODOLOGIE DE RECHERCHE		44
4.1	Énoncé de la problématique	44
4.2	Hypothèses	45
4.3	Objectifs	46
4.4	Méthodologie générale	46
4.4.1	Bases de données	48
CHAPITRE 5 ARTICLE 1 : RADIOMICS USING COMPUTED TOMOGRAPHY TO PREDICT CD73 EXPRESSION AND PROGNOSIS OF COLORECTAL CAN- CER LIVER METASTASES		51
5.1	Abstract	52
5.1.1	Background	52
5.1.2	Methods	52
5.1.3	Results	52
5.1.4	Conclusions	52
5.1.5	Keywords	53
5.2	Background	53
5.3	Methods	54
5.3.1	Study population	54
5.3.2	Evaluation of CD73 expression	55
5.3.3	Image preprocessing and radiomics workflow	56
5.3.4	TabNet training and evaluation	57

5.3.5	Model interpretability analysis	57
5.3.6	Statistical analysis	58
5.4	Results	59
5.4.1	Patient cohort	59
5.4.2	Prediction of CD73 expression from preoperative CT images	59
5.4.3	Interpretability of the predictive model	61
5.4.4	Association of rad-CD73 with clinicopathological characteristics and oncological outcomes	64
5.5	Discussion	67
5.6	Conclusions	70
5.7	Abbreviations	70
5.8	Acknowledgments	71
5.9	Authors' contributions	71
5.10	Funding	71
5.11	Availability of data and materials	72
5.12	Ethics approval and consent to participate	72
5.13	Consent for publication	72
5.14	Competing interests	72
CHAPITRE 6 ARTICLE 2 : ENSEMBLE TABNET PREDICTING A T-CELL/MHC-I-BASED IMMUNE PROFILE BIOMARKER FOR COLORECTAL LIVER METASTASES FROM CT IMAGES		
6.1	Abstract	75
6.2	Introduction	76
6.3	Materials and Methods	77
6.3.1	Patient cohort	77
6.3.2	Data preparation	77
6.3.3	Radiomic feature extraction	79
6.3.4	Prediction framework for T-cell/MHC-I profile	79
6.4	Results and discussion	81
6.4.1	Evaluation methodology	81
6.4.2	Prediction of the CD3/MHC-I-based immune profile	82
6.4.3	Prognostic value of the predicted immune profile	83
6.5	Conclusion	83
6.6	Compliance with ethical standards	84
6.7	Acknowledgments	84

CHAPITRE 7	ARTICLE 3 : MULTI-INSTITUTIONAL CT SCAN-BASED RADIO-	
	MICS FOR PREDICTING TUMOR PD-L1 EXPRESSION IN PATIENTS WITH	
	ADVANCED AND LIMITED NON-SMALL CELL LUNG CANCER	85
7.1	Abstract	85
	7.1.1 Keywords	86
7.2	Introduction	86
7.3	Materials and methods	88
	7.3.1 Study population and data collection	88
	7.3.2 CT image preparation and pre-processing	89
	7.3.3 Radiomics workflow	90
	7.3.4 Self-training-based PD-L1 predictive pipeline	90
	7.3.5 Interpretability of the AI-based pipeline	92
	7.3.6 Statistical analysis	92
7.4	Results	93
	7.4.1 Patient cohort	93
	7.4.2 Prediction of PD-L1 expression and validation on the neoadjuvant cohort	93
	7.4.3 Interpretability analysis	94
	7.4.4 Decision curve analysis	95
	7.4.5 Association of rad-PDL1 with clinical outcomes	96
7.5	Discussion	96
7.6	Conclusion	98
7.7	Acknowledgments	99
7.8	Declaration of interest statement	99
7.9	Funding	99
7.10	Data availability statement	99
7.11	Authors' contributions	100
CHAPITRE 8	ARTICLE 4 : MULTI-OMICS FUSION NETWORK FOR PREDIC-	
	TION OF EARLY RECURRENCE IN COLORECTAL LIVER METASTASES .	105
8.1	Abstract	105
	8.1.1 Keywords	106
8.2	Introduction	106
8.3	Background	108
	8.3.1 CRLM outcome prediction from CT	108
	8.3.2 CRC outcome prediction from RNA sequencing	109
	8.3.3 CRLM outcome prediction from clinical models	110

8.4	Materials and methods	110
8.4.1	Patient cohort	110
8.4.2	Data preparation	111
8.4.3	RNA sequencing module	112
8.4.4	Imaging module	117
8.4.5	Multi-omics feature fusion	118
8.4.6	Statistical and interpretability analyses	119
8.4.7	Training and implementation details	119
8.5	Results	120
8.5.1	Comparison across modalities	120
8.5.2	Performance of the imaging component	120
8.5.3	Prognostic value of the proposed pipeline	121
8.5.4	Clinical utility of the proposed pipeline	122
8.5.5	Comparison with state-of-the-art tabular learning models	123
8.5.6	Comparison of different fusion approaches	123
8.5.7	Interpretability analysis	124
8.6	Discussion	126
8.7	Conclusion	131
CHAPITRE 9	DISCUSSION	135
9.1	Prédiction des marqueurs immunitaires non invasifs dans les MHCR	136
9.2	Prédiction de l'expression de PD-L1 dans le CPNPC	138
9.3	Stratification du risque de récurrence multi-omique dans le cas des MHCR	140
9.4	Résumé	141
CHAPITRE 10	CONCLUSION ET RECOMMANDATIONS	144
10.1	Contribution à l'avancement des connaissances	144
10.2	Limites	145
10.3	Améliorations futures	146
RÉFÉRENCES		148
ANNEXES		185

LISTE DES TABLEAUX

Tableau 5.1	Patient clinicopathological characteristics and correlation with rad-CD73	60
Tableau 5.2	Performance of the different models on the hold-out test set. AUC, Area Under the Curve; TabNet, Attentive Interpretable Tabular Learning; XGBoost, Extreme Gradient Boosting; SVM, Support Vector Machine; RF, Random Forest; LR, Logistic Regression.	62
Tableau 5.3	Univariate analyses of outcomes according to clinicopathological variables and rad-CD73. HR, Hazard Ratio; CI, Confidence interval. . .	66
Tableau 5.4	Multivariate analyses of outcomes according to clinical risk score and rad-CD73. HR, Hazard Ratio; CI, Confidence interval.	67
Tableau 6.1	Performance on the hold-out test set of the different models on the prediction of the CD3/MHC-I-based immune profile. Top results are shown in bold. 95% confidence intervals are represented between brackets.	82
Tableau 7.1	Clinicopathological characteristics of the primary and neoadjuvant cohorts.	94
Tableau 7.2	Multivariate analysis of the progression-free survival (PFS) according to the rad-PDL1 score and clinical features. The rad-PDL1 score was a predictor of the PFS, independently of the clinical features.	96
Tableau 8.1	Ablation study results. The mean and standard deviation of the performance metrics in cross-validation are reported in the table. The performance observed with the proposed multi-omics approach was the only one that exhibited a statistically significant difference with the model trained using clinical features (marked with an asterisk). SA ; synthetic augmentation.	123
Tableau 8.2	Comparison of the performance of the different state-of-the-art tabular models for recurrence prediction using RNA sequencing data. Models with statistically significant differences with the FT-Transformer are marked with an asterisk ($p < 0.05$)	128
Tableau 8.3	Comparison of different feature fusion approaches. Statistically significant differences ($p < 0.05$) with the proposed pipeline (using intermediate fusion) are marked by an asterisk.	128

LISTE DES FIGURES

Figure 2.1	Exemple d'une image histologique de lame entière (gauche) et d'un bloc de tissu microarray (droite) de la biobanque des MHCR du CHUM	9
Figure 2.2	Les cellules T (CD3+) peuvent se différencier en cellules T cytotoxiques (CD8+) ou cellule T auxiliaires (CD4+) (gauche); Cellule T cytotoxique identifiant les peptides présentés par la cellule tumorale (droite)	12
Figure 2.3	Rôle de CD73 (ecto-5'-nucléotidase) dans l'immunosuppression tumorale [1]	13
Figure 2.4	La voie de régulation négative de PD-1/PD-L1	13
Figure 2.5	Algorithme de prise en charge des MHCR (Centre suisse d'hépatobiliaire et pancréatique, Zurich, Suisse [2])	15
Figure 2.6	Traitement du CPNPC selon le stade de la maladie selon la Société Américaine du Cancer [3]; IO : immunothérapie, chimio : chimiothérapie.	16
Figure 3.1	Le flux de travail radiomique adopté dans la littérature pour la prédiction des cibles thérapeutiques dans les MHCR à l'aide de l'imagerie médicale [4]. Le flux de travail commence généralement par la collecte d'images médicales, suivie de la segmentation de la région d'intérêt et de l'extraction des caractéristiques radiomiques. Ces dernières sont utilisées pour former des modèles d'apprentissage automatique pour la prédiction de biomarqueurs pronostiques, de la réponse au traitement ou des résultats cliniques.	23
Figure 3.2	Les informations multimodales, y compris l'imagerie (TDM, IRM, TEP, histopathologie, etc.), les données omiques (transcriptomique, génomique, etc.) et les données cliniques, sont exploitées dans des approches de modélisation algorithmique. Ces algorithmes offrent des perspectives sur de nombreux résultats de santé différents, tels que le diagnostic du cancer, son pronostic et la découverte de biomarqueurs. Ces informations peuvent être utilisées comme outils d'assistance par les centres de santé et les oncologues pour fournir des soins aux patients plus efficaces et personnalisés. Source [5]	31

Figure 3.3	L'approche multi-omique adoptée par Pan et al. [6] pour la prédiction de la réponse du CPNPC à la chimio-radiothérapie et la SG à 1 an, intégrant des données radiomiques, pathomiques, dosimétriques et cliniques.	34
Figure 3.4	L'entraînement préalable du modèle de fondation : un encodeur convolutionnel profond a été préalablement entraîné en utilisant l'apprentissage contrastif basé sur des volumes avec et sans lésions [7]	42
Figure 4.1	Présentation du cadre méthodologique proposé. (I) Prédiction de cibles thérapeutiques des MHCR, notamment CD73, à partir d'images de TDM de préopératoires ; (II) Proposition d'un pipeline pour prédire à partir des images de TDM si les MHCR présentent une faible infiltration de cellules T CD3+, associée à une expression élevée de CMH-I. (III) Prédiction du niveau d'expression de PD-L1 dans le CPNPC à partir d'images de TDM ; (IV) Proposition d'une approche multio- mique pour la prédiction de la récurrence à deux ans suivant la résection des MHCR.	49
Figure 4.2	Organisation méthodologique de la thèse	50
Figure 5.1	Isotropic spatial resampling was applied on preoperative CT-scan images to segment 160 colorectal liver metastases (CRLM) resected in 122 colorectal cancer (CRC) patients who underwent partial hepatectomy. Matching CRLM, identified by their pathology report block numbers and anatomical description, were included in tissue microarrays for automated quantification of CD73 intra-tumoral expression by immunofluorescence. Radiomic features were extracted from the resulting three-dimensional regions of interest (ROI) (18 first-order statistics, 14 shape features and 75 textural features) and preprocessed. Subsequently, an Attentive Interpretable Tabular Learning (TabNet) model was trained to predict CD73 expression dichotomized as $CD73^{High}$ vs. $CD73^{Low}$ based on the surface area expressing CD73 over the total surface of assessable tissue. The model was then evaluated, and its predictions were interpreted using ROC curve and the Shapley Additive Explanations technique (SHAP). The association between radiomic CD73 (rad-CD73) and oncological survival outcomes was analyzed by Kaplan-Meier and logrank test.	61

Figure 5.2	Performance of TabNet on the hold-out test set. A Comparison of the ROC curves of TabNet with the other baseline models. TabNet outperformed all the trained baseline models as mirrored by the area under the ROC curves (AUC). B Confusion matrix of TabNet for CD73 classification. LR, Logistic Regression ; RF, Random Forest ; SVM, Support Vector Machine ; TabNet, Attentive Interpretable Tabular Learning ; XGBoost, Extreme Gradient Boosting.	62
Figure 5.3	Assessment of the predicted TabNet probabilistic score, rad-CD73. A Violin plot depicting the distribution of rad-CD73 score by CD73 expression level. A statistically significant difference was observed in the radiomic score between the $CD73^{High}$ and $CD73^{Low}$ groups (Wilcoxon signed rank test ; $P < 0.0001$). B Spearman's correlation between rad-CD73 and the actual CD73 expression (the percentage of CD73 positive area per metastasis).	63
Figure 5.4	Decision curve analysis. The red line shows the treat none as $CD73^{High}$ approach while the blue line shows the treat all as $CD73^{High}$ approach. For threshold values greater than 0.08, rad-CD73 had a higher net benefit than both the "treat all" and "treat none" approaches.	64
Figure 5.5	Representative cases with high and low CD73 expression. A A preoperative CT image of a $CD73^{High}$ CRLM (arrow) in the right hemiliver. B A preoperative CT image of a $CD73^{Low}$ CRLM (arrow) in the left hemiliver. C-D Representative tissue microarray cores of the high C and low D CD73 expression of the respective CRLM. Hematoxylin and eosin staining (left) shows the integrity and general architecture of the tissue. Immunofluorescent staining (right) shows cell nucleus (blue, DAPI), CD73 expression (red), and the cancer cells with the pan-cytokeratins epithelial cell marker (green). Scale bars represent 100 μ m.	65
Figure 5.6	TabNet global interpretability analysis using the Shapley Additive Explanations (SHAP) technique. A Summary plot listing radiomics features from top to bottom in a decreasing order of their impact on the model's decision. The top four most significant features were texture-related features. B Variation of the SHAP value of four selected features with respect to the actual feature value. Homogeneous textures were associated with a higher CD73 expression, as mirrored by the effect of the top-ranking features.	73

Figure 5.7	Visualization of TabNet learnable mask from the third decision step. The dependence non uniformity normalized (DNUN) feature, designated as the top 1 feature in SHAP analysis, was selected by the learnable mask among the most salient features for almost all test set instances. Brighter colors indicate higher feature importance.	74
Figure 5.8	Prognostic value of rad-CD73 in colorectal liver metastases (CRLM). A Disease-specific survival according to rad-CD73 and B time-to-recurrence after the initial complete surgical resection of CRLM. In patients with more than one CRLM, the mean rad-CD73 was used to classify patients as low or high. The lower tertile was used as a cut-off value (rad-CD73 > 0.362).	74
Figure 6.1	(a) Attentive Interpretable Tabular Learning (TabNet) model architecture. The attentive transformer outputs a feature selection mask which aim at discarding irrelevant features at a given decision step. The feature transformer processes the selected features through shared and decision step-specific layers. The model's output is obtained by adding a fully connected layer on top of the encoder's output. (b) Proposed Multi-TabNet ensemble. The multi-TabNet ensemble aggregates separately trained TabNet models through voting. Avg, Averaging; BN, Batch Normalization; FC, Fully Connected layer; ReLU, Rectified Linear Unit.	78
Figure 6.2	Representative CT images with (a) $CD3^{Low}MHC-I^{High}$ CLM and (b) $CD3^{Low}MHC-I^{Low}$ CLM. (c) Representative TMA cores with CD3 and MHC-I staining.	79
Figure 6.3	Evolution of Multi-TabNet ensemble performance in terms of AUC, accuracy and balanced accuracy, by ensemble size (n). For $n < 5$, the average performance across all possible combinations of n-sized ensembles is reported.	81
Figure 6.4	Prognostic value of the predicted immune profile : (a) DSS and (b) TTR following the initial CLM resection.	82

- Figure 7.1 Overview of the AI-based PD-L1 classification pipeline. Pre-treatment CT images of patients with advanced NSCLC treated with ICIs are collected and pre-processed. The corresponding PD-L1 expression is associated and dichotomized into two classes : $<50\%$ vs $\geq 50\%$. The ensuing images are segmented and radiomics features (shape, texture and histogram-based features) are extracted from the resulting lesions after the application of a Laplacian of Gaussian filter or a wavelet decomposition to the image. A self-training framework is then adopted in which the deep learning model, the Feature Tokenizer Transformer, is first trained using labeled data and then leveraged to generate pseudolabels for instances with no PD-L1 expression information. Labeled and pseudolabeled data are subsequently combined to fine tune the model and improve its predictive performance. The model is finally validated on the internal test set and a separate cohort of patients who underwent neoadjuvant treatment combining ICIs and chemotherapy. 91
- Figure 7.2 PD-L1 classification results of the predictive pipeline. (a) Receiver operating characteristics curve on the primary testing dataset (green) and the external validation cohort consisting of patients who underwent neoadjuvant treatment (orange). (b) Spider plot showing the performance metrics of the predictive pipeline on the internal and external validation cohorts. AUC ; area under the receiver operating characteristic curve. (c-d) Evaluation of the probabilistic rad-PDL1 score obtained using the predictive pipeline. The box plots present the distribution of rad-PDL1 by PD-L1 expression level (high vs low). The data indicated a statistically significant difference in rad-PDL1 between the PD-L1High and PD-L1Low groups in both (c) the primary testing cohort (Mann-Whitney U test ; $p < 0.0001$) and (d) the external neoadjuvant cohort ($p = 0.0298$). 101

Figure 7.3	Interpretability analysis of the predictive pipeline. (a) Summary plot showing the top features obtained using the Shapley Additive Explanations (SHAP) technique, according to their average Shapley values, which reflect their impact on the model's output. The top six features were found to be textural features, which highlights the importance of texture-related properties of NSCLC lesions in classifying NSCLC. (b) Variation in SHAP values of the top two radiomics features as a function of their actual values. The results show that heterogeneous NSCLC lesions were associated with a high PD-L1 expression.	102
Figure 7.4	Representative cases with low and high PD-L1 expression. (a) A pretreatment CT image of a PD-L1High NSCLC (arrow) in the right lung. (b) A pretreatment CT image of a PD-L1Low NSCLC (arrow) in the left lung.	103
Figure 7.5	Decision curve analysis by modality on (a) the primary and (b) the external (neoadjuvant chemotherapy) testing cohorts. The net benefit is computed for each threshold probability, representing the acceptable risk beyond which acting (treatment, further testing, etc.) is advised. The dashed orange curve represents the approach where all patients are considered to have a high expression of PD-L1 and the black x-axis represents the approach where no patient is considered to have a high PD-L1 expression. As shown in the plots, the net benefit of the proposed pipeline (green curve) is higher than the "treat all patients as high PD-L1" (orange curve) and "treat no patient as high PD-L1" (black line) approaches in both the primary and external (neoadjuvant) testing cohorts.	103
Figure 7.6	Prognostic significance of the predicted rad-PDL1 score in (a) the primary testing cohort and (b) the external testing cohort composed of patients receiving neoadjuvant treatment. The progression-free survival was significantly shorter in the rad-PDL1Low group compared to the rad-PDL1High in the primary testing cohort ($p = 0.02$).	104

Figure 8.1	Overall workflow of the proposed approach which combines transcriptomics, imaging and clinical data to predict the early recurrence of resected CRLM (two years after complete resection). RNA sequencing and CT images were first processed separately and the ensuing transcriptomics and imaging scores were leveraged along with clinical features (Fong CRS, age and gender) to predict the likelihood of disease recurrence. The RNA sequencing module comprised a Generative Adversarial Imputation Network (GAIN) which performed generative imputation of missing gene expression counts, followed by a Conditional Tabular Generative Adversarial Network (CTGAN) which synthesized RNA sequencing instances that were combined with the original samples and used to train a Feature Tokenizer Transformer (FT-Transformer). The CT component incorporated a foundation model leveraged as a feature extractor, along with a radiomics feature extraction module. The resulting features were used to train cascaded random forest models. The scores of the respective models were combined in a multilayer perceptron to predict the 2-year recurrence.	113
Figure 8.2	Generative imputation of missing gene expression counts. The GAIN model [8] was leveraged to impute missing gene expressions. A generator synthesized imputations of missing gene expression counts and a discriminator was trained to predict whether expression counts were real or synthetic imputations proposed by the generator. The imputed gene expression matrix was subsequently deployed for recurrence prediction.	114
Figure 8.3	Architecture of the FT-Transformer model trained on the synthetically augmented gene expression dataset for the prediction of 2-year recurrence. Features are first embedded using the feature tokenizer module and the embeddings are fed into the Transformer encoder including self-attention and fully-connected layers.	116
Figure 8.4	Receiver operating characteristic (ROC) curves by modality using the proposed deep learning pipeline for the prediction of 2-year recurrence in colorectal liver metastasis surgery cases ($n = 129$, 84 events). Figures show the AUC with the corresponding standard deviation across cross-validation fold ($k = 5$). The diagonal dashed line represents a performance equivalent to random chance.	121

Figure 8.5	Comparison of the areas under the receiver operating characteristic curves by modality against the baseline clinical-only model. Box plots show the Q1, median and Q3, with box whisker's representing the 1.5 interquartile range. Statistically significant differences ($p < 0.05$) are highlighted with an asterisk. The dashed horizontal line depicts the median AUC of the clinical-only model. The brown dots indicate the mean AUC and the brown errors bars the standard deviations. . . .	122
Figure 8.6	Ablation study results of the imaging module. The boxes cover the interquartile range, the whiskers the 1.5 interquartile range and the dots indicate the outliers. The dashed and straight lines represent the mean and median performance metrics, respectively.	124
Figure 8.7	Comparison of the performance of different machine learning models for recurrence prediction using radiomics features (left) and features derived from the foundation model (right). The boxes indicate the interquartile range, the whiskers the 1.5 interquartile range and the dots represent the outliers.	125
Figure 8.8	Kaplan-Meier curves obtained with the models trained using different combinations of modalities : RNA sequencing; CT features; Clinical features; RNA sequencing + CT features; RNA sequencing + clinical features; Clinical + CT features; RNA sequencing + clinical + CT features (proposed pipeline). The vertical dashed line highlights the stratification at 2 years post-hepatectomy.	126
Figure 8.9	Decision curve analysis by modality. The gray curve represents the approach where all patients are considered to have a high recurrence risk and the black x-axis represents the approach where no patient is considered to have a high recurrence risk. The blue curve represents the net benefit of the proposed multi-omics pipeline.	127
Figure 8.10	Volcano plot of the differential gene expression results. The gray dots represent the genes that are not differentially expressed, the green dots denote the differentially expressed genes, and the brown dots represent six chosen genes among the top differentially expressed. The fold change of a given gene is computed as the ratio of the its expression in the high versus the low risk groups. FDR : False Discovery Rate. The dashed horizontal and vertical lines mark the thresholds for considering a gene as differentially expressed.	129
Figure 8.11	SHAP analysis of the clinical features.	130

Figure 8.12	SHAP analysis of the radiomics features.	133
Figure 8.13	Correlation matrix between top differentially expressed genes, CT radiomics features and the Fong CRS. While some RNA sequencing features are inter-correlated, different modalities did not exhibit a high correlation between each other. DNUN-HLL, DependenceNonUniformityNormalized-HighLowLow ; DNUN-LHL, DependenceNonUniformityNormalized-LowHighLow ; HGLZE, HighGrayLevelZoneEmphasis ; GLNUN, GrayLevelNonUniformityNormalized ; SZNUN, SizeZoneNonUniformityNormalized. The clinical score is presented is blue, genes in brown and imaging features in green ...	134
Figure B.1	Distribution du score probabiliste TabNet prédit, rad-CD73, parmi les patients. Une distribution trimodale a été observée pour rad-CD73. Le logiciel X-tile a retourné un point seuil basé de 0,383 pour cette distribution (proche de la valeur du premier tertile de 0,362).	195

LISTE DES SIGLES ET ABRÉVIATIONS

ACE	Antigène carcinoembryonnaire
AMT	Attention multi-têtes
AUC	Aire sous la courbe de caractéristique de fonctionnement du récepteur
CCR	Cancer colorectal
CD3	Cluster de différenciation 3
CD4+	Cellules T auxiliaires
CD8+	Cellules T cytotoxiques
CHC	Carcinome hépatocellulaire
CHM	Centre hospitalier de l'Université de Montréal
CMH	Complexe majeur d'histocompatibilité
CPNPC	Cancer du poumon non à petites cellules
CTGAN	Réseau antagoniste génératif tabulaire conditionnel
EGFR	Récepteur du facteur de croissance épidermique
FDA	Food and Drug Administration
FRF	Futur Reste du Foie
FT-Transformer	Feature Tokenizer Transformer
GAIN	Réseau d'imputation antagoniste générative
GAN	Réseau antagoniste génératif
ICI	Inhibiteur de points de contrôle immunitaires
IHC	Immunohistochimie
IMS	Instabilité Microsatellitaire
IRM	Imagerie par résonance magnétique
LASSO	Opérateur de sélection et de rétrécissement absolu minimum
MHCR	Métastases hépatiques colorectales
MIT	Microenvironnement immunitaire tumoral
PD-1	Protéine 1 de la mort cellulaire programmée
PD-L1	Ligand de mort programmé 1
RECIST	Critères d'évaluation de la réponse dans les tumeurs solides
RF	Forêt aléatoire
RNC	Réseaux neuronaux convolutionnels
RPKM	Lectures par kilobase de transcrit par million de lectures mappées
SG	Survie globale
SMOTE	Technique de suréchantillonnage des minorités synthétiques

SRC	Score de risque clinique
SSC	Survie spécifique au cancer
SVM	Machine à vecteurs de support
TabNet	Apprentissage Attentif Interprétable Tabulaire
TAC	Thérapie adoptive par cellules T
TDM	Tomodensitométrie
TIGIT	Lymphocytes T avec domaines Immunoglobuline et ITIM
TTR	Temps jusqu'à la récurrence
ViT	Transformateur de vision
XGB	Classificateur de renforcement du gradient extrême

LISTE DES ANNEXES

Annexe A	Paramètres d'extraction de caractéristiques radiomiques	185
Annexe B	Traitement des données et entraînement des modèles	192

CHAPITRE 1 INTRODUCTION

Le cancer colorectal (CCR) et le cancer du poumon figurent parmi les types de cancer les plus mortels au Canada, le cancer du poumon occupant la première place en termes de mortalité liée au cancer et le CCR la deuxième place chez les hommes canadiens et la troisième chez les femmes canadiennes [9]. En 2024, le cancer du poumon était le type de cancer le plus diagnostiqué au Canada, tandis que le CCR était le quatrième type de cancer le plus diagnostiqué [10]. Les projections mondiales de 2040 de l'Organisation mondiale de la santé indiquent que l'incidence du CCR augmentera pour atteindre 3,2 millions de nouveaux cas par an et entraînera 1,6 million de décès chaque année [11]. De plus, moins de 10% des patients atteints de CCR avec métastases synchrones survivent plus de 5 ans [12].

Parmi les métastases du CCR, les métastases hépatiques colorectal (MHCR) se développent fréquemment en raison du rôle majeur du foie dans le filtrage du sang intestinal. Les MHCR sont réputés pour leur agressivité. En fait, plus de 99% des patients atteints de MHCR non traitées succombent à la maladie dans les 5 ans. Pour les MHCR opérables, la chirurgie est considérée comme la première option de traitement. Cependant, même après une hépatectomie, la survie globale (SG) à 5 ans ne dépasse pas 45% [13]. Cela met en évidence les limites des systèmes de stratification des risques actuels qui reposent principalement sur des variables cliniques.

Le cancer du poumon non à petites cellules (CPNPC) est le type de cancer du poumon le plus courant et la survie des patients varie en fonction du sous-type de CPNPC et du stade de la maladie. L'avènement des inhibiteurs de points de contrôle immunitaires (ICI) a modifié la gestion du CPNPC, entraînant une amélioration significative du taux de survie des patients depuis l'introduction de l'immunothérapie. La SG à 5 ans des patients atteints de CPNPC métastatiques est passée de 9% à l'ère pré-immunothérapie à 14,3% après l'approbation de l'immunothérapie [14].

Néanmoins, les résultats cliniques de ces cancers restent largement hétérogènes. En fait, plus de 70% des patients atteints de MHCR connaissent une récurrence dans les deux ans suivant la résection chirurgicale [15]. De même, le taux de réponse objectif du CPNPC avancé aux ICI est inférieur à 42,9% [16]. Cette incertitude pronostique souligne le besoin crucial de résoudre un problème plus large en oncologie : le décalage entre la biologie tumorale et la prise de décision clinique. La médecine personnalisée vise à combler cet écart, mais sa mise en œuvre dépend des biomarqueurs permettant de décoder le microenvironnement tumoral. Les biomarqueurs pronostiques conventionnels du cancer reposent sur l'évaluation des échantillons

histopathologiques, qui sont confrontés à des problèmes d’hétérogénéité spatiale, de risques procéduraux et de l’absence de surveillance dynamique.

Les avancées en radiomique offrent des alternatives non invasives. Des études récentes ont révélé que les caractéristiques d’imagerie sont corrélées avec les phénotypes moléculaires des tumeurs et le microenvironnement immunitaire, montrant ainsi que les biomarqueurs basés sur l’imagerie pourraient être utilisés comme substituts non invasifs, fournissant des informations sur la biologie tumorale. L’intégration de l’intelligence artificielle et de la radiomique en oncologie a ouvert des opportunités sans précédent pour interroger la biologie du cancer. Ces outils promettent particulièrement de permettre la mise au point de proxies d’imagerie non invasifs qui pourraient compléter les évaluations basées sur les tissus. Néanmoins, le domaine manque encore de biomarqueurs d’imagerie fiables prédisant les résultats cliniques des patients, en particulier dans le cas des MHCR et du CPNPC. De plus, les études récentes en multi-omique intégrant des données médicales provenant de multiples sources suggèrent que la combinaison d’images médicales avec des données génétiques, cliniques et histologiques pourrait prédire de manière fiable la biologie tumorale grâce à des outils d’apprentissage profond de pointe.

Cependant, des lacunes critiques persistent, en particulier dans la pronostication des MHCR et du CPNPC. Premièrement, les modèles d’apprentissage profond sont souvent utilisés dans des pipelines à modalité unique : les modèles pronostiques pour les MHCR restent insuffisamment explorés, en particulier en ce qui concerne l’intégration des textures d’imagerie pré-opératoires, comme celles des tomodensitogrammes, avec les facteurs transcriptomiques de récurrence et les caractéristiques cliniques. Deuxièmement, les problèmes de généralisation des modèles sont fréquemment négligés dans les études existantes : l’effet des différents stades de CPNPC et des régimes de traitement sur la performance des prédictors de biomarqueurs de réponse aux ICI est souvent ignoré. Troisièmement, l’interprétabilité biologique des modèles d’apprentissage profond reste insaisissable, car ces modèles sont développés sans une compréhension complète de leurs prédictions, ce qui entrave leur adoption clinique. Quatrièmement, les artefacts techniques, tels que la grande dimensionnalité des données transcriptomiques et les transcrits non détectés en raison de la faible pureté tumorale ou de la profondeur de séquençage, affecte les analyses en aval.

1.1 Contributions

Le thème global de la thèse est l’amélioration du pronostic des MHCR et du CPNPC grâce à la prédiction non invasive de biomarqueurs et à l’intégration multi-omique, basées sur l’apprentissage automatique.

La première contribution établit un pipeline radiomique basé sur la tomodensitométrie (TDM) qui classe l'expression de l'ecto-5'-nucléotidase, ou CD73, dans le MHCR de manière non invasive. La surexpression de CD73 dans le MHCR favorise les environnements immunosuppresseurs et la croissance métastatique. En entraînant des modèles d'apprentissage tabulaire profond sur des caractéristiques radiomiques extraites de tomodensitogrammes avant traitement, ce travail associe des caractéristiques d'imagerie et des métriques d'hétérogénéité de texture à des scores d'immunohistochimie (IHC). Le biomarqueur basé sur la radiomique, rad-CD73, peut identifier un sous-groupe de patients caractérisés par une récurrence plus précoce, une survie plus courte et une expression intratumorale accrue de CD73. L'utilité clinique potentielle de rad-CD73 justifie une étude prospective afin de valider son rôle en tant qu'outil d'assistance pour prédire la réponse aux immunothérapies ciblant la voie de l'adénosine et le pronostic des patients atteints d'une MHCR réséqué.

La deuxième contribution est l'identification d'un score immunitaire non invasif du MHCR basé sur la densité d'infiltration des lymphocytes T CD3+ combinée à l'expression du complexe majeur d'histocompatibilité I (CMH-I). Comme indiqué précédemment, une faible densité d'infiltration des lymphocytes T CD3+ associée à une forte expression du CMH-I ($CD3^{Low}CMH^{High}$) est associée à un pronostic plus sombre et à une réponse plus faible à la chimiothérapie systémique. En proposant une approche d'apprentissage profond par ensemble appelée Multi-TabNet, nous mettons en avant un score immunologique non invasif basé sur des images TDM préopératoires qui est associé aux résultats cliniques des patients atteints de MHCR. Lorsqu'il est validé de manière prospective sur de grands ensembles de données multi-institutionnels, ce score peut également être utilisé pour guider de manière dynamique les thérapies adjuvantes et la fréquence des suivis après l'opération.

Troisièmement, la thèse développe un pipeline d'apprentissage profond qui met en correspondance les caractéristiques des images de TDM avant traitement avec les niveaux d'expression du ligand de mort programmé 1 (PD-L1), qui est actuellement le seul biomarqueur approuvé par la Food and Drug Administration (FDA) pour guider le traitement par ICI dans le CPNPC avancé. Nous avons exploité des techniques avancées, notamment l'auto-apprentissage et les transformeurs. Nous avons validé le modèle dans une cohorte indépendante de patients atteints de CPNPC à un stade limité traités par ICI néoadjuvant et chimiothérapie, la norme de soins actuelle pour les patients éligibles atteints de CPNPC résécable. De plus, nous avons fourni une analyse d'interprétabilité du modèle, montrant que les scores radiomiques Rad-PD-L1 générés par le modèle sont associés avec les résultats cliniques. Ces résultats soulignent le potentiel de Rad-PD-L1 en tant que biomarqueur non invasif et robuste pour prédire la réponse à l'ICI et éclairer les stratégies de traitement du CPNPC. Notre pipeline démontre une grande précision prédictive et une bonne généralisa-

bilité. À notre connaissance, il s’agit du premier travail à démontrer la généralisabilité d’un score radiomique de PD-L1 non invasif à différents stades de la maladie, en particulier dans le contexte émergent de la chimiothérapie-ICI.

Quatrièmement, nous abordons les défis liés à la manipulation des ensembles de données transcriptomiques, à savoir les transcrits d’ARN non détectés et la grande dimensionnalité des ensembles de données d’expression génique. En tirant parti de l’apprentissage génératif tabulaire, nous fournissons des imputations plausibles des comptes d’expression génique non détectés et validons leurs avantages sur le sous-typage histologique du CPNPC en tant que tâche en aval. Nous tirons également parti de l’apprentissage génératif tabulaire conditionnel pour générer des données transcriptomiques synthétiques permettant de surmonter le problème de la taille limitée des échantillons étant donné la forte dimensionnalité des ensembles de données d’expression génique. Nous garantissons également l’interprétabilité de nos approches en analysant les gènes les plus influents pour les tâches en aval. Ces méthodologies améliorent collectivement la robustesse et l’applicabilité des données transcriptomiques, en particulier pour le pronostic du cancer.

Cinquièmement, la thèse est pionnière dans le domaine des plateformes de pronostic multi-omiques, en fusionnant la radiomique préopératoire par TDM, les données de séquençage de l’ARN et les variables cliniques pour évaluer le risque de récurrence précoce des MHCR dans les deux ans suivant l’hépatectomie. Le pipeline proposé s’appuie sur les progrès récents des modèles de fondation, des transformeurs et des techniques d’apprentissage génératif pour créer une approche globale de la prédiction des récurrences des métastases hépatiques d’origine colorectale. Un élément clé de ce pipeline est le module de prédiction des récurrences basé sur la TDM, qui utilise des modèles en cascade formés à la fois sur des représentations dérivées d’un modèle de fondation et sur des caractéristiques radiomiques interprétables, fournissant une caractérisation riche et polyvalente du microenvironnement tumoral. L’étude met également l’accent sur l’interprétabilité du modèle grâce à des analyses post-hoc. Le pipeline multi-omique a notamment démontré des performances supérieures à celles d’un score pronostique cliniquement établi, le score clinique de Fong, soulignant son utilité clinique potentielle.

1.2 Structure de la thèse

Cette thèse se compose de 10 chapitres. Le chapitre 1 est une introduction générale à l’importance des outils de pronostic du cancer multi-omiques. Le chapitre 2 présente les informations cliniques de base utiles pour mieux comprendre les méthodes développées tout au long de cette thèse et une introduction à la radiomique. Le chapitre 3 est une revue de la littérature couvrant les études et les techniques d’apprentissage profond exploitant les différentes

données omiques dans la recherche sur le cancer, à savoir l'imagerie, la transcriptomique et les données cliniques, tout en mettant l'accent sur la prédiction du pronostic du cancer. La problématique de recherche, les objectifs et l'hypothèse, ainsi que la méthodologie générale, sont présentés au chapitre 4.

Les principales conclusions de cette thèse sont présentées aux chapitres 5 à 8, qui sont basés sur des publications évaluées par des pairs. Le chapitre 5 présente le premier article intitulé : « Radiomics using computed tomography to predict CD73 expression and prognosis of colorectal cancer liver metastases » publié dans le Journal of Translational Medicine. Le chapitre 6 décrit le deuxième article évalué par des pairs publié dans le 2023 IEEE 20th International Symposium on Biomedical Imaging et intitulé « Ensemble TabNet predicting a T-cell/CMH-I-based immune profile biomarker for colorectal liver metastases from CT images ». Le troisième article est présenté au chapitre 7, intitulé « Multi-institutional CT scan-based radiomics for predicting tumor PD-L1 expression in patients with advanced and limited non-small cell lung cancer », soumis à la revue OncoImmunology. Le chapitre 8 présente le quatrième article soumis au Journal of Biomedical and Health Informatics et intitulé « Multi-omics fusion network for prediction of early recurrence in colorectal liver metastases ».

Le chapitre 9 apporte une discussion globale sur les développements introduits tout au long de cette thèse. Enfin, le chapitre 10 conclut par un résumé des travaux de cette thèse, donne un aperçu des limites générales et présente des suggestions pour les travaux futurs.

CHAPITRE 2 CONTEXTE

Dans ce chapitre, un résumé des connaissances de base liées au sujet de recherche est présenté. La section commence par une introduction du MHCR et du CPNPC ainsi que des techniques diagnostiques associées. Ensuite, les biomarqueurs pronostiques les plus pertinents pour ces deux types de cancer sont présentés. Puis, les traitements standards du MHCR et du CPNPC sont décrits. Enfin, un aperçu technique du concept de radiomique est fourni.

2.1 Les métastases hépatiques colorectales

Le CCR en stade précoce passe souvent inaperçu, ce qui entraîne une progression de la maladie et la formation de métastases [13]. En particulier, les MHCR se développent fréquemment comme premier site de métastase. Il est rapporté que moins d'un tiers des patients atteints de MHCR sont de bons candidats pour une chirurgie curative, principalement en raison de la charge tumorale importante et de l'insuffisance du reste hépatique post-chirurgical [17]. De plus, la récurrence de la maladie survient chez environ les deux tiers des patients après hépatectomie [2]. Cela a motivé les chercheurs à explorer les options de traitement des MHCR et à étudier de nouveaux traitements pour améliorer les résultats des patients atteints de la maladie.

2.1.1 Le score de risque clinique

Plusieurs scores pronostiques ont été proposés dans la littérature pour stratifier les patients atteints de MHCR en fonction de leurs résultats de survie, tels que les scores de réponse pathologique à la chimiothérapie préopératoire [18, 19], le stade TNM et le score de risque clinique (SRC), proposé par Fong et al. [20]. Ce dernier est calculé en ajoutant un point chaque fois que l'une des cinq conditions suivantes est remplie : (1) l'intervalle sans maladie entre le diagnostic de la tumeur primaire et les métastases hépatiques est inférieur à un an ; (2) le nombre de métastases est supérieur à 1 ; (3) le niveau préopératoire de l'antigène carcinoembryonnaire (ACE) est supérieur à 200 ng/mL ; (4) la plus grande métastase mesure plus de 5 cm ; (5) la tumeur primaire est positive pour les ganglions lymphatiques.

2.2 Le cancer pulmonaire non à petites cellules

Selon la Société Canadienne du Cancer, le cancer du poumon est la principale cause de décès par cancer chez les hommes et les femmes au Canada [10]. En 2024, la maladie représentait

environ 13 % de tous les cas de cancer et a causé 20 700 décès au Canada. Les statistiques montrent que les taux de survie spécifiques au cancer (SSC) sur cinq ans sont de 19 % chez les hommes canadiens et de 26 % chez les femmes canadiennes. Bien que le CPNPC progresse plus lentement que le cancer du poumon à petites cellules, il représente 85 % de tous les cas de cancer du poumon [21].

Les principaux sous-types histologiques du CPNPC sont l'adénocarcinome, le carcinome épidermoïde et le carcinome à grandes cellules, représentant respectivement environ 40 %, 25 % et 10 % de tous les cas de cancer du poumon [22]. De plus, certains sous-ensembles d'adénocarcinomes ont été associés à des variants spécifiques dans les gènes qui codent pour les composants du récepteur du facteur de croissance épidermique (EGFR). Des études suggèrent également que la survie diffère entre les deux sous-types les plus répandus du CPNPC. Néanmoins, le pronostic est largement influencé par le stade auquel le diagnostic est effectué, les stades plus précoces étant associés à de meilleurs résultats cliniques. Le dépistage est donc essentiel pour les personnes à haut risque, car le CPNPC à un stade précoce est souvent asymptomatique. Les facteurs de risque incluent l'âge et le tabagisme [23].

2.3 Outils diagnostiques et pronostiques

2.3.1 Imagerie médicale

L'imagerie par TDM est largement utilisée dans le diagnostic du cancer, en particulier pour les MHCR et les CPNPC [24]. En effet, les scanners abdominaux permettent de détecter et de délimiter les lésions hépatiques, car ils fournissent des images tridimensionnelles avec une bonne résolution et une couverture complète du foie et des organes adjacents. De plus, l'imagerie par TDM joue un rôle essentiel dans la stadification du CCR, car elle permet aux radiologues d'évaluer l'étendue de la propagation de la maladie. De plus, les scanners thoraciques jouent un rôle clé dans le diagnostic initial et la surveillance du CPNPC. D'autres techniques d'imagerie, telles que l'imagerie par résonance magnétique (IRM) et la tomographie par émission de positons (TEP), ont également été utilisées dans le diagnostic des MHCR et des CPNPC [25].

D'un point de vue technique, l'acquisition d'images TDM consiste à générer des images en coupe transversale en quantifiant l'atténuation des rayons X lorsqu'ils traversent le patient à partir d'orientations multiples [26]. Un tube à rayons X émet un faisceau dirigé vers le patient, couvrant une fine section transversale du corps. Lorsque le faisceau de rayons X traverse le corps, il est atténué puis détecté par un réseau de détecteurs. Les détecteurs les plus couramment utilisés dans les TDM sont des ensembles scintillateur-photodiode à l'état

solide. Lorsque les rayons X interagissent avec le scintillateur, ils produisent de la lumière visible, qui est ensuite convertie en courant électrique par la photodiode. Ce courant est proportionnel à l'intensité du faisceau de rayons X transmis. Le signal analogique résultant est numérisé à l'aide de convertisseurs analogiques-numériques.

Le tube à rayons X et le détecteur sont montés sur un support rotatif qui tourne autour du patient, effectuant généralement une rotation complète de 360°. Cette configuration permet l'acquisition de données de projection sous plusieurs angles, ce qui est essentiel pour la reconstruction de l'image.

Le TDM génère une image numérique composée de voxels, chacun représentant un élément de volume discret dans le tissu du patient. Lorsque le faisceau de rayons X incident d'intensité initiale I_0 traverse le corps, il est atténué par chaque voxel i , caractérisé par un coefficient d'atténuation μ_i , le long d'une trajectoire de longueur x . L'intensité I du faisceau transmis reçu par le détecteur peut être exprimée mathématiquement comme suit :

$$I = I_0 \cdot \exp \left(-x \cdot \sum_{i=1}^n \mu_i \right)$$

où n est le nombre de voxels le long du trajet du faisceau et x_i est la distance parcourue dans le voxel i .

Pour reconstruire l'image, le coefficient d'atténuation μ_i pour chaque voxel doit être estimé. Cette estimation est généralement réalisée à l'aide de l'algorithme de rétroprojection filtrée, qui combine les données provenant de plusieurs angles de projection.

Avant la visualisation, les coefficients d'atténuation reconstruits sont convertis en nombres CT standardisés, appelés unités Hounsfield (HU), en normalisant par rapport au coefficient d'atténuation de l'eau (μ_w) à l'aide de la relation suivante :

$$\text{HU} = 1000 \cdot \left(\frac{\mu_i - \mu_w}{\mu_w} \right)$$

Par conséquent, l'eau a une HU de 0, l'air d'environ -1000 et l'os cortical dense d'environ +1000.

2.3.2 Histopathologie

L'histopathologie est l'évaluation des modifications des tissus accompagnant une maladie, telles qu'observées sur des images microscopiques [27]. Les images histologiques de lames entières (Figure 2.1, à gauche) sont des images à haute résolution permettant la caractérisation

des tissus et la confirmation de la maladie. De plus, l'histopathologie comprend l'évaluation de plusieurs caractéristiques pathologiques [28], telles que la positivité des marges tumorales, la réponse à la chimiothérapie à travers la classification de la régression tumorale dans le cas des MHCR [18], et les dommages liés à l'administration de chimiothérapie. De plus, les lames histologiques permettent l'évaluation de paramètres supplémentaires tels que le pattern de croissance histopathologique [29].

En outre, le tissu microarray est un outil histopathologique récent dans lequel plusieurs petits échantillons de tissus, provenant de différentes lésions, sont disposés dans un seul bloc sous forme de noyaux circulaires [30] (Figure 2.1, à droite). Cette technique permet l'évaluation simultanée des caractéristiques génomiques et protéomiques des tissus, telles que l'évaluation du niveau d'expression de marqueurs immunologiques par IHC.

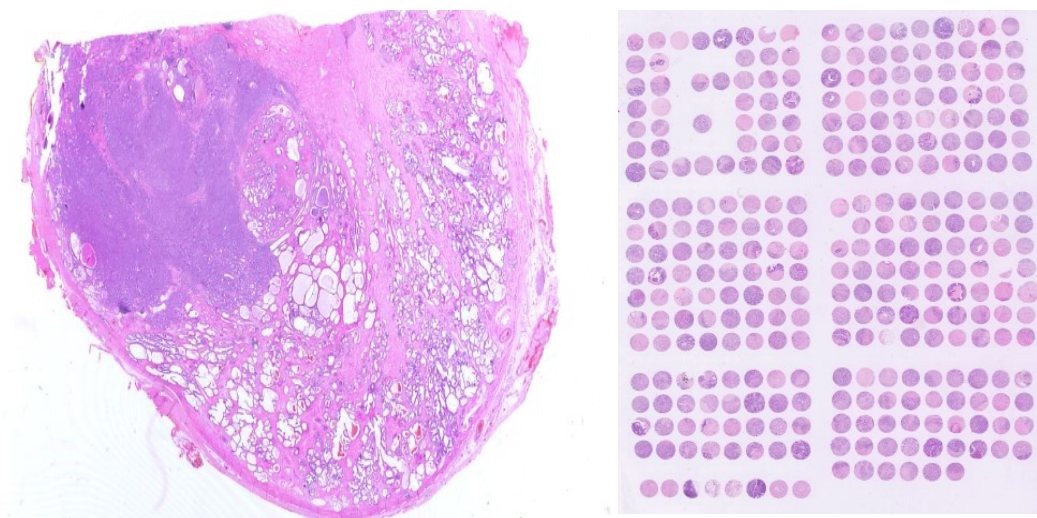


FIGURE 2.1 Exemple d'une image histologique de lame entière (gauche) et d'un bloc de tissu microarray (droite) de la biobanque des MHCR du CHUM

2.3.3 Séquençage d'ARN

Bien que moins utilisé dans la pratique clinique courante, le séquençage de l'ARN est une technique récente à haut débit qui peut jouer un rôle considérable dans le diagnostic et le pronostic du cancer. Cette technologie permet de cartographier l'ensemble du transcriptome et peut ainsi fournir des biomarqueurs pertinents. En effet, le séquençage de l'ARN permet d'étudier diverses modifications affectant les cellules tumorales, telles que l'expression génique et les mutations [31].

Techniquement, le séquençage de l'ARN comprend trois étapes : (1) l'ARN est d'abord extrait des échantillons de tissus et purifié, (2) une librairie de séquençage de nouvelle génération (next-generation sequencing) est ensuite préparée en convertissant l'ARN en ADN complémentaire qui est plus stable et compatible avec les plateformes de séquençage, et (3) des lectures de séquences brutes sont produites. Ensuite, les lectures sont alignées sur un génome de référence. Cette étape détermine l'origine de chaque fragment d'ARN, en associant les séquences à leurs gènes correspondants. Les comptes bruts de gènes sont par suite générés, reflétant l'expression des transcrits. La normalisation des comptes bruts est nécessaire pour prendre en compte la taille de la librairie et la profondeur de séquençage. Différentes méthodes de normalisation existent telles que la médiane des ratios et la méthode de lectures par kilobase de transcrit par million de lectures mappées (RPKM). Lorsqu'un effet de lot existe en raison de différences techniques lors de la mesure ou le traitement des données, une correction de lot devrait être effectuée en utilisant des méthodes appropriées telles que ComBat [32].

La matrice d'expression génique prétraitée peut être utilisée pour des analyses ultérieures, telles que l'analyse différentielle de l'expression génique ou l'analyse d'enrichissement des ensembles de gènes. L'analyse différentielle de l'expression génique permet d'identifier les gènes qui présentent des différences d'expression significatives entre deux groupes. L'analyse d'enrichissement des ensembles de gènes, basée sur les résultats de l'analyse différentielle, détermine ensuite les voies biologiques associées à ces gènes. Plus récemment, des algorithmes d'apprentissage automatique ont été utilisés pour prédire les différences dans l'expression des gènes en utilisant des modèles de classification [33].

Le séquençage de l'ARN est utilisé dans divers domaines, y compris la recherche sur le cancer pour découvrir les mécanismes de la maladie, la médecine personnalisée pour prédire les réponses thérapeutiques, ainsi que dans les maladies infectieuses et la médecine de la transplantation [34]. Cependant, plusieurs défis se posent lors de l'analyse des données de séquençage de l'ARN, tels que (1) la gestion des volumes importants de données générés par chaque séquenceur, nécessitant des ressources informatiques considérables en raison des millions de lectures produites ; (2) les variations techniques survenant lors du traitement des données dues à des protocoles de librairies ou des conditions expérimentales différentes, entraînant des effets de lot ; et (3) les transcrits géniques non détectés en raison de pertes lors de la préparation de la librairie ou d'une profondeur de séquençage insuffisante [35].

2.4 Biomarqueurs immunitaires du cancer

Dans cette section, nous décrivons les biomarqueurs immunologiques et génotypiques pertinents qui ont été associés au pronostic du MHCR. Dans le cas des biomarqueurs immunologiques, les lymphocytes infiltrant les tumeurs et l'immunoscore ont récemment fait l'objet d'une attention accrue en raison de leur association avec l'évolution des patients. En outre, d'autres cibles thérapeutiques tels que les protéines de contrôle immunitaire ont également été associés à la réponse immunitaire au cancer.

2.4.1 Infiltrats immunitaires et immunoscore

Le nombre et la répartition des lymphocytes ont été corrélés avec le pronostic des patients et la réponse à l'immunothérapie [36, 37]. Les marqueurs de surface tels que le cluster de différenciation 3 (CD3), le cluster de différenciation 8 (CD8) et le CMH-I sont des exemples qui ont été étudiés. En fait, les lymphocytes T, ou cellules CD3+, peuvent être différenciées en cellules T cytotoxiques (CD8+) ou en cellules T auxiliaires (CD4+) (Figure 2.2). En particulier, le CD8+ a été rapporté comme un marqueur dans le CPNPC [38, 39].

L'immunoscore qui implique la quantification des cellules CD3+ et CD8+ infiltrées dans la tumeur et la marge invasive dans le cas des MHCR, a également été évalué dans ce contexte [40, 41]. Baldin et al. [29] ont trouvé qu'une faible valeur de l'immunoscore était fortement associée au temps jusqu'à la récurrence (TTR) et à la SG. La même étude a rapporté que l'immunoscore était le principal facteur contribuant au TTR et à la SG par rapport à toutes les autres caractéristiques histopathologiques et génotypiques évaluées dans l'étude. Van Den Eynde et al. [12] ont mené une autre étude confirmant l'association de l'immunoscore élevé avec la survie sans maladie.

De plus, il a été décrit que la régulation négative de CMH-I permet aux cellules tumorales de « se cacher » du système immunitaire. L'identification des peptides CMH-I est la première étape pour l'activation des cellules CD8+. La combinaison des niveaux d'expression de CD3+ et de CMH-I a également été rapportée comme un biomarqueur pronostique indépendant dans les MHCR [42].

2.4.2 Cibles thérapeutiques

L'ecto-5'-nucléotidase, ou CD73

Les caractéristiques immunitaires des MHCR ont été associées à la récurrence et à la survie après résection, indépendamment des caractéristiques cliniques, pathologiques et génomiques

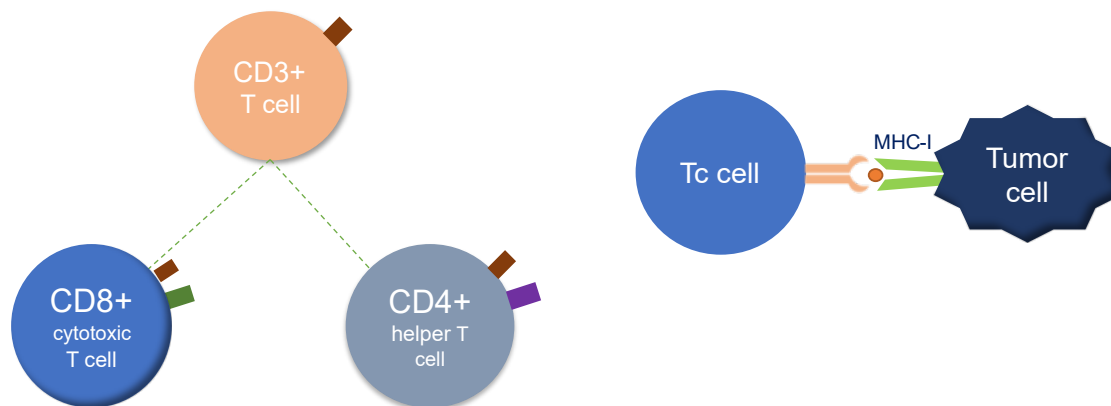


FIGURE 2.2 Les cellules T (CD3+) peuvent se différencier en cellules T cytotoxiques (CD8+) ou cellule T auxiliaires (CD4+) (gauche) ; Cellule T cytotoxique identifiant les peptides présentés par la cellule tumorale (droite)

des tumeurs [43]. L'adénosine extracellulaire dans le microenvironnement immunitaire tumoral (MIT) émerge comme un mécanisme immunosuppresseur particulièrement pertinent au sein du microenvironnement des MHCR [44]. L'ecto-5'-nucléotidase, également connue sous le nom de CD73, exprimée à la fois par les cellules cancéreuses et stromales au sein des MHCR, fonctionne comme une enzyme limitante qui favorise la dégradation de l'adénosine monophosphate extracellulaire dérivée de l'ATP en adénosine immunosuppressive [1]. Cette adénosine se lie à ses quatre récepteurs distincts, dont A2A et A2B exprimés sur les cellules T, les cellules tueuses naturelles (Natural Killer) et d'autres cellules immunitaires, inhibant ainsi leurs fonctions cytotoxiques anti-tumorales [45]. L'activation du récepteur A2A par l'adénosine entraîne l'inhibition de la signalisation du récepteur des cellules T [46], et l'activation du récepteur A2B conduit à la suppression de l'élimination des cellules tumorales médiée par le système immunitaire (Figure 2.3). Une expression intratumorale élevée de CD73 est fortement associée à un mauvais pronostic tant dans le CCR [47] primaire que dans les MHCR [44].

La protéine 1 de la mort cellulaire programmée 1 (PD-1) et le ligand de mort programmé 1 (PD-L1)

PD-1 et PD-L1 sont deux glycoprotéines exprimées respectivement à la surface des cellules T immunologiques et des cellules tumorales. L'interaction entre ces deux marqueurs de surface entraîne la régulation négative des cellules T et, par conséquent, la multiplication des cellules tumorales (Figure 2.4) [48]. Dans le cas du traitement anti-PD1, les anticorps anti-PD1 interagissent avec PD-1 et bloquent le mécanisme d'évasion immunitaire [49]. L'expression

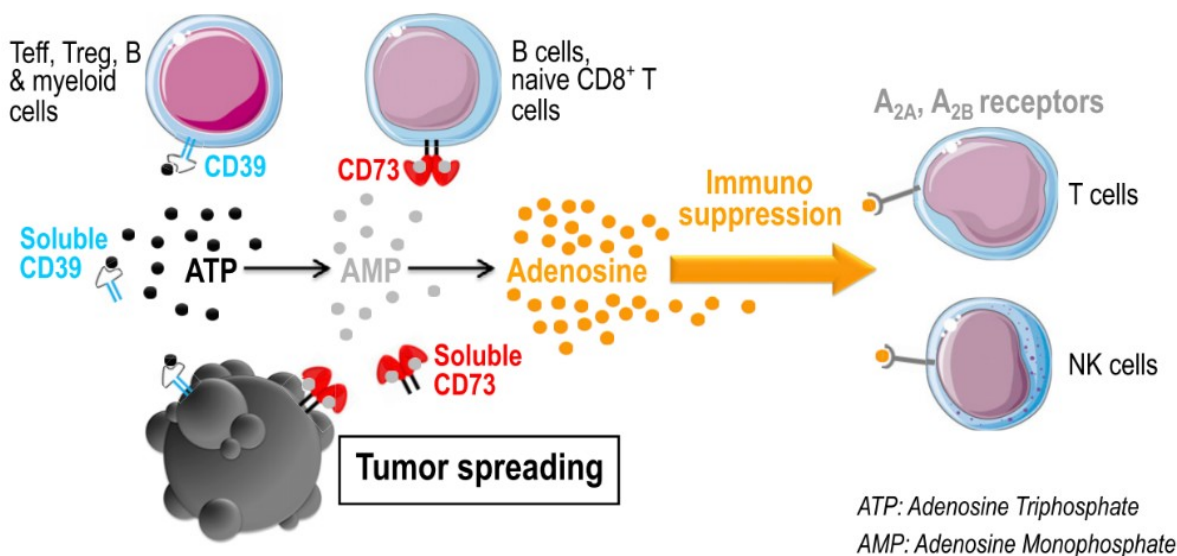


FIGURE 2.3 Rôle de CD73 (ecto-5'-nucléotidase) dans l'immunosuppression tumorale [1]

de PD-L1 est donc utilisée comme biomarqueur pour la thérapie anti-PD1 et joue un rôle important dans la réponse à l'immunothérapie contre le cancer [36, 37, 50].

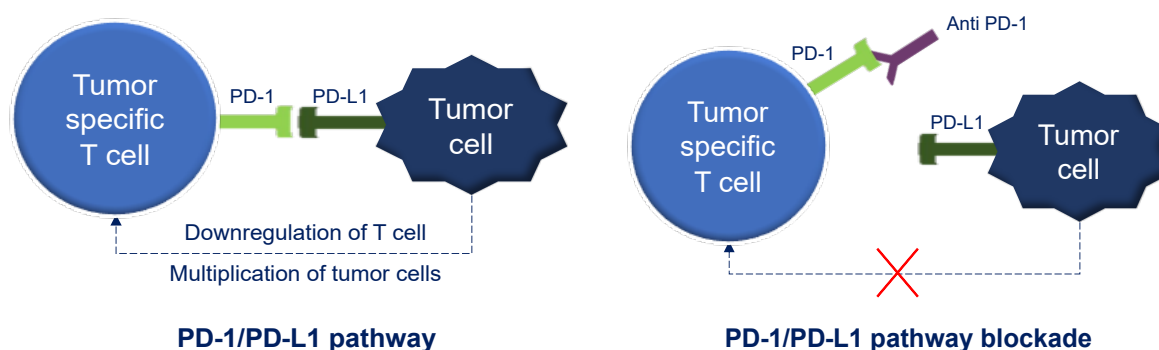


FIGURE 2.4 La voie de régulation négative de PD-1/PD-L1

D'autres protéines de point de contrôle immunitaire sont exprimées dans le CCR et contribuent au mécanisme d'évasion immunitaire, telles que la protéine 4 associée aux lymphocytes T cytotoxiques, le gène d'activation lymphocytaire 3, l'immunoglobuline mucine-3 des lymphocytes T et les immunorécepteurs des lymphocytes T avec domaines Immunoglobuline et ITIM (TIGIT).

2.5 Prise en charge des métastases hépatiques colorectales

Dans le cas des MHCR résécables, une résection curative est généralement réalisée [2]. La résécabilité est souvent déterminée par le nombre et la distribution des métastases, le volume du futur reste du foie (FRF) et la fonction hépatique. Une hépatectomie à une ou deux étapes est souvent réalisée. Le choix de la chirurgie à une étape dépend du volume du FRF et peut inclure une augmentation du tissu par embolisation de la veine porte. En cas d'échec de cette dernière, une hépatectomie à deux étapes est considérée pour potentiellement effectuer une élimination complète de la tumeur.

Lorsque les métastases ne sont pas résécables, une chimiothérapie de conversion est administrée pour augmenter la possibilité d'une intervention chirurgicale ou contrôler les MHCR. Les patients présentant des MHCR non résécables peuvent être évalués pour une transplantation hépatique si la tumeur est bien contrôlée par chimiothérapie (Figure 2.5). Dans certains cas particuliers, certains patients subissent des procédures ablatives ou une radiothérapie [2].

Après la chirurgie initiale, une chimiothérapie adjuvante est administrée aux patients qui n'ont pas reçu de traitement adjuvant préalable [51]. Les patients présentant des MHCR synchrones, un nombre élevé ou une grande taille de MHCR, ainsi que des métastases extra-hépatiques sont aussi envisagés pour une chimiothérapie périopératoire.

D'autre part, des efforts récents ont été déployés pour étudier l'utilisation de l'immunothérapie dans le traitement des MHCR, notamment : (1) les ICI, (2) la thérapie adoptive par cellules T (TAC), et (3) les vaccins contre le cancer [13]. Bien que l'efficacité des ICI n'ait pas été satisfaisante pour tous les CCR métastatiques, les patients présentant une instabilité microsatellitaire (IMS) élevée étaient généralement plus réactifs, ce qui a conduit la FDA des États-Unis à approuver le traitement anti-PD-1 pour ce sous-ensemble de patients [52].

2.6 Les normes de traitement du CPNPC

Le traitement du cancer du poumon a considérablement évolué ces dernières années. Pour le CPNPC, les options thérapeutiques comprennent la chirurgie, la chimiothérapie, la radiothérapie, l'immunothérapie et les thérapies ciblées (Figure 2.6). Le choix du traitement dépend du stade de la maladie, des caractéristiques biologiques et moléculaires de la tumeur, ainsi que du profil de santé global du patient, notamment la présence de contre-indications à la chirurgie.

La Société Canadienne du Cancer a établi des lignes directrices spécifiques soulignant les options de traitement du CPNPC selon le stade de la maladie et la biologie [53]. La chirurgie

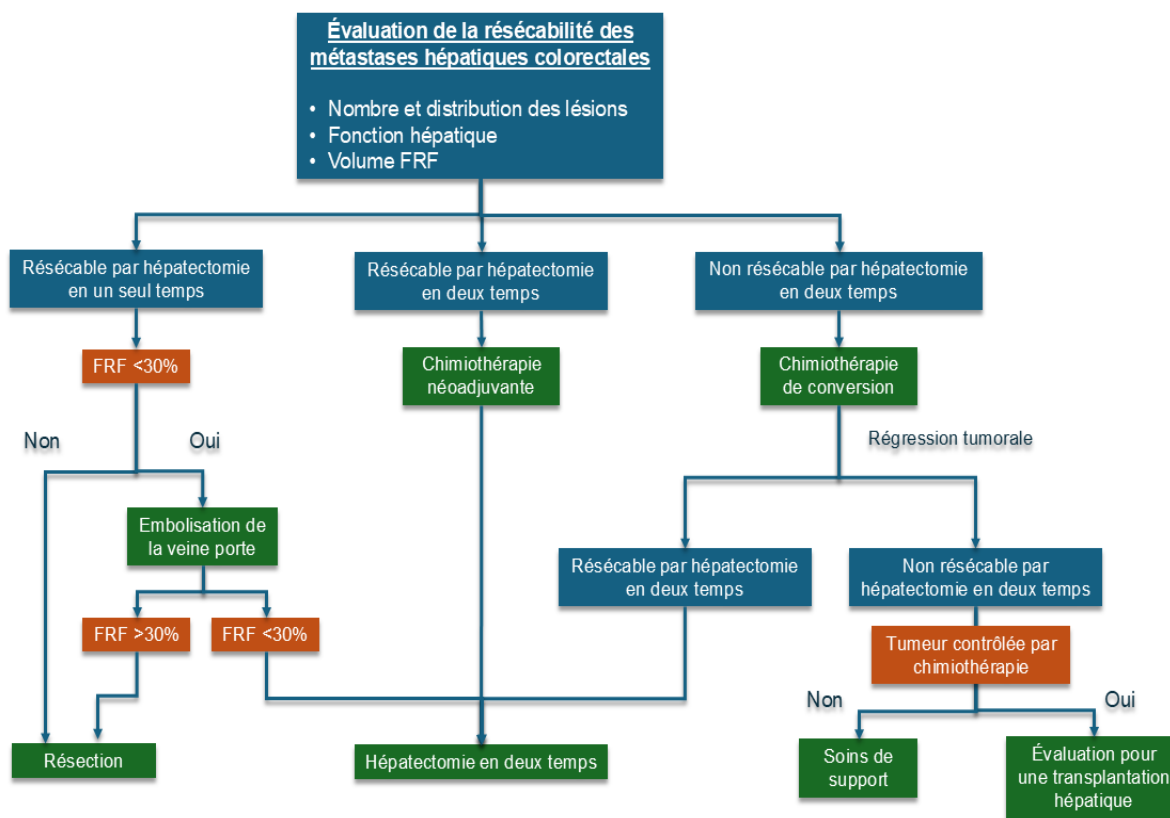


FIGURE 2.5 Algorithme de prise en charge des MHCR (Centre suisse d'hépatobiliaire et pancréatique, Zurich, Suisse [2])

demeure l'option curative principale pour le CPNPC résécable (stades I et II ainsi que certains stades IIIA résécables) [21]. La radiothérapie peut également être administrée aux patients atteints de CPNPC au stade I qui sont incapables ou non désireux de subir une intervention chirurgicale [54].

L'utilisation des ICI a considérablement amélioré les résultats cliniques du CPNPC à divers stades de la maladie. Dans le cas du CPNPC avancé et métastatique, l'administration de l'immunothérapie repose sur le profil d'expression de la PD-L1 : pour les patients chez lesquels l'expression de la PD-L1 est supérieure ou égale à 50 %, l'immunothérapie est administrée seule. Pour les patients dont l'expression est inférieure à 50 % (<1 % et 1-49 %), l'immunothérapie est souvent combinée avec la chimiothérapie [55]. Le CPNPC présentant certaines mutations (par exemple, KRAS, EGFR, ALK, etc.) peut également être traité par thérapie ciblée [21]. Les découvertes récentes ont montré que les patients atteints de CPNPC résécable pourraient bénéficier d'un traitement de chimiothérapie-immunothérapie néoadjuvante, qui est devenu la norme, en particulier pour ceux sans mutations oncogéniques [56]. La thérapie néoadjuvante présente plusieurs avantages par rapport au traitement adjuvant,

notamment une administration plus cohérente et l'élimination précoce des micrométastases. L'immunothérapie peut aussi être administrée en tant que thérapie de consolidation pour les patients atteints de CPNPC localement avancé, non résécable, au stade III, après une chimiothérapie-radiothérapie définitive [57].

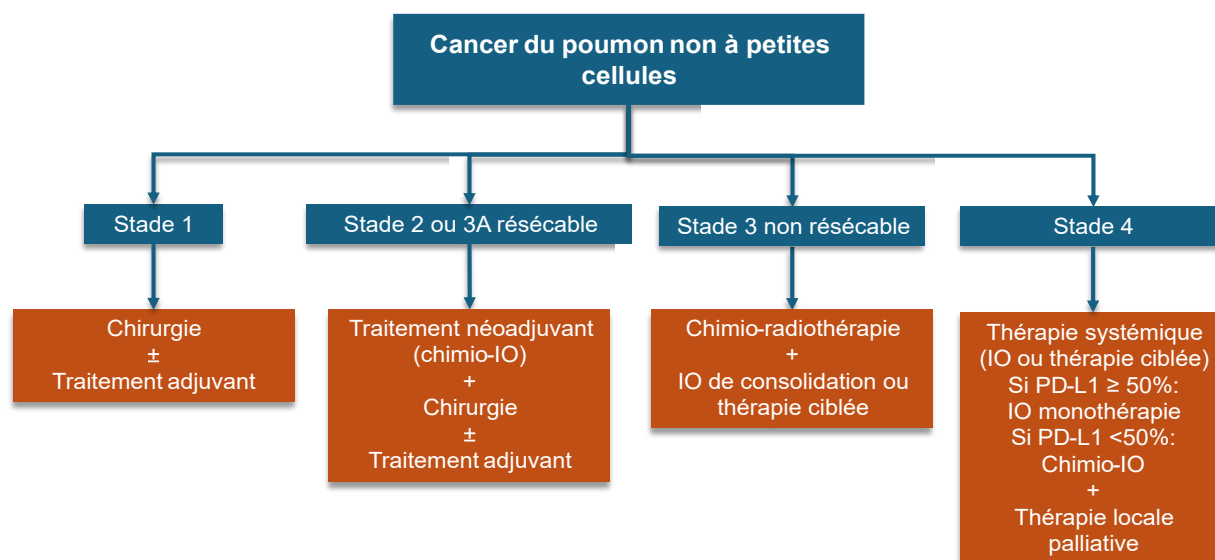


FIGURE 2.6 Traitement du CPNPC selon le stade de la maladie selon la Société Américaine du Cancer [3]; IO : immunothérapie, chimio : chimiothérapie.

2.7 La radiomique

La radiomique est la quantification des informations d'imagerie médicale par l'extraction d'une combinaison de caractéristiques conçues manuellement [58]. Le concept repose sur l'hypothèse selon laquelle de telles caractéristiques quantitatives peuvent refléter les caractéristiques au niveau microscopique des cellules, telles que les propriétés génomiques et protéomiques. Le pipeline radiomique inclut généralement les étapes suivantes : (1) l'acquisition d'images, (2) la définition de la région d'intérêt par segmentation manuelle ou automatique, (3) l'extraction des caractéristiques radiomiques et (4) la génération de modèles prédictifs dans lesquels les résultats cliniques ou les biomarqueurs de différentes sources (pathologie, génomique, antécédents des patients) sont prédits [59].

Les caractéristiques radiomiques incluent : (1) les statistiques de premier ordre, qui sont des caractéristiques basées sur des histogrammes reflétant la distribution des valeurs de voxels dans la région d'intérêt (par exemple, l'entropie, l'asymétrie, l'intensité moyenne, minimale, maximale, etc.), (2) les caractéristiques de forme (par exemple, la sphéricité, la surface, le

rapport aire/volume, etc.), et (3) les caractéristiques texturales extraites par des matrices spécifiques telles que la matrice de co-occurrence des niveaux de gris, la matrice de dépendance des niveaux de gris, la matrice de longueur de parcours des niveaux de gris, la matrice des tailles de zones des niveaux de gris et la matrice de différence des tons de gris voisins [60]. Ces caractéristiques radiomiques peuvent être extraites de différents types de modalités d'imagerie médicale telles que les images CT, la résonance magnétique, la tomographie par émission de positron. Elles peuvent aussi être utilisées seules ou en combinaison avec d'autres données omiques telles que les données génomiques, histologiques ou les propriétés cliniques.

Le prétraitement d'images joue un rôle crucial dans l'analyse radiomique. L'application de filtres tels que le Laplacien de Gauss et la décomposition en ondelettes avant l'extraction des caractéristiques peut améliorer la qualité informative des caractéristiques radiomiques extraites [61]. De plus, les filtres Laplacien de Gauss et les décompositions en ondelettes permettent d'améliorer la reproductibilité et la répétabilité des caractéristiques radiomiques, ce qui est d'une grande importance pour les images prises dans différents centres avec des paramètres d'acquisition variés, tels que l'épaisseur des coupes [62]. D'une part, les filtres Laplacien de Gauss permettent de réduire le bruit en détectant les régions de changement rapide d'intensité, ce qui permet d'améliorer les contours et les textures. D'autre part, l'opération de décomposition en ondelettes permet de décomposer l'image en bandes de fréquences distinctes, ce qui peut permettre de séparer le bruit de l'image et de réaliser une analyse multi-résolution.

Divers paramètres d'extraction peuvent influencer les valeurs des caractéristiques radiomiques obtenues. Le premier hyperparamètre concerne la largeur des intervalles (« bins ») utilisés lors de la discrétisation des images. Avant l'extraction des caractéristiques, les intensités de gris au sein de la région d'intérêt sont discrétisées. La largeur des bins définit la taille des intervalles lors de la construction de l'histogramme, influençant ainsi la répartition des niveaux de gris.

Le second paramètre est la distance entre un voxel central et ses voxels voisins. Cette distance intervient dans la définition de la notion de voisinage et a un impact direct sur les caractéristiques dérivées des matrices texturales. Elle détermine la distance spatiale maximale à considérer entre un voxel de référence et ses voisins.

Le troisième paramètre porte sur le choix entre une matrice de co-occurrence des niveaux de gris symétrique ou asymétrique. Dans le cas symétrique, les co-occurrences sont évaluées dans les deux directions pour chaque angle, contrairement au cas asymétrique où elles ne sont prises en compte que dans une seule direction.

Un quatrième paramètre est le seuil de dépendance α , utilisé dans le calcul de la matrice de

dépendance des niveaux de gris. Deux voxels voisins, de niveaux de gris respectifs i et j , sont considérés comme dépendants si la condition $|i - j| < \alpha$ est satisfaite.

En ce qui concerne le filtre Laplacien de Gauss, le paramètre sigma (σ) contrôle le niveau de détail textural mis en évidence. Une faible valeur de σ met en exergue les textures fines, caractérisées par des variations rapides des niveaux de gris, tandis qu'une valeur élevée de σ révèle des textures plus grossières, associées à des changements sur de plus grandes distances spatiales.

Enfin, pour la décomposition en ondelettes, deux paramètres permettent de personnaliser l'opération : le niveau de départ, qui indique à partir de quel niveau d'ondelette commence la décomposition, et le nombre total de niveaux de décomposition appliqués.

2.8 Apprentissage automatique

L'apprentissage automatique est une branche de l'intelligence artificielle qui permet aux systèmes d'apprendre de manière autonome à partir de données et de prendre des décisions avec une intervention humaine minimale [63]. Cette technologie constitue le pilier de nombreuses applications modernes, y compris l'aide au diagnostic dans le domaine de la santé et les recommandations personnalisées dans le domaine du commerce électronique. L'apprentissage automatique utilise des techniques statistiques pour apprendre des modèles importants à partir des données, améliorant progressivement les performances sur des tâches spécifiques sans être explicitement programmé pour chaque scénario. En utilisant les modèles appris, les modèles d'apprentissage automatique peuvent ensuite faire des prédictions sur des données non vues. Les applications de l'apprentissage automatique comprennent la reconnaissance d'images et le traitement du langage naturel.

Il existe différentes approches de l'apprentissage automatique, principalement l'apprentissage supervisé, semi-supervisé et non supervisé. Dans l'apprentissage supervisé, les modèles sont généralement formés à l'aide de données entièrement annotées, composées d'entrées et de sorties associées, afin d'effectuer des tâches spécifiques. Le modèle est donc formé pour apprendre une équation mathématique qui associe les entrées aux sorties correspondantes. L'apprentissage supervisé est particulièrement utile pour les tâches de classification et de régression.

L'apprentissage non supervisé exploite uniquement des données non annotées et vise à former le modèle pour qu'il apprenne les tendances et les relations inhérents aux données sans avoir accès aux sorties correspondantes. Les algorithmes de regroupement (clustering) et l'analyse en composantes principales sont des exemples largement utilisés d'apprentissage

non supervisé. Ce dernier trouve des applications dans la détection des anomalies, ainsi que dans le pré-entraînement des modèles pour des tâches d'apprentissage supervisé en aval.

Les principales étapes de l'apprentissage supervisé sont les suivantes :

1. Collecte, nettoyage et prétraitement de grandes quantités de données annotées (paires entrée-sortie) ;
2. Ingénierie des caractéristiques dans laquelle les caractéristiques des données sont collectées ;
3. Sélection des caractéristiques dans laquelle seules les caractéristiques les plus saillantes sont incluses pour l'entraînement du modèle ;
4. Entraînement des modèles d'apprentissage automatique permettant ainsi d'affiner les paramètres des modèles pour effectuer une tâche spécifique ;
5. Évaluation des performances du modèle en vue d'un déploiement potentiel.

Les architectures suivantes sont des exemples de modèles d'apprentissage automatique :

1. Les arbres de décision qui séparent l'espace des caractéristiques en effectuant une succession de divisions binaires.
2. Les forêts aléatoires qui combinent plusieurs arbres de décision formés sur différents sous-ensembles d'instances.
3. Les algorithmes de boosting de gradient qui combinent plusieurs arbres de manière à ce que chaque arbre soit formé pour corriger les erreurs résiduelles de l'arbre précédent à l'aide de la descente de gradient.
4. La régression linéaire, qui utilise des équations mathématiques entre les entrées et les sorties pour prédire des valeurs continues, et la régression logistique, qui effectue des tâches de classification.
5. Les machines à vecteurs de support qui sont formées pour trouver l'hyperplan optimal qui maximise la marge entre les différentes classes dans l'espace des caractéristiques.
6. Les réseaux neuronaux construits en organisant les nœuds connectés en une séquence de couches, convertissant ainsi les données en liaisons pondérées et en fonctions d'activation.

L'apprentissage profond est un sous-type d'apprentissage automatique qui exploite les architectures basées sur les réseaux neuronaux pour extraire des représentations significatives à partir de données brutes, en évitant souvent la nécessité d'une ingénierie manuelle de caractéristiques élaborées à la main.

CHAPITRE 3 REVUE DE LITTÉRATURE

3.1 Apprentissage automatique en imagerie médicale du cancer

L'utilisation d'algorithmes d'apprentissage automatique avec l'imagerie médicale a révolutionné la détection du cancer, le diagnostic, la segmentation et la planification du traitement, en particulier pour les MHCR et le cancer du poumon. Les approches récentes de l'apprentissage automatique permettent d'améliorer de multiples aspects de l'imagerie du cancer à travers diverses modalités, en améliorant les performances diagnostiques et en fournissant des informations pronostiques qui étaient auparavant impossibles à obtenir.

Les tâches de classification en imagerie médicale sont très répandues, en particulier dans le cadre du diagnostic assisté par ordinateur. Romero et al. [64] ont exploité l'architecture InceptionV3 [65] pour classer les lésions hépatiques en MHCR et en kystes. Une étude s'est appuyée sur les caractéristiques radiomiques extraites des lésions colorectales primaires sur les images de TDM de base pour prédire la probabilité d'apparition de métastases hépatiques métachrones [66]. Une autre étude s'est concentrée sur la caractérisation des lésions colorectales en extrayant des caractéristiques radiomiques de polypes segmentés manuellement sur des images colonographiques par TDM et en appliquant une classification par forêt aléatoire (RF) à ces caractéristiques. Le pipeline a pu prédire le caractère histopathologique de polypes individuels, démontrant le potentiel de la radiomique pour l'identification des lésions précancéreuses dans les programmes de dépistage colorectal [67]. En ce qui concerne le CPNPC, Jacob et al. [68] ont prédit les sous-types de cancer du poumon à l'aide de caractéristiques radiomiques extraites d'images de TDM. Cette approche permet une caractérisation plus accessible des sous-types de tumeurs, ce qui peut conduire à des stratégies de traitement plus ciblées.

En outre, un grand nombre de travaux de classification ont été réalisés sur des images histopathologiques, principalement en raison de leur grande disponibilité. Kapil et al. [69] ont proposé une approche semi-supervisée pour automatiser le calcul du niveau d'expression de PD-L1 dans l'IHC afin de s'affranchir de la nécessité d'annotations manuelles fastidieuses. Le modèle a surpassé le calcul visuel en termes de répétabilité. En ce qui concerne le cancer du sein, Krijgsman et al. [70] ont développé un pipeline combinant des images H&E et IHC recalées pour quantifier automatiquement la densité des CD8 TIL au centre et à la marge invasive des tumeurs du cancer du sein et ont étudié la corrélation entre les nombres obtenus et les résultats cliniques.

En outre, la segmentation des tumeurs représente une application critique où les modèles d'apprentissage automatique peuvent être appliqués pour délimiter les frontières du cancer. Un algorithme de segmentation automatique [71] a été proposé pour segmenter les lésions MHCR. Le modèle de segmentation consiste en deux réseaux neuronaux convolutionnels (RNC) entraînés bout à bout afin de segmenter conjointement le foie et les lésions qu'il contient. Une étude récente visait à remédier au manque de généralisation des modèles de segmentation entre différentes modalités d'imagerie. À cette fin, une stratégie d'apprentissage générative semi-supervisée pour la segmentation des tumeurs sur des ensembles de données bimodales non appariées a été proposée et évaluée sur un ensemble de données de segmentation de tumeurs cérébrales [72]. L'approche s'appuie sur un objectif non supervisé permettant au modèle de démêler les tumeurs de l'arrière-plan, en parallèle à la tâche de segmentation. En utilisant les annotations au niveau des pixels disponibles dans la modalité source, la segmentation est effectuée dans la modalité cible non annotée en entraînant le modèle à convertir les images d'une modalité à l'autre.

L'apprentissage automatique s'est avéré très prometteur dans l'imagerie des MHCR et CPNPC, avec des applications couvrant toutes les tâches d'analyse d'images médicales conventionnelles. La prochaine étape importante est l'utilisation de l'apprentissage automatique et de l'imagerie médicale pour le pronostic, où des algorithmes avancés prédisent la progression de la maladie et les résultats pour le patient. La section suivante traite de ce domaine émergent et de son potentiel d'amélioration des stratégies de traitement personnalisées.

3.1.1 Pronostication basée sur l'imagerie du cancer

La pronostication du cancer, en particulier dans le cas des MHCR et du CPNPC, est cruciale pour la sélection des patients et la planification du traitement. Les recherches récentes ont produit des applications prometteuses de l'apprentissage automatique pour répondre à ce besoin. Plusieurs techniques d'apprentissage automatique sont utilisées pour prédire des résultats cliniques importants, tels que la survie, la récurrence, la progression de la maladie et la réponse au traitement dans ces cancers (Figure 3.1).

Une étude prospective réalisée sur 17 patients avec des MHCR non résécables traités pour prédire la réponse à la thérapie biologique ciblée (chimiothérapie avec bevacizumab ou cetuximab). Des modèles de mémoire à long court terme (Long Short Term Memory) ont été entraînés à l'aide de caractéristiques radiomiques extraites des images IRM à trois moments (baseline, et deux suivis) pour prédire la réponse au traitement après 8 semaines post-chimiothérapie. En atteignant une précision de 76,5 %, l'étude montre comment l'apprentissage automatique peut potentiellement identifier plus tôt les patients répondant au

traitement que les méthodes conventionnelles comme les critères RECIST (Critères d'évaluation de la réponse dans les tumeurs solides), permettant ainsi des modifications de traitement plus opportunes [73]. Qu et al. ont prédit la réponse à la thérapie antiangiogénique en extrayant des caractéristiques de "radiomique dynamique" à partir d'images de TDM avec contraste acquises avant traitement. En extrayant des caractéristiques radiomiques à partir de différentes séquences TDM et en les combinant en caractéristiques dynamiques, les auteurs ont montré que l'utilisation de la radiomique dynamique pourrait conduire à de meilleurs résultats que la radiomique dérivée d'une seule séquence TDM [74]. Une autre étude menée sur 79 patients atteints de MHCR s'est concentrée sur la prédiction de la progression locale de la tumeur après un traitement par ablation thermique à l'aide d'images de TDM [75]. Li et al. ont proposé une approche basée sur un RNC multi-dimension pour stratifier le risque de récurrence des patients atteints de CCR en utilisant des images de TDM préopératoires [76]. Les auteurs ont recruté 808 patients et ont atteint une aire sous la courbe de caractéristique de fonctionnement du récepteur (AUC) de 0,610 dans la cohorte de validation. Une étude menée sur des patients atteints de CCR de stade II a proposé de fusionner les caractéristiques TDM extraites des lésions primaires et des ganglions lymphatiques périphériques pour la prédiction de la SSC et de la SG [77]. D'autres travaux se sont concentrés sur la prédiction de la réponse des MHCR à la chimiothérapie. Maaref et al. ont entraîné un RNC sur des images de TDM pour prédire la réponse au traitement FOLFOX avec Bevacizumab, en se basant sur le grade de régression tumorale évalué par histopathologie. Les auteurs ont atteint une précision de 78 %, mettant en évidence le potentiel pour la détection précoce des patients non répondants [78]. Une autre étude a utilisé des caractéristiques radiomiques extraites d'images de TDM pour entraîner un modèle de machine à vecteurs de support (SVM) pour la prédiction de la réponse à la chimiothérapie [79]. En évaluant le changement du plus grand diamètre des lésions, les patients ont été divisés en répondeurs et non-répondeurs. Le modèle a ensuite été entraîné pour classifier la réponse des patients et a atteint une précision de 61,5 % sur l'ensemble de validation. Une étude menée sur 85 patients a atteint une AUC de 0,818 pour la prédiction de la réponse à la chimiothérapie, telle qu'évaluée selon les critères RECIST Version 1.1. Des paramètres conventionnels tels que l'intensité du signal ainsi que des caractéristiques radiomiques extraits d'IRM en phase veineuse porte et en phase hépatobiliaire ont été utilisés pour entraîner un SVM linéaire à cet effet. En utilisant un modèle de régression de Cox, les auteurs ont trouvé une association statistiquement significative entre la réponse des MHCR prédite par le modèle et la survie sans progression [80].

La pronostication du CPNPC a également bénéficié des applications de l'apprentissage automatique, avec plusieurs approches innovantes proposées à cet effet. Chen et al. ont proposé une méthode basée sur un RNC pour classifier la SG à 2 ans en utilisant des images de

TDM [81]. Dans le cas des patients traités par immunothérapie et chimiothérapie, une étude réalisée sur 60 patients atteints de CPNPC a rapporté des résultats prometteurs pour la classification de la survie sans progression (SSP) à 12 mois. Les auteurs ont extrait des caractéristiques radiomiques des images de TDM et ont entraîné des modèles SVM pour prédire si la SSP est supérieure à 12 mois [82]. Néanmoins, les auteurs d’une autre étude menée sur 102 patients atteints de CPNPC traités par immunothérapie ont rapporté une précision limitée dans la prédiction de la progression de la maladie à l’aide d’un modèle de RF entraîné sur des caractéristiques radiomiques dérivées des images de TDM [83]. Cela souligne la nécessité d’études à grande échelle avec des cohortes de validation indépendantes afin de garantir la reproductibilité des résultats et la généralisation des modèles proposés à travers les différentes caractéristiques des patients et les paramètres d’acquisition.

Récemment, l’approche consistant à cibler des proxies pronostiques a gagné en importance, avec plusieurs études se concentrant sur la prédiction de biomarqueurs cliniquement établis. Lee et al. ont entraîné deux modèles d’apprentissage profond pour la prédiction des caractéristiques histopathologiques à partir de TDM, à savoir l’invasion viscéropleurale et l’invasion lymphovasculaire, en utilisant une approche de moyenne des sorties du modèle pour la prédiction de la SG et de la survie sans récurrence [84]. L’étude suggère qu’un score composite obtenu à partir des TDM et basé sur des caractéristiques histopathologiques, pourrait avoir une pertinence pronostique. D’autres études se sont concentrées sur la prédiction de biomarqueurs immunologiques en tant que proxies des résultats cliniques des patients.

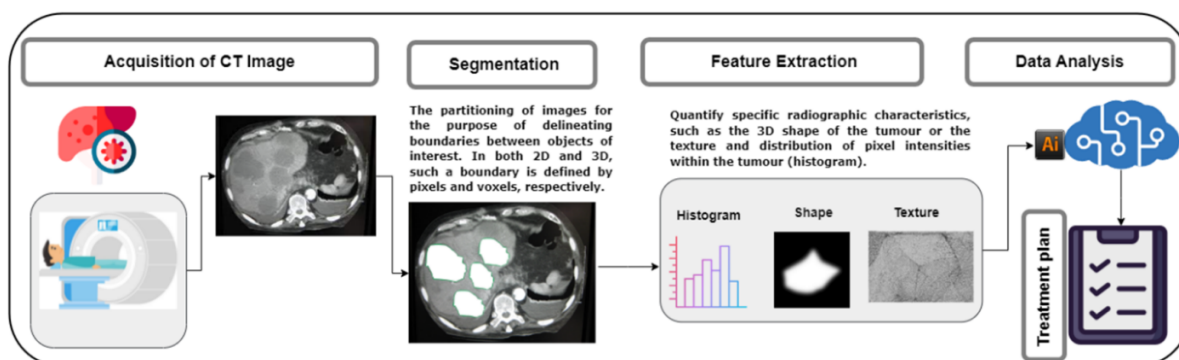


FIGURE 3.1 Le flux de travail radiomique adopté dans la littérature pour la prédiction des cibles thérapeutiques dans les MHCR à l’aide de l’imagerie médicale [4]. Le flux de travail commence généralement par la collecte d’images médicales, suivie de la segmentation de la région d’intérêt et de l’extraction des caractéristiques radiomiques. Ces dernières sont utilisées pour former des modèles d’apprentissage automatique pour la prédiction de biomarqueurs pronostiques, de la réponse au traitement ou des résultats cliniques.

3.1.2 Prédiction de biomarqueurs pronostiques à l'aide de l'imagerie du cancer

Les techniques d'apprentissage automatique pour la prédiction des marqueurs pronostiques du cancer à partir des images médicales ont montré un grand potentiel à travers divers types de cancer, biomarqueurs et modalités d'imagerie. Ces approches offrent des alternatives non invasives à la biopsie tissulaire conventionnelle, fournissant des informations sur les phénotypes moléculaires, les cibles thérapeutiques, les mutations génétiques et les lymphocytes infiltrant la tumeur. De telles méthodes peuvent faciliter une évaluation plus complète de la tumeur et soutenir une planification du traitement plus précoce et plus précise.

Prédiction des cibles thérapeutiques basée sur l'imagerie

Des efforts ont été faits pour dériver les caractéristiques immunologiques du cancer à partir des données d'imagerie médicale en tirant parti des progrès de la radiomique et de l'apprentissage automatique.

Tian et al. [85] ont conçu et mis en œuvre un pipeline pour la prédiction de l'expression de PD-L1 dans le CPNPC à partir des images de TDM avant traitement. Les tests IHC, les images et les données cliniques ont été inclus dans l'étude. Le pipeline se compose de deux étapes principales : une étape d'extraction des caractéristiques et une étape de classification. Les caractéristiques d'imagerie ont été concaténées et ajoutées à quatre caractéristiques cliniques (âge, sexe, antécédents de tabagisme, antécédents familiaux) afin d'intégrer les connaissances du domaine médical dans le pipeline. L'étape de classification consiste en réseau neuronal produisant des classes d'expression : PD-L1^{high} et PD-L1^{low}. Une AUC de 0,76 a été atteinte sur le jeu de test.

Pour la même tâche de classification, Wen et al. ont obtenu une AUC de 0,730 en utilisant un modèle de régression logistique entraîné sur des caractéristiques radiomiques extraites des images de TDM de 120 patients atteints de CPNPC [86]. Pour prédire l'expression de PD-L1, Yoon et al. ont obtenu une AUC de 0,661 en utilisant des caractéristiques radiomiques et une régression logistique multivariée effectuée sur 153 patients atteints de CPNPC avancés [87].

Yolchuyeva et al. ont mené une étude multicentrique sur 385 patients atteints de CPNPC avancé susceptibles de bénéficier des ICI, recrutés dans deux centres académiques, et ont entraîné différents modèles pour prédire l'expression de PD-L1 et la SSP. Lors de l'évaluation de leur modèle sur la cohorte de validation, ils ont atteint une AUC de 59 % avec un modèle de régression logistique pour la prédiction de PD-L1 et de 63 % avec un SVM pour la prédiction de la SSP [88].

De même, Iwatate et al. [89] ont prédit le statut de mutation p53 et l'expression de PD-L1 à

partir d'images de TDM avec contraste acquis avant traitement du cancer pancréatique. En conséquence, les auteurs ont extrait des caractéristiques radiomiques à partir des volumes d'intérêt délimités manuellement (segmentations tumorales avec ou sans régions péri-tumorales) et ont entraîné un classificateur de renforcement du gradient extrême (XGB) pour produire des classes binaires. Après avoir inclus les régions péri-tumorales dans les volumes d'intérêt, l'AUC de classification est passée de 0,705 à 0,795 pour la mutation p53 et de 0,660 à 0,683 pour l'expression de PD-L1. Les auteurs ont également démontré que le statut prédit de p53 était associé à la SG des patients.

D'autre part, plusieurs chercheurs se sont concentrés sur la prédiction des mutations géniques à partir des images de TDM. Arafat et al. [90] ont prédit quatre mutations géniques dans le carcinome rénal à cellules claires. De plus, de nombreux travaux ont été rapportés sur la prédiction du statut de la mutation EGFR chez les patients atteints de cancer du poumon à partir des images de TDM en utilisant des caractéristiques radiomiques extraites des régions d'intérêt [91–95]. Des études comparables ont été menées sur la prédiction des mutations KRAS dans différents types de cancers, y compris le CPNPC [96], le CCR primaire [97–99], et les MHCR [100]. Dans ce dernier travail, les auteurs ont combiné des caractéristiques radiomiques avec des caractéristiques sémantiques extraites des images de TDM avec contraste et ont entraîné un réseau de neurones pour prédire le statut de mutation (RAS ou BRAS). Une AUC de 0,79 a été atteinte sur le jeu de validation.

Cao et al. [101] ont proposé un modèle d'apprentissage profond basé sur des images de TDM de pré-traitement pour prédire le statut de réparation des erreurs de l'ADN (mismatch repair) chez les patients atteints de CCR. Leur modèle a atteint une AUC de 0,986 dans la cohorte de validation interne de 482 patients et de 0,915 dans la cohorte de validation externe de 206 patients. Li et al. ont développé une signature radiomique pour la prédiction de l'expression génique élevée vs faible du gène RRM2 en entraînant un modèle de régression logistique à partir des images de TDM [102]. Ils ont également rapporté que l'expression élevée de RRM2 était un facteur de risque pour l'OS.

Prédiction des lymphocytes infiltrant la tumeur basée sur l'imagerie

Plusieurs chercheurs ont tenté de prédire les comptes stratifiés de TIL à partir des images de TDM, avec des applications incluant les cancers du foie, du poumon, du pancréas et de l'estomac.

Liao et al. ont proposé un pipeline pour évaluer l'infiltration des cellules T CD8+ dans le carcinome hépatocellulaire (CHC) en utilisant des images de TDM présurgicales avec produit de contraste de 142 patients [103]. À cette fin, les auteurs ont réalisé une régression régula-

risée par réseau élastique pour générer un « Rad score » incorporant sept caractéristiques radiomiques dérivées des images de TDM. Ils ont également démontré l'association du « Rad score » avec la densité des cellules CD8+, l'expression de PD1 et de PD-L1, ainsi que la survie sans maladie et l'OS. Le travail pionnier de Sun et al. a adopté une approche similaire pour dériver une signature radiomique CD8 à partir des images de TDM avec produit de contraste de 135 patients ayant plusieurs types de cancers [104]. Le score a atteint une AUC de 0,67 sur un ensemble de validation externe et a pu différencier les tumeurs « immune inflamed » et les tumeurs « immune desert » ($AUC = 0,76$). La signature radiomique était également associée à la SG des patients et à leur réponse au traitement anti-PD-1/PD-L1.

En ce qui concerne les applications dans le cancer du poumon, Liu et al. ont entraîné un réseau neuronal pour prédire l'infiltration CD8+ dans le CPNPC à l'aide de l'imagerie TDM et ont atteint une AUC de 0,721 [105]. Zerka et al. ont présenté une approche radiomique basée sur le TDM pour la stratification de l'infiltration CD8+, démontrant une performance raisonnable sur des ensembles de tests indépendants avec des AUC de 0,67 et 0,64 [106]. Yoon et al. ont prédit les comptes stratifiés des cellules T auxiliaires de type 2 dans le CPNPC et ont atteint une AUC de 0,68 sur l'ensemble de test [107].

De même, Li et al. ont extrait 1409 caractéristiques radiomiques des images de TDM des tumeurs segmentées de l'adénocarcinome canalaire pancréatique et ont entraîné un classificateur XGB pour prédire les niveaux d'expression binaires des TIL CD8+ [108]. Une AUC de 0,67 a été atteinte sur l'ensemble de validation. Des études similaires ont été menées sur le cancer gastrique. Jiang et al. ont combiné des caractéristiques radiomiques intratumorales et péri-tumorales pour prédire un score immunologique binaire comprenant CD3, CD8, CD45RO et CD66b [109]. Les auteurs ont mis en œuvre un modèle de régression logistique régularisée atteignant une AUC variant entre 0,766 et 0,786 sur trois ensembles de validation indépendants. Gao et al. ont mis en place un pipeline radiomique pour évaluer les cellules T régulatrices infiltrant la tumeur et ont obtenu une AUC de 0,847 sur l'ensemble de test [110].

Une autre étude innovante a proposé d'évaluer les changements dynamiques des cellules T CD8+ infiltrant la tumeur et des macrophages dans le cancer du col de l'utérus après chimiothérapie et radiothérapie. Cette approche longitudinale représente une avancée par rapport aux évaluations statiques, permettant potentiellement une surveillance en temps réel des réponses immunitaires au traitement [111]. Ces résultats soutiennent la faisabilité de l'utilisation des caractéristiques de texture TDM pour différencier les niveaux d'infiltration élevés et faibles, offrant une méthode non invasive pour analyser l'ensemble de la tumeur et potentiellement identifier les patients susceptibles de répondre à l'immunothérapie.

D'autre part, certains auteurs se sont intéressés à l'application des images IRM pour la pré-

diction des TIL [112–115] avec des AUC variant entre 0,67 et 0,90. Bien que les travaux décrits ci-dessus fournissent des preuves de la relation entre les caractéristiques radiomiques et la densité des TIL dans plusieurs types de cancers, des investigations supplémentaires sont nécessaires pour élucider la capacité des caractéristiques dérivées des images à refléter les profils immunologiques du cancer. Des études de grande envergure avec validation indépendante sont essentielles pour garantir la fiabilité des modèles proposés. De plus, ce type d'études est insuffisamment exploré dans le contexte des MHCR.

3.2 Apprentissage automatique en génomique du cancer

L'apprentissage automatique a émergé comme une approche puissante pour analyser les données complexes de génomique du cancer, en particulier les données transcriptomiques dérivées du séquençage de l'ARN. Des études récentes montrent des progrès remarquables dans l'utilisation de modèles d'apprentissage automatique pour exploiter les données d'expression génique à des fins de classification du cancer, de sous-typage moléculaire et de prédiction des résultats des patients. Les algorithmes traditionnels d'apprentissage automatique tels que les RF, les algorithmes de boosting de gradient et les SVM continuent de démontrer des performances élevées. Avec l'avènement de techniques avancées comme le mécanisme d'attention et l'apprentissage génératif, les approches d'apprentissage profond sont de plus en plus appliquées pour gérer la haute dimensionnalité des données génomiques et transcriptomiques.

La classification du cancer, en particulier l'identification de sous-types moléculaires cliniquement pertinents qui guident les décisions de traitement, a été l'objectif de nombreuses études exploitant les données transcriptomiques et l'apprentissage automatique. Wu et al. [116] ont entraîné différents modèles d'apprentissage automatique traditionnels à l'aide de données de séquençage de l'ARN pour classer le sous-type moléculaire du cancer du sein en triple négatif et non triple négatif. L'étude montre que le modèle SVM a surpassé les autres algorithmes d'apprentissage automatique et a mis en évidence l'importance d'inclure davantage de gènes de séquençage de l'ARN afin de réduire la variabilité des performances des modèles. Néanmoins, les instances avec des données manquantes ont été supprimées dans l'analyse. Bobbée et al. ont entraîné un modèle RF pour classer les catégories les plus fréquentes des lymphomes B non hodgkiniens à l'aide de données d'expression génique [117]. Dans une autre étude, trois classificateurs binaires ont été entraînés à l'aide de gènes différentiellement exprimés dérivés du séquençage de l'ARN pour prédire les sous-types du cancer du sein [118].

Malgré ces progrès prometteurs, d'importants défis restent à relever, notamment les problèmes liés à la qualité des données et au prétraitement, tels que la gestion des transcrits de gènes d'ARN non détectés, les tailles d'échantillons limitées par rapport à la haute di-

mensionnalité des données génomiques et les préoccupations concernant l’interprétabilité des modèles avant leur mise en œuvre clinique.

3.2.1 Pronostication du cancer basée sur la transcriptomique

Les méthodes traditionnelles d’apprentissage automatique ainsi que les approches d’apprentissage profond ont démontré une valeur substantielle dans l’extraction de modèles significatifs à partir des données de séquençage de l’ARN, chacune offrant des avantages distincts pour différentes applications et contextes de données qui peuvent être exploités pour améliorer le pronostic du cancer.

Une étude a développé des signatures pronostiques de MHCR en utilisant des données de séquençage de l’ARN. À cette fin, les auteurs ont effectué une analyse d’expression génique différentielle pour identifier les gènes différentiellement exprimés entre les tissus sains et le MHCR. Ensuite, les auteurs ont identifié les gènes les plus pronostiques à l’aide d’un modèle de régression de Cox opérateur de sélection et de rétrécissement absolu minimum (Least Absolute Shrinkage and Selection Operator, LASSO). Sur la base des gènes sélectionnés, un modèle multivarié de prédiction du risque par régression de Cox a été construit avec la SG et la survie sans progression comme critères d’évaluation [119]. Xie et al. ont formé plusieurs modèles d’apprentissage automatique pour la prédiction de la SG dans le CCR en utilisant des données de séquençage de l’ARN en vrac et unicellulaire. À cette fin, les auteurs ont identifié les gènes différentiellement exprimés entre les patients sains et les patients atteints de CCR en utilisant le séquençage de l’ARN en vrac. Ils ont ensuite identifié les gènes liés aux macrophages qui étaient exprimés de manière différentielle entre les macrophages et les autres populations cellulaires. En choisissant l’intersection des deux ensembles de gènes obtenus, les auteurs ont obtenu des gènes liés aux macrophages dans le CCR qui ont été utilisés dans une analyse de régression de Cox univariée pour rechercher des gènes liés aux macrophages ayant une signification pronostique potentielle. Les gènes pronostiques obtenus ont été utilisés pour former des modèles d’apprentissage automatique pour la prédiction de la survie [120]. Ce travail élucide le potentiel des modèles basés sur la transcriptomique en tant qu’outil de prédiction personnalisée du pronostic chez les patients atteints de CCR.

Un pipeline d’apprentissage profond proposé par Huang et al. a combiné une architecture d’autoencodeur avec la régression de Cox pour la prédiction de la survie dans plusieurs types de cancer, y compris le CHC et deux sous-types de CPNPC : l’adénocarcinome pulmonaire et le carcinome épidermoïde [121]. L’étude montre que les performances du pipeline sont fortement influencées par le type de cancer, ce qui souligne la nécessité d’adapter les approches prédictives d’une application à l’autre. Les résultats obtenus dans une autre étude menée sur

les cancers du poumon, du rein et du sein soulignent également la nécessité d’adapter les approches à chaque type de cancer. Les auteurs ont formé différents modèles d’apprentissage automatique pour la classification du pronostic dans les trois types de cancer (survie supérieure à 5 ans contre moins de 2 ans) en utilisant des données d’expression ARNm et ont obtenu des performances sensiblement différentes par type de cancer [122].

Une autre étude visait à prédire la récurrence du CPNPC à l’aide de données de séquençage de l’ARN chez 130 patients atteints [123]. En utilisant la technique de suréchantillonnage des minorités synthétiques (SMOTE) pour atténuer le déséquilibre des classes et des algorithmes d’apprentissage automatique supervisé, les auteurs ont obtenu des résultats prometteurs pour la prédiction de la récurrence basée sur l’expression génique. En outre, Cao et al. ont entraîné différents modèles de survie en utilisant des valeurs d’expression génique préalablement sélectionnées pour la prédiction de la survie des patients atteints d’adénocarcinome pulmonaire. Les auteurs ont constaté que le modèle de RF de survie (random survival model) était plus performant que les autres modèles. Ils ont également identifié un ensemble de gènes systématiquement associés à des résultats cliniques plus défavorables chez les patients atteints d’adénocarcinome pulmonaire [124].

Malgré ces avancées prometteuses, des défis importants restent à relever, notamment les questions liées à la qualité des données et au prétraitement, la taille limitée des échantillons par rapport à la dimensionnalité élevée des données génomiques, les problèmes d’interprétabilité des modèles et le besoin critique d’une validation externe solide avant la mise en œuvre clinique.

3.3 Pronostication du cancer basé sur les caractéristiques cliniques

Les modèles d’apprentissage automatique appliqués aux caractéristiques cliniques pour prédire le pronostic du cancer constituent une approche prometteuse qui complète les méthodes basées sur l’imagerie et la génomique. Ces modèles s’appuient sur des données de patients facilement accessibles pour prédire les résultats et guider les décisions thérapeutiques.

Une étude réalisée sur 40 870 patients atteints de CCR visait à prédire la probabilité de développement de métastases hépatiques à l’aide de caractéristiques cliniques, notamment le stade de la maladie et les informations relatives au traitement [125]. Une autre étude s’est concentrée sur la prédiction du CCR avec réparation déficiente des mésappariements (deficient mismatch repair). À cette fin, ils ont formé cinq modèles d’apprentissage automatique en utilisant 12 caractéristiques clinicopathologiques de base et ont atteint une AUC de 0,86 avec un modèle de RF [126].

Des modèles d'apprentissage automatique basés sur les caractéristiques cliniques ont également été exploités pour le pronostic du CPNPC. Janik et al. ont formé des modèles d'apprentissage automatique tabulaires et graphiques pour la prédiction de la récurrence du CPNPC au stade précoce (I-II) à l'aide de données clinicopathologiques. Les auteurs ont inclus des caractéristiques telles que les données démographiques des patients, les caractéristiques des tumeurs, les biomarqueurs, y compris PD-L1, et les comorbidités. Ils ont également inclus les détails de la chirurgie et les caractéristiques des autres traitements tels que la chimiothérapie (régime, type, heure de début). Les auteurs rapportent une précision de 76 % obtenue avec le modèle de RF et une précision de 68 % obtenue avec le modèle de graphe. Le modèle proposé pourrait potentiellement servir d'outil d'assistance pour guider l'utilisation des traitements adjuvants dans le cancer du poumon à un stade précoce [127].

L'accessibilité des données cliniques et leur formatage normalisé constituent l'un des principaux avantages par rapport aux données d'imagerie ou de génomique. En outre, les modèles basés sur les caractéristiques cliniques sont plus rentables, car ils exploitent les dossiers existants sans nécessiter de procédures supplémentaires coûteuses. En outre, les analyses d'interprétabilité peuvent être réalisées plus facilement sur ces modèles, ce qui facilite leur intégration clinique. Néanmoins, les données cliniques peuvent ne pas rendre compte de toute la complexité de la biologie ou du microenvironnement tumoral et peuvent être sujettes à des biais potentiels introduits par le clinicien déclarant.

Bien que chaque approche ait ses mérites, l'intégration des données cliniques, d'imagerie et génomiques peut s'avérer plus efficace pour développer des modèles prédictifs complets reposant sur les caractéristiques cliniques et subcliniques du cancer.

3.4 Approches multimodales d'apprentissage profond

Les approches multimodales combinant l'apprentissage automatique et diverses sources de données sont de plus en plus utilisées dans la recherche sur le cancer et l'oncologie clinique [5]. Ces approches intègrent des informations provenant de modalités multiples afin d'améliorer le diagnostic, le pronostic et la planification du traitement du cancer (Figure 3.2).

Dans le contexte de la classification et du sous-typage des cancers, une étude a combiné des images histologiques avec des données de séquençage de l'ARN pour la classification des sous-types histologiques du CPNPC. Les auteurs ont formé un modèle SVM sur des données d'expression génique sélectionnées à partir du séquençage de l'ARN en utilisant la méthode de la redondance minimale et de la pertinence maximale (minimum redundancy maximum relevance). Ensuite, un modèle ResNet-18 a été entraîné à l'aide de tuiles d'images (patches)

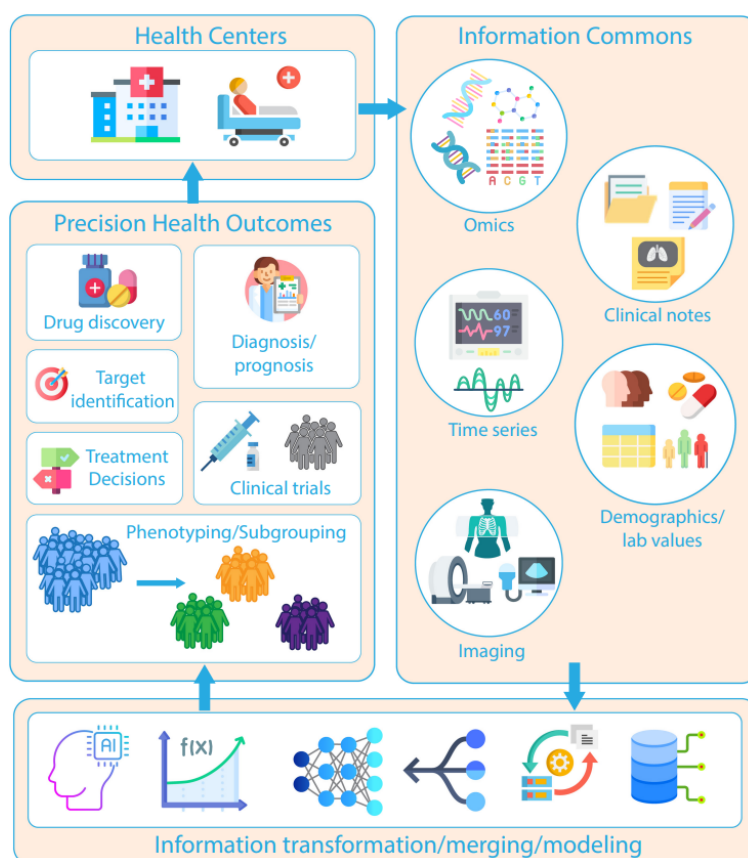


FIGURE 3.2 Les informations multimodales, y compris l'imagerie (TDM, IRM, TEP, histopathologie, etc.), les données omiques (transcriptomique, génomique, etc.) et les données cliniques, sont exploitées dans des approches de modélisation algorithmique. Ces algorithmes offrent des perspectives sur de nombreux résultats de santé différents, tels que le diagnostic du cancer, son pronostic et la découverte de biomarqueurs. Ces informations peuvent être utilisées comme outils d'assistance par les centres de santé et les oncologues pour fournir des soins aux patients plus efficaces et personnalisés. Source [5]

de lames entières et un vote majoritaire a été effectué pour dériver des prédictions au niveau des lames. Le modèle intégratif obtenu par une approche de fusion tardive a permis une amélioration de 24 % par rapport au classificateur d'expression génique seul et de 64 % par rapport au classificateur d'histologie seul [128].

Plusieurs approches multimodales ont été conçues pour prédire le pronostic du cancer, y compris la prédiction des biomarqueurs, de la survie des patients, de la progression ou de la récurrence de la maladie et de la réponse au traitement. Ces travaux comprennent la combinaison de plusieurs modalités d'imagerie en tant que données d'entrée pour les modèles d'apprentissage automatique, et l'incorporation de différentes données omiques, y compris la radiomique, la transcriptomique, la génomique, la pathomique et les données cliniques.

Une étude comportant 447 patients a utilisé l'algorithme de Boruta pour sélectionner les caractéristiques étroitement associées aux mutations du gène KRAS, suivi d'une régression logistique multivariée pour développer des modèles radiomiques, cliniques et combinés clinique-radiomique [129]. Le modèle de fusion clinique-radiomique intégré a démontré de bonnes performances avec une AUC de 0,755 dans la cohorte de validation, ce qui suggère que ces modèles pourraient servir de marqueurs d'imagerie potentiels pour la détection pré-opératoire du statut de mutation du gène KRAS.

Une autre étude a testé les avantages de la combinaison de la radiomique et des caractéristiques cliniques du MHCR pour la prédiction par TDM de la progression locale de la tumeur après une thérapie d'ablation thermique. L'étude a été réalisée sur 90 patients et a montré que la combinaison de la radiomique et des caractéristiques cliniques dans un modèle de risques proportionnels de Cox a conduit à une amélioration de 1,28 % par rapport au modèle radiomique seul en termes d'indice de concordance sur l'ensemble de validation [130]. Néanmoins, la généralisation des modèles développés est cruciale pour leur incorporation dans la pratique clinique. Une étude a souligné l'importance d'une validation indépendante en exploitant trois modèles précédemment publiés pour la prédiction de la progression locale de la tumeur par TDM. Les auteurs ont indiqué que les résultats obtenus à l'aide des modèles cliniques et radiomiques n'étaient pas reproductibles dans le cadre d'une validation interne et externe indépendante [131].

En ce qui concerne la tâche de prédiction du PD-L1 dans le CPNPC, une étude a montré que l'incorporation d'informations métaboliques complémentaires provenant de la TEP et de données anatomiques provenant du TDM améliorait le pouvoir prédictif du modèle, en saisissant à la fois les aspects structurels et fonctionnels de la biologie tumorale [132]. Une autre étude a montré que la combinaison d'images TEP et TDM permettait d'obtenir de meilleures performances pour la classification de PD-L1, quel que soit le stade de fusion (fusion précoce ou tardive) [133]. Liang et al. ont proposé un pipeline multimodal combinant la TEP/TDM et les données cliniques pour réaliser la même tâche. Les auteurs ont montré que la combinaison des deux modalités d'imagerie était bénéfique et que l'incorporation des caractéristiques cliniques améliorait les performances de prédiction du PD-L1 par rapport à la TEP/TDM seule [134].

Plusieurs travaux ont été réalisés pour prédire la récurrence post-chirurgicale du CPNPC à l'aide du TDM, de la TEP, des caractéristiques cliniques et d'une combinaison de ces éléments [135]. Une étude a combiné des caractéristiques radiomiques extraites d'images de TDM et des caractéristiques démographiques pour prédire la récurrence du CPNPC de stade I-IIIa réséqué chirurgicalement et a atteint une AUC de 61,4 % [136]. Une autre étude a combiné

les caractéristiques radiomiques du TDM et de la TEP avec des données cliniques (stade du cancer) pour la prédiction de la récurrence post-chirurgicale du CPNPC. Les auteurs ont obtenu un indice de concordance de 0,76 en entraînant un modèle de survie aléatoire sur les caractéristiques multimodales [137].

En ce qui concerne la prédiction de la réponse pathologique majeure à la chimio-immunothérapie néoadjuvante chez les patients atteints de CPNPC résécable, She et al. ont formé un pipeline d'apprentissage profond basé sur un RNC à l'aide d'images de TDM, atteignant une AUC de 0,72 dans la validation externe. En combinant le modèle d'imagerie avec un modèle clinique en utilisant la régression logistique, les auteurs ont atteint une AUC de 0,76 pour la classification de la réponse pathologique majeure dans la validation externe. Ces résultats révèlent le potentiel des caractéristiques combinées imagerie-clinique-pathologie dans la prédiction de la réponse à la chimio-immunothérapie. Néanmoins, l'étude se concentre uniquement sur la réponse pathologique majeure, ce qui ne décrit pas complètement le résultat des patients après le traitement. L'extension de ces études pour couvrir la progression de la maladie et la survie des patients est donc justifiée pour évaluer de manière complète l'efficacité prédictive de l'approche proposée [138].

Plusieurs études combinant les caractéristiques radiomiques et transcriptomiques ont également été réalisées pour prédire le pronostic du CPNPC. Bouhamama et al. ont combiné des caractéristiques radiomiques dérivées du TDM avec des données de séquençage de l'ARN pour prédire la progression du TDM trois mois après le traitement par ICI. Ils ont également testé la combinaison des caractéristiques radiomiques avec un statut immunitaire chaud/froid, basé sur une signature d'expression de 27 gènes.

Grossman et al. [139] ont montré que la combinaison des données radiomiques, de l'expression génique et des caractéristiques cliniques du CPNPC dans un modèle de risques proportionnels de Cox a amélioré la prédiction de la SG par rapport à l'utilisation individuelle de ces types de caractéristiques ou de la combinaison de deux d'entre eux, mettant en évidence la nature complémentaire de ces données.

Une étude menée sur 220 patients atteints de CPNPC a mis en évidence les avantages de la combinaison de différents types de données pour prédire le pronostic du cancer du poumon. En intégrant la radiomique, la pathomique, la dosiomique et les caractéristiques cliniques basées sur le TDM, les auteurs ont obtenu une AUC de 0,91 pour la prédiction de la réponse du CPNPC à la chimioradiothérapie et de 0,80 pour la prédiction de la SG à un an [6] (Figure 3.3). Starke et al. [140] ont proposé une approche multimodale combinant le TDM et les images FDG-TEP acquises à plusieurs niveaux pour la prédiction de la récurrence loco-régionale du cancer localement avancé de la tête et du cou traité par chimioradiothérapie. Une autre

étude visait à prédire la réponse du CPNPC à l'ICI en intégrant les données radiomiques, pathomiques et génomiques du TDM dans un pipeline proposé appelé DyAM [141]. Les auteurs ont montré que la combinaison des caractéristiques radiologiques (radiomique) et génomiques (altération des gènes et charge mutationnelle de la tumeur) permettait d'obtenir la performance bimodale la plus élevée ($AUC=0,76$). Ils ont également montré que l'utilisation des trois modes de données avec le score PD-L1 permettait d'obtenir la plus grande précision avec une AUC de 0,80.

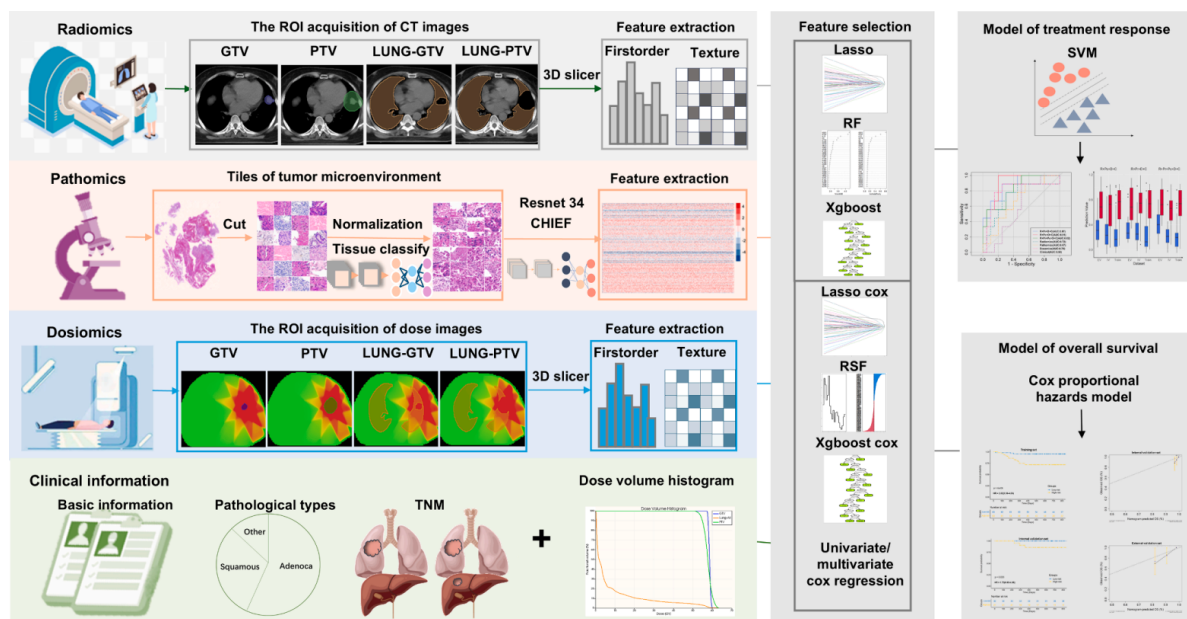


FIGURE 3.3 L'approche multi-omique adoptée par Pan et al. [6] pour la prédiction de la réponse du CPNPC à la chimio-radiothérapie et la SG à 1 an, intégrant des données radiomiques, pathomiques, dosiométriques et cliniques.

Captier et al. [142] ont combiné des données cliniques, histopathologiques, TEP/TDM et transcriptomiques pour la prédiction de la SG et de la progression du CPNPC métastatique après une immunothérapie de première ligne. Les auteurs ont comparé les performances des modèles linéaires (régression logistique et modèle des risques proportionnels de Cox) et des modèles arborescents (XGB et Random Survival Forest). Ils ont également comparé les approches de fusion précoce ou tardive des caractéristiques. La supériorité des approches multimodales a été confirmée par la comparaison des performances moyennes des différentes stratégies d'intégration pour toutes les combinaisons possibles d'une, deux, trois et quatre modalités. Cette étude souligne l'importance capitale de la combinaison de données multimodales pour prédire le pronostic du cancer. Néanmoins, les quatre modalités n'étaient disponibles que pour 80 patients dans cette étude et tous les patients ont été recrutés dans le même centre. Par conséquent, des études à grande échelle incluant plusieurs centres médicaux

sont nécessaires pour l'intégration des modèles multimodaux dans le flux de travail clinique.

3.5 Techniques d'apprentissage profond pour la recherche sur le cancer

L'apprentissage profond a transformé la recherche sur le cancer en permettant une analyse avancée de données biomédicales complexes. Dans cette section, nous présentons un aperçu des architectures d'apprentissage profond déployées dans la recherche sur le cancer, en mettant particulièrement l'accent sur les modèles basés sur les transformeurs qui ont démontré des performances remarquables à travers plusieurs types de données sur le cancer, y compris les images, les séries temporelles et les données tabulaires.

Les premières applications de l'apprentissage profond dans la recherche sur le cancer ont principalement exploité les réseaux neuronaux, en particulier les RNC pour les images médicales et les réseaux neuronaux récurrents pour les données de séries temporelles. Les RNC sont devenus la pierre angulaire de l'analyse d'images médicales en raison de leur capacité à extraire des caractéristiques hiérarchiques à partir des images radiologiques et histopathologiques. Malgré leur succès, les architectures basées sur les RNC ont rencontré des limitations pour capturer les dépendances à long terme.

Le modèle transformeur a révolutionné l'analyse des données de séries temporelles en remplaçant les opérations de récurrence et de convolution par des mécanismes d'attention. En exploitant le mécanisme d'auto-attention, le transformeur évalue simultanément l'importance des différents composants d'entrée, quel que soit leur éloignement positionnel. Cette caractéristique le rend particulièrement adapté à la recherche sur le cancer, où les relations entre des caractéristiques éloignées revêtent souvent une signification diagnostique.

Le transformeur de vision (ViT) [143] s'inspire de l'architecture du transformeur initialement conçue pour les tâches de traitement du langage naturel [144]. Ce dernier a obtenu des résultats de pointe en traduction automatique, ce qui a motivé les chercheurs à l'adapter pour les applications de vision par ordinateur.

Un transformeur traditionnel combine un encodeur et un décodeur, tous deux incorporant un mécanisme d'attention. Cependant, le ViT ne comprend que la partie encodeur, et le décodeur est remplacé par un perceptron multicouche qui effectue la prédiction de classe. Dans le ViT, chaque image est divisée en patches qui sont ensuite aplatis et envoyés à une fonction de projection linéaire avec des paramètres entraînaables. Les embeddings de classe résultants sont précédés par un token de classe. Des embeddings positionnels sont ensuite ajoutés afin de préserver les informations spatiales. Les embeddings obtenus sont envoyés à l'encodeur du transformeur qui traite simultanément les entrées à l'aide d'un mécanisme

d’attention et génère un vecteur d’attention pour chaque entrée. La tâche de classification est finalement effectuée par le perceptron multicouche qui prédit la classe de l’image en fonction du vecteur d’attention de l’embedding de classe.

Le module le plus important de l’encodeur du transformeur est le bloc d’attention multi-têtes (AMT). Dans ce bloc, trois paramètres entraînaables sont attribués à chaque entrée : (1) une requête q qui peut être interprétée comme une question posée à propos de l’entrée donnée (patch d’image dans le ViT), (2) une clé k qui aide à trouver, parmi toutes les autres entrées, celle qui a la plus grande similarité avec la requête, et (3) une valeur v associée à l’entrée. Un bloc d’attention à tête unique (ATU) peut être décrit par l’équation suivante :

$$\text{ATU} = \text{softmax} \left(\frac{QK^T}{s} \right) V \quad (3.1)$$

Où les requêtes, clés et valeurs de toutes les entrées sont combinées respectivement dans les matrices Q , K et V , et s est un facteur d’échelle. Dans l’AMT, l’opération ATU est répétée plusieurs fois et les vecteurs d’attention résultants sont concaténés et projetés linéairement.

De nombreuses variantes avancées de transformeurs ont émergé dans la littérature, avec des applications dans la recherche sur le cancer. TransMed est un exemple d’architecture hybride combinant les transformeurs et les RNC pour capturer à la fois les relations globales et les caractéristiques locales [145]. TransUNet combine également l’extraction de caractéristiques basée sur les RNC avec la modélisation du contexte global basée sur les transformeurs, suivie d’opérations de suréchantillonnage pour les tâches de segmentation des tumeurs [146]. Une autre variante améliorée des transformeurs est le SwinTransformers, qui effectue une extraction de caractéristiques multi-échelles, se montrant particulièrement efficace pour détecter des lésions de tailles variables [147].

Les architectures basées sur les transformeurs ont démontré des performances supérieures dans les analyses d’imagerie, de génomique et multimodales. Néanmoins, certaines architectures avancées de transformeurs nécessitent des ressources computationnelles substantielles, en particulier pour le traitement d’images médicales à haute résolution. De plus, bien que les transformeurs excellent généralement avec de grands ensembles de données, les applications en oncologie font souvent face à une rareté des données, nécessitant des stratégies de pré-entraînement spécialisées et des techniques d’augmentation des données.

3.5.1 Apprentissage profond pour les données tabulaires

Bien que l'apprentissage profond ait montré un succès remarquable dans le traitement des données non structurées telles que les images et les textes, les données tabulaires, largement utilisées dans le domaine de la santé où les données cliniques structurées sont abondantes, ont traditionnellement été analysées à l'aide de méthodes basées sur les arbres telles que les RF et les algorithmes de boosting de gradient. Ces méthodes ont tendance à surpasser les réseaux neuronaux profonds conventionnels pour les données tabulaires.

En effet, le traitement des données tabulaires présente des défis uniques en raison de leur nature hétérogène et de l'absence de structure spatiale. Les récentes avancées intégrant des mécanismes d'attention, des embeddings de caractéristiques et des approches hybrides ont conduit à des architectures spécialisées conçues pour traiter efficacement les données tabulaires. Ces architectures répondent aux défis associés au traitement des données tabulaires tout en préservant l'explicabilité et l'efficacité computationnelle, ce qui les rend particulièrement adaptées aux applications médicales où la fiabilité et l'interprétabilité sont cruciales [148].

Le modèle d'Apprentissage Attentif Interprétable Tabulaire (TabNet) est une innovation importante dans le traitement des données tabulaires [149]. Son architecture repose sur des réseaux neuronaux, des mécanismes d'attention avec des stratégies d'ensemble pour améliorer les performances tout en préservant une grande interprétabilité. Le mécanisme d'attention séquentielle permet une sélection instance par instance des caractéristiques les plus saillantes, tout en maintenant l'interprétabilité et une meilleure capacité d'apprentissage.

Les architectures basées sur le transformeur pour les données tabulaires, adaptées du traitement du langage naturel et de la vision par ordinateur, utilisent le mécanisme d'attention pour capturer des associations complexes entre les caractéristiques, ce qui les rend bien adaptées aux données médicales, notamment la génomique et les données cliniques où des interactions entre variables sont souvent impliquées et non linéaires. Le modèle TabTransformer convertit les caractéristiques catégorielles en représentations contextuelles en les traitant avec des couches de transformeurs, comprenant les mécanismes d'AMT, avant de les combiner avec les caractéristiques numériques [150].

Le modèle Feature Tokenizer Transformer (FT-Transformer) traite à la fois les caractéristiques catégorielles et numériques comme des tokens et applique des mécanismes d'attention pour capturer les interactions impliquées [151]. Le FT-Transformer a généralement surpassé les modèles d'apprentissage machine traditionnels et le TabTransformer à plusieurs niveaux de classification, établissant une base solide pour les approches basées sur l'apprentissage

profond dans divers domaines, en particulier dans le domaine médical où les caractéristiques sont fortement corrélées.

3.5.2 Modèles génératifs

Principes des réseaux antagonistes génératifs

Les réseaux antagonistes génératifs (Generative Adversarial Networks, GAN) permettent de générer des données synthétiques en utilisant un entraînement antagoniste [152]. Les deux principaux composants du modèle GAN sont le générateur et le discriminateur. Le premier a pour objectif de générer des données synthétiques qui suivent la même distribution que les données réelles, tandis que le rôle du second est de distinguer entre les instances réelles et celles générées artificiellement. Les deux modèles sont entraînés de manière itérative à l'aide d'une fonction de perte min-max. Le discriminateur est entraîné pour maximiser la probabilité de classer correctement les instances comme réelles ou générées artificiellement, tandis que le générateur est entraîné pour minimiser cette probabilité.

Les GAN sont particulièrement utiles dans le domaine médical, où les données annotées sont rares et où les caractéristiques présentent des relations complexes. En apprenant la distribution complexe des données, les GAN peuvent être utilisés dans la recherche biomédicale pour la génération d'images médicales réalistes qui pourraient être exploitées pour augmenter les jeux de données limités afin d'entraîner des modèles diagnostiques et pronostiques ou pour des tâches de segmentation médicale [153]. Parmi les applications, on trouve la génération d'images de mammographie, l'amélioration des images médicales de faible qualité et la normalisation des images pathologiques [154, 155].

Génération de données tabulaires

Alors que les premières applications biomédicales des GAN ont émergé et gagné en importance dans la génération d'images médicales, l'utilisation des GAN sur des données médicales tabulaires offre des opportunités uniques. Par exemple, les GAN pourraient être utilisés pour générer des dossiers de santé électroniques tabulaires, surmontant ainsi les préoccupations liées au partage de données et à la confidentialité [156]. Des modèles spécialisés sont nécessaires pour les données tabulaires, car elles présentent des défis uniques tels que les différents types de caractéristiques, les distributions hétérogènes et les relations complexes entre les caractéristiques.

Parmi ces modèles, le modèle tableGAN utilise des architectures de RNC pour le générateur et le discriminateur, et inclut un classificateur ayant la même architecture que le discrimi-

nateur, dont le rôle est d’assurer l’intégrité sémantique des données générées. Il est donc entraîné à l’aide des étiquettes de vérité terrain issues des données réelles [157]. Le med-GAN gère les variables binaires et continues en utilisant un autoencodeur qui mappe les dossiers de patients de haute dimension dans un espace latent. Le générateur est donc entraîné à synthétiser des représentations latentes à partir de bruit aléatoire, qui sont ensuite envoyées au décodeur pour reconstruire les données dans l’espace réel [158]. Le réseau antagoniste génératif tabulaire conditionnel (Conditional Tabular Generative Adversarial Network, CTGAN) utilise un générateur conditionnel pour préserver les relations entre les variables (distributions conditionnelles), ce qui donne des données synthétiques de bonne qualité tout en conservant les propriétés statistiques du jeu de données d’origine. Il effectue également un rééchantillonnage efficace durant le processus d’entraînement, de manière à permettre un échantillonnage uniforme de toutes les classes des colonnes discrètes. Cela permet de prévenir l’effondrement de mode (mode collapse) et le problème du déséquilibre des classes dans les colonnes discrètes [159].

Bien que chacun de ces modèles présente ses propres avantages pour le traitement des données tabulaires, des analyses comparatives ont montré que le CTGAN offre des performances supérieures [159].

Imputation tabulaire générative

Le problème des données manquantes est particulièrement courant dans le domaine médical en raison de la rareté des jeux de données annotées de haute qualité. Cela peut potentiellement introduire un biais et compromettre la validité des analyses d’apprentissage automatique en aval. Une des façons dont les données manquantes peuvent impacter la validité d’une étude est en réduisant la puissance statistique et en augmentant la probabilité d’erreurs de type II, définie comme la probabilité de détecter un effet significatif lorsqu’il existe, en raison de la réduction de la taille de l’échantillon [160]. Le besoin d’approches fiables d’imputation va au-delà des jeux de données cliniques et s’étend à diverses données omiques, y compris la génomique, la protéomique, la transcriptomique et l’imagerie médicale.

Les méthodes conventionnelles d’imputation, telles que les techniques d’imputation par zéro, moyenne et k-plus proches voisins (k-nearest neighbors), ne tiennent pas compte des relations complexes entre les caractéristiques en raison des hypothèses simplifiées concernant les mécanismes de données manquantes. Les approches génératives antagonistes sont bien positionnées pour résoudre efficacement ce problème, car elles peuvent apprendre des distributions de données complexes sans se baser sur une modélisation explicite des probabilités. Le Réseau d’imputation antagoniste générative (Generative Adversarial Imputation Network,

GAIN) est apparu comme un outil prometteur pour l'imputation des données tabulaires en adaptant le GAN standard spécifiquement pour imputer les données manquantes [8]. En se basant sur la distribution des données, GAIN utilise un générateur qui propose des imputations plausibles pour les données manquantes. Le discriminateur vise à distinguer entre les valeurs observées et les valeurs imputées, plutôt qu'entre des échantillons réels et synthétiques comme dans les GAN classiques. Les résultats récents montrent que GAIN surpasse les autres techniques d'imputation sur de grands jeux de données réels et qu'il est plus résistant aux taux élevés de données manquantes [161].

3.6 Paradigmes d'apprentissage avancés

3.6.1 Apprentissage auto-supervisé et semi-supervisé

L'apprentissage auto-supervisé est une approche qui permet aux modèles d'apprendre des représentations significatives à partir de données non étiquetées sans avoir besoin de labels. Cette méthodologie permet aux modèles d'apprendre de manière autonome des caractéristiques inhérentes riches. Les représentations apprises peuvent être utilisées pour des tâches en aval [162]. SimCLR est une mise en œuvre récente d'une approche auto-supervisée qui utilise l'apprentissage contrastif pour extraire des représentations indépendantes du bruit pour des tâches supervisées ultérieures [163].

L'apprentissage semi-supervisé utilise à la fois des données étiquetées et non étiquetées. Il est particulièrement utile lorsque les données non étiquetées sont abondantes alors que les données étiquetées sont limitées. En combinant les paradigmes d'apprentissage supervisé et non supervisé, les approches semi-supervisées permettent de maximiser l'utilité des données étiquetées [164]. L'algorithme du mean teacher est une approche d'apprentissage semi-supervisé qui utilise un modèle étudiant entraîné avec des données étiquetées et non étiquetées, et un modèle enseignant dont les poids sont mis à jour en tant que moyennes mobiles (moving averages) des poids de l'étudiant et utilisés pour générer des pseudotags afin d'entraîner le modèle étudiant [165].

L'auto-apprentissage est une technique d'apprentissage semi-supervisé qui commence par entraîner un modèle sur un ensemble de données étiquetées limité. Le modèle entraîné est ensuite utilisé pour générer des prédictions, appelées pseudotags, pour des instances non étiquetées. Les pseudotags générés avec une forte confiance sont intégrés en tant que nouveaux exemples d'entraînement et le modèle est réentraîné de manière itérative sur des instances étiquetées et pseudotagées.

Un défi important de l'auto-apprentissage est d'éviter la propagation des erreurs en s'assurant

que des pseudotags de haute qualité sont utilisés pour réentraîner le modèle. Une solution pourrait être d'appliquer le pseudotagging uniquement sur les instances pour lesquelles le modèle était le plus confiant, car la plupart des prédictions erronées faites par le classificateur se produisent près de la frontière de décision [166]. Dans ce cas, le pseudotag est défini comme suit :

$$\hat{y} = \begin{cases} 1 & , \text{ if } f(c) > \delta^+ \\ 0 & , \text{ if } f(c) < \delta^- \\ \emptyset & , \text{ otherwise} \end{cases} \quad (3.2)$$

où $\hat{y}$ est le pseudotag, $f(c)$ est la sortie probabiliste du modèle avec une instance non étiquetée c , et δ^+ et δ^- sont respectivement les seuils des classes positives et négatives.

3.6.2 Modèles de fondation dans le domaine biomédical

Les modèles de fondation constituent un changement de paradigme dans l'apprentissage profond appliqué aux soins de santé. Ces modèles vastes se caractérisent par un grand nombre de paramètres et sont entraînés sur un grand nombre de jeux de données divers. Cela leur permet de démêler des informations médicales complexes et de développer une compréhension approfondie des relations contextuelles dans les données [167]. De plus, les modèles de fondation sont souvent entraînés à l'aide de stratégies d'apprentissage auto-supervisé, ce qui leur permet d'extraire des représentations pertinentes à partir de données non étiquetées. Cela est particulièrement utile dans le domaine médical, où les grands ensembles de données annotées sont souvent rares. Les modèles qui en résultent peuvent être utilisés comme base pour une large gamme de tâches en aval, permettant ainsi de surmonter le besoin de jeux de données annotés suffisamment grands pour l'entraînement de modèles spécifiques à des tâches. Les modèles de fondation ont été conçus en utilisant différents types de données médicales, y compris des séries temporelles telles que les électrocardiogrammes [168] et la documentation clinique, où les grands modèles de langage peuvent être utilisés pour la réponse à des questions médicales [169].

Dans l'analyse d'images médicales, où les approches traditionnelles reposent sur l'apprentissage supervisé nécessitant des données annotées par des experts, les modèles de fondation ont montré des capacités uniques à apprendre des représentations généralisables. Leurs applications incluent la segmentation, la détection de tumeurs, la classification et la prédiction de la réponse au traitement [170].

Dans les tâches de segmentation, les modèles de fondation ont atteint de meilleures perfor-

mances que les approches supervisées lorsqu'ils ont été entraînés avec des données limitées. Le modèle MedSAM, une adaptation du Segment Anything Model (SAM) conçu pour les images naturelles, a été conçu pour la segmentation universelle des images médicales. Le modèle utilise un encodeur d'images basé sur l'architecture ViT, un encodeur de commandes d'utilisateurs (user prompt) et un décodeur de masques. Entraîné sur plus de 1,57 million de paires image-masque, incluant 30 types de cancers et 10 modalités d'imagerie, dont les images de TDM, MedSAM a systématiquement surpassé le modèle SAM et atteint des résultats équivalents ou meilleurs que le modèle spécialisé U-Net [171].

La découverte de biomarqueurs et le soutien à la décision clinique est une autre application des modèles de fondation dans l'analyse d'images médicales. Un modèle de fondation, entraîné sur 11 467 lésions radiographiques provenant de divers sites à l'aide de l'apprentissage contrastif auto-supervisé d'un encodeur convolutif, a montré des propriétés diagnostiques et pronostiques [7] (Figure 3.4). Le modèle a démontré des capacités dans la caractérisation non invasive des tumeurs, car il a été validé sur différentes applications cliniquement pertinentes, y compris la classification de la malignité et la prédiction de la SG du CPNPC. Le modèle de fondation a également surpassé les approches supervisées et d'autres modèles d'état de l'art pré-entraînés sur les tâches susmentionnées. De plus, les caractéristiques d'imagerie identifiées par le modèle de fondation ont montré des associations plus fortes avec les profils d'expression génique que celles dérivées des méthodes conventionnelles.

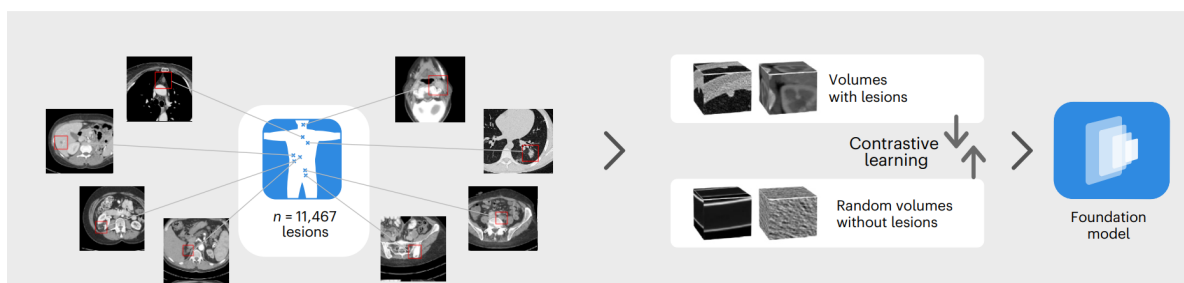


FIGURE 3.4 L'entraînement préalable du modèle de fondation : un encodeur convolutif profond a été préalablement entraîné en utilisant l'apprentissage contrastif basé sur des volumes avec et sans lésions [7]

3.7 Résumé

Cette revue de littérature explore l'application de l'apprentissage automatique dans la recherche sur le cancer, en se concentrant sur diverses modalités de données et des techniques avancées d'apprentissage automatique. La revue commence par examiner l'apprentissage automatique dans l'imagerie du cancer pour le diagnostic et la pronostication, en incluant à

la fois des applications directes et l'utilisation de l'imagerie comme substitut pour les cibles thérapeutiques et les marqueurs immunitaires. Elle approfondit ensuite les applications de l'apprentissage automatique en génomique du cancer, avec un accent particulier sur la pronostication et le sous-typage basés sur la transcriptomique. La revue aborde également les modèles d'apprentissage automatique basés sur des caractéristiques cliniques pour la pronostication du cancer, mettant en évidence leur potentiel pour améliorer les soins aux patients.

Une part importante de la revue est consacrée aux approches multimodales qui combinent l'imagerie, la génomique et les données cliniques, mettant en valeur le potentiel d'intégration de divers types de données pour améliorer la performance prédictive. La revue se concentre ensuite sur les algorithmes d'apprentissage profond dans la recherche sur le cancer, couvrant les transformeurs, l'apprentissage profond tabulaire et les modèles génératifs. Une attention particulière est accordée aux modèles génératifs dans l'apprentissage tabulaire, en explorant leur utilisation dans la génération de données synthétiques et l'imputation de valeurs manquantes.

La revue discute également des approches auto-supervisées, semi-supervisées et l'auto-apprentissage, abordant le défi des données étiquetées limitées dans la recherche médicale. Enfin, elle introduit le concept de modèles de fondation dans le domaine médical, soulignant leur potentiel à révolutionner la recherche sur le cancer et la pratique clinique.

Bien que la revue souligne l'évolution rapide et le potentiel de l'apprentissage automatique dans la recherche sur le cancer, elle met également en évidence les défis et limitations importants qui persistent. Ceux-ci incluent la variabilité des performances des modèles d'apprentissage automatique selon les types de cancer. La revue note que les applications de l'apprentissage automatique pour la pronostication des MHCR après résection curative sont particulièrement limitées, ce qui indique un besoin de recherches plus ciblées dans ce domaine. Dans le contexte du CPNPC, la revue souligne le manque d'études évaluant la généralisabilité des modèles de prédiction de la réponse à l'immunothérapie à travers différents stades de la maladie. De plus, elle aborde les défis liés à la gestion des données transcriptomiques dans les modèles d'apprentissage automatique, notamment les tailles d'échantillons limitées et la présence de transcrits non détectés, ce qui peut compliquer l'analyse et le développement des modèles. La revue touche également des problèmes de qualité des données, la nécessité de grandes bases de données diversifiées pour l'entraînement et la validation des modèles, ainsi que la nature "boîte noire" de nombreux algorithmes avancés d'apprentissage automatique, ce qui peut entraver l'interprétabilité et la confiance clinique.

CHAPITRE 4 MÉTHODOLOGIE DE RECHERCHE

4.1 Énoncé de la problématique

Le cancer demeure une cause majeure de mortalité dans le monde, avec les MHCR et le CPNPC posant des défis cliniques importants en raison de réponses hétérogènes aux traitements et de risques de récurrence variables. Un nombre significatif de patients sont confrontés à des complications potentielles, des effets secondaires et des toxicités associées à des traitements qui ne tiennent pas compte du pronostic attendu, sans bénéfices sur la survie. Il existe donc un besoin urgent de biomarqueurs permettant d'identifier les patients à haut risque de progression ou de récurrence de la maladie.

Dans le cas des patients atteints de MHCR, il manque des biomarqueurs fiables pour identifier ceux susceptibles de présenter une récurrence précoce après une résection chirurgicale, pour lesquels la chirurgie pourrait être inutile, qui pourraient bénéficier d'une chimiothérapie adjuvante, et qui devraient être suivis de manière plus rapprochée en raison du risque de récurrence. En ce qui concerne le CPNPC, les ICI sont devenus une approche thérapeutique clé pour les patients atteints de CPNPC avancé, et des essais cliniques récents ont également démontré leur efficacité pour le traitement du CPNPC résécable en situation néoadjuvante. Cependant, le bénéfice clinique de ces agents est limité à un sous-ensemble de patients, moins de la moitié d'entre eux ayant des réponses durables.

Alors que des biomarqueurs tels que le CD73 dans les MHCR, les infiltrats immunitaires et le PD-L1 dans le CPNPC montrent une valeur pronostique, l'évaluation actuelle nécessite une analyse histopathologique du tissu tumoral, une approche invasive qui est difficile à répéter de manière longitudinale. Dans le cas des MHCR, l'évaluation histopathologique est obtenue uniquement après la résection. Les marqueurs basés sur l'histologie ne sont également pas idéaux pour l'évaluation préopératoire par biopsie en raison de leur expression hétérogène au sein de la tumeur. De plus, les biopsies sont associées à des défis et limitations procéduraux tels que les différences entre les tumeurs primaires et les métastases, la variabilité de la qualité des échantillons et les incohérences dans les protocoles de coloration. L'incapacité à surveiller de manière non invasive l'expression des biomarqueurs limite l'adaptation dynamique du traitement. Par conséquent, la première question de recherche qui pourrait être abordée est la suivante : Les caractéristiques radiomiques dérivées du TDM peuvent-elles servir de proxys pronostiques non invasifs pour les biomarqueurs des MHCR et du CPNPC, surmontant ainsi les limitations des prélèvements tissulaires ?

Bien que l'imagerie par TDM contienne des informations biologiques latentes précieuses, les approches radiomiques peuvent être insuffisantes pour capturer pleinement la complexité de la biologie tumorale. Les données transcriptomiques peuvent saisir les facteurs biologiques de la progression et de la récurrence du cancer, y compris les caractéristiques du microenvironnement tumoral. Néanmoins, des défis importants doivent être relevés, notamment la gestion des transcrits d'ARN non détectés et la prise en charge de la haute dimensionnalité des données de séquençage de l'ARN. De plus, les préoccupations concernant l'interprétabilité des modèles doivent être résolues avant que ces approches puissent être efficacement mises en œuvre dans la pratique clinique. La deuxième question de recherche peut donc être formulée comme suit : Quelles stratégies d'apprentissage profond peuvent analyser efficacement les données transcriptomiques tout en préservant l'interprétabilité clinique ?

Les modèles existants sous-exploitent souvent la valeur pronostique complémentaire des données radiologiques, transcriptomiques et cliniques. Cela est particulièrement vrai dans le cas de la prédiction de la récurrence précoce des MHCR, qui manque d'approches multimodales pour prédire la récurrence précoce après résection chirurgicale. Cela met en évidence le besoin d'une recherche plus ciblée dans ce domaine. Ainsi, la question suivante se pose : L'intégration des données TDM préopératoires, du séquençage de l'ARN et des variables cliniques améliore-t-elle la prédiction de la récurrence précoce des MHCR par rapport aux techniques de pronostication conventionnelles ?

4.2 Hypothèses

En se basant sur les problèmes exposés dans la section précédente, les hypothèses suivantes sont proposées.

Hypothèse 1. Les images de TDM de pré-traitement contiennent des informations intrinsèques reflétant les niveaux d'expression protéiniques des marqueurs immunitaires CD73 et CD3/CMH-I dans le microenvironnement tumoral des MHCR, ainsi que ceux de PD-L1 dans le CPNPC.

Hypothèse 2. Les substituts d'imagerie non invasifs des biomarqueurs CD73 et CD3/MHC-I des MHCR et PD-L1 du CPNPC conservent la valeur pronostique de leurs homologues histologiques.

Hypothèse 3. Les données transcriptomiques et cliniques contiennent des informations pronostiques complémentaires à celles des biomarqueurs d'imagerie dans la prédiction de la récurrence précoce des MHCR.

Hypothèse 4. Un pipeline computationnel multi-omique, combinant des images de TDM,

des données transcriptomiques et cliniques, permet de prédire de manière fiable la récurrence post-chirurgicale des MHCR.

4.3 Objectifs

L'objectif global est le développement de pipelines computationnels et de stratégies d'intégration multi-omiques pour améliorer le pronostic du cancer, la personnalisation des traitements et la découverte de biomarqueurs à travers des approches non invasives.

En utilisant des modèles d'apprentissage profond basés sur l'attention, ainsi que des données multi-omiques, à savoir des images de TDM, des microréseaux de tissus, du séquençage d'ARN et des données cliniques, nous effectuons une caractérisation du microenvironnement tumoral immunitaire. Cela offre aux oncologues une évaluation biologique préliminaire des tumeurs à partir de données médicales largement disponibles, leur permettant d'effectuer une évaluation pronostique plus éclairée et de sélectionner un traitement plus adapté.

Les objectifs spécifiques qui seront abordés sont donc les suivants (Figure 4.1) :

Objectif 1. Concevoir, mettre en œuvre et évaluer un pipeline pour obtenir, à partir des images de TDM de pré-traitement, des substituts non invasifs de biomarqueurs immunitaires dans les MHCR ayant une valeur pronostique, notamment l'expression de CD73.

Objectif 2. Proposer un pipeline pour prédire de manière non invasive à partir des images de TDM préopératoires, si les MHCR présentent une faible infiltration de cellules T CD3+, associée à une expression élevée de CMH-I.

Objectif 3. Développer et évaluer un outil pour inférer, à partir des images de TDM de pré-traitement, le niveau d'expression de PD-L1 dans le CPNPC avancé et limité traité avec des ICI seuls ou en combinaison avec la chimiothérapie.

Objectif 4. Introduire une nouvelle approche multi-omique pour prédire le risque de récurrence du MHCR à deux ans après une résection chirurgicale en intégrant des images de TDM préopératoires, des données de séquençage d'ARN et des variables cliniques.

4.4 Méthodologie générale

La structure méthodologique de cette thèse est illustrée dans la Figure 4.2. Le premier objectif est de proposer des substituts non invasifs, pertinents sur le plan pronostique, des biomarqueurs immunitaires dans les MHCR. Le principal marqueur immunitaire des MHCR étudié dans cette thèse est l'expression du biomarqueur CD73. La CD73 est une ectonucléotidase qui catalyse la dégradation de l'AMP extracellulaire en adénosine immunosuppressive, qui

peut être bloquée par des anticorps thérapeutiques. Une forte expression de la CD73 dans les MHCR résequés avec l'intention curative est associée à une récurrence précoce et à une survie plus courte des patients. La méthodologie, présentée dans le Chapitre 5 et publiée dans le *Journal of Translational Medicine* (Saber et al. [172]), propose un pipeline pour prédire un substitut d'imagerie de l'expression de la CD73 dans les MHCR résequés chirurgicalement, appelé rad-CD73, basé sur des approches d'apprentissage tabulaire attentif et des images de TDM préopératoires. L'association de rad-CD73 avec les issues cliniques après résection chirurgicale des MHCR, à savoir la TTR et la SSC, a également été évaluée. Le deuxième objectif présente un travail connexe publié dans *International Symposium of Biomedical Imaging 2023* (Saber et al. [173]), dans lequel nous avons prédit si les MHCR présentent une faible densité d'infiltration des cellules T CD3+, avec une forte expression de CMH-I, puisque ce phénotype a été prouvé comme étant associé à de mauvais résultats cliniques [174]. À cet effet, nous avons proposé un modèle d'ensemble basé sur l'apprentissage profond tabulaire pour prédire ce phénotype immunitaire basé sur les caractéristiques radiomiques extraites des images de TDM préopératoires. Le travail susmentionné faisait suite à une étude précédente dans laquelle nous avons prédit l'expression de CD3 dans les MHCR en utilisant des images de TDM de prétraitement (SPIE Medical Imaging 2022 Saber et al. [175]).

Le troisième objectif est de développer un pipeline entraîné sur un ensemble de données multi-institutionnelles pour prédire l'expression de PD-L1 chez les patients atteints de CPNPC traités par ICI. Notre modèle utilise des techniques avancées d'apprentissage profond, y compris l'auto-apprentissage et les Transformeurs, pour améliorer sa précision prédictive et sa généralisabilité. L'approche est présentée dans le Chapitre 7. Nous avons également validé le modèle dans une cohorte indépendante de patients atteints de CPNPC à un stade limité, traités par ICI néoadjuvants et chimiothérapie, et fourni une analyse de l'interprétabilité des scores probabilistes rad-PDL1. Les travaux préliminaires sont publiés dans le congrès *International Symposium on Biomedical Imaging 2024* (Saber et al. [176]) et les résultats finaux ont été soumis au journal *OncoImmunology*.

Le quatrième et dernier objectif de cette thèse, présenté dans le Chapitre 8 et soumis au *Journal of Biomedical and Health Informatics*, est d'exploiter les données multiomiques, à savoir le séquençage de l'ARN, les images de TDM et les données cliniques telles que le SRC de Fong, pour prédire la récurrence des MHCR dans les deux ans suivant leur résection chirurgicale. Pour ce faire, le composant transcriptomique intègre l'apprentissage profond génératif pour imputer les transcrits d'ARN géniques manquants et mitiger le problème de la haute dimensionnalité des ensembles de données d'expression génique. Le problème des transcrits d'ARN non détectés impacte la performance des modèles d'apprentissage automatique sur les tâches en aval, telles que la classification ou la pronostication. La méthodologie, publiée dans

le congrès International Conference on Biomedical and Health Informatics (BHI 2023) [177], utilise l'apprentissage génératif pour fournir des imputations plausibles des transcrits d'ARN manquants. Le pipeline est également évalué sur la tâche de sous-typage du CPNPC, étant donné que les sous-types histologiques de cancer peuvent influencer le MIT, tels que les infiltrats immunitaires, entre autres, et affecter par la suite le pronostic du patient. Cette approche a été incorporée dans le quatrième objectif.

Nous avons également utilisé un modèle de fondation pour l'extraction de caractéristiques, qui complète les caractéristiques radiomiques interprétables dans l'extraction de caractéristiques d'imagerie pertinentes sur le plan pronostique. Nous avons en outre fourni des analyses d'interprétabilité de chaque composant du pipeline multiomique et appliqué la technique bioinformatique d'expression génique différentielle pour dériver les gènes significativement associés à la récurrence des MHCR.

4.4.1 Bases de données

Les deux premières études ont été menées sur un ensemble de données conservées de manière prospective et comprenaient 160 lésions de 122 patients ayant subi une résection complète de MHCR entre 2011 et 2014 au Centre hospitalier de l'Université de Montréal (CHUM). La base de données comprenait des images TDM préopératoires, des données immunologiques provenant des microréseaux de tissus et des caractéristiques cliniques, y compris des caractéristiques démographiques, de traitement et de survie.

La troisième étude a été effectuée sur un ensemble de données multi-institutionnelles de 482 patients atteints d'un CPNPC avancé traités par ICI entre 2015 et 2022 dans trois centres médicaux différents : le CHUM, le Centre Hospitalier Universitaire de Sherbrooke et l'Institut Universitaire de Cardiologie et de Pneumologie de Québec-Université Laval. La validation indépendante a été effectuée sur une cohorte de 51 patients atteints de CPNPC à un stade limité traités au CHUM par un régime néoadjuvant combinant ICI et chimiothérapie.

La quatrième étude a été menée sur l'ensemble de données du CHUM conservées prospectivement et a inclus 129 patients traités entre 2010 et 2022 pour des MHCR. L'ensemble de données comprenait des images de TDM préopératoires, des données de séquençage de l'ARN et des données cliniques, notamment des données démographiques, le SRC de Fong et des données de survie et de progression de la maladie.

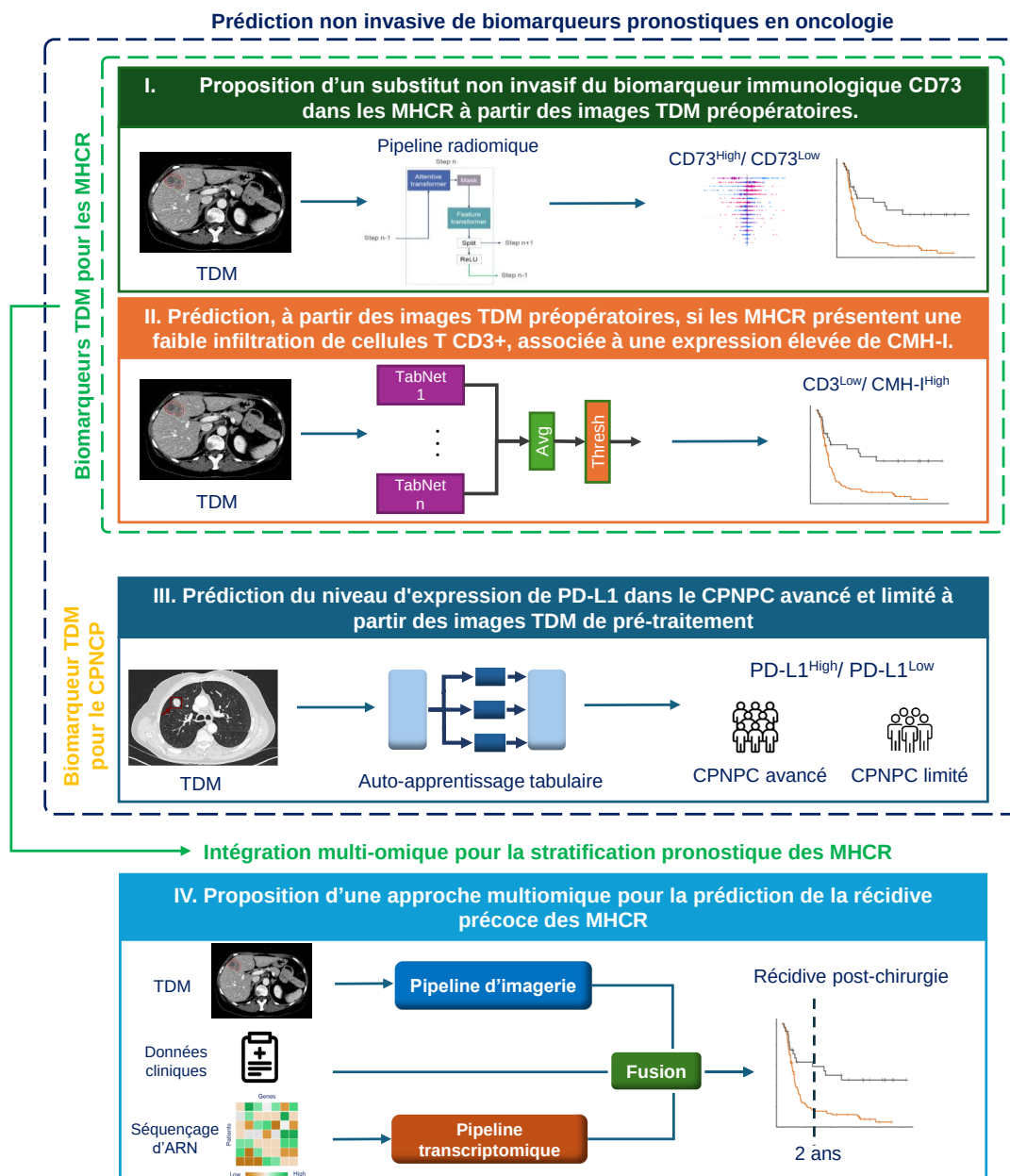


FIGURE 4.1 Présentation du cadre méthodologique proposé. (I) Prédiction de cibles thérapeutiques des MHCR, notamment CD73, à partir d'images de TDM de préopératoires; (II) Proposition d'un pipeline pour prédire à partir des images de TDM si les MHCR présentent une faible infiltration de cellules T CD3+, associée à une expression élevée de CMH-I. (III) Prédiction du niveau d'expression de PD-L1 dans le CPNPC à partir d'images de TDM; (IV) Proposition d'une approche multiomique pour la prédiction de la récurrence à deux ans suivant la résection des MHCR.

Objectif global: Développer des pipelines computationnels et des approches multi-omiques pour améliorer le pronostic du cancer, la personnalisation des traitements et la découverte de biomarqueurs non invasifs.

O1. Concevoir un pipeline pour obtenir des substituts non invasifs des biomarqueurs immunitaires dans les MHCR à partir des images TDM.

- **Article 1:** Radiomics using computed tomography to predict CD73 expression and prognosis of colorectal cancer liver metastases

O2. Proposer un pipeline pour prédire à partir des images TDM si les MHCR présentent une faible infiltration de cellules T CD3+ avec une expression élevée de CMH-I.

- **Article 2:** Ensemble TabNet predicting a T-cell/CMH-I-based immune profile biomarker for colorectal liver metastases from CT images

O3. Développer un outil pour inférer, à partir des images TDM de pré-traitement, le niveau d'expression de PD-L1 dans le CPNPC avancé et limité.

- **Article 3:** Multi-institutional CT-based radiomics for predicting tumor PD-L1 expression in patients with advanced and limited non-small cell lung cancer

O4. Introduire une approche multi-omique pour prédire le risque de récurrence du MHCR à deux ans après une résection chirurgicale en intégrant des images TDM préopératoires, des données de séquençage d'ARN et des variables cliniques.

- **Article 4:** Multi-omics fusion network for prediction of early recurrence in colorectal liver metastases

Discussion et conclusion : Discussion générale sur la méthodologie proposée dans la thèse et ses avantages pour guider le traitement des patients atteints de MHCR et de CPNPC.

FIGURE 4.2 Organisation méthodologique de la thèse

CHAPITRE 5 ARTICLE 1 : RADIOMICS USING COMPUTED TOMOGRAPHY TO PREDICT CD73 EXPRESSION AND PROGNOSIS OF COLORECTAL CANCER LIVER METASTASES

Cet article a été publié dans le Journal of Translational Medicine en juillet 2023.

Remarque : L'étudiant a produit la contribution majeure à l'article : Conceptualisation, Prétraitement des données, Logiciel, Analyse formelle, Validation, Enquête, Visualisation, Revue de la littérature, Méthodologie, Rédaction (version originale), Rédaction (relecture). La contribution de l'étudiant constitue 90% du travail.

Radiomics using computed tomography to predict CD73 expression and prognosis of colorectal cancer liver metastases

Ralph Saber^{1,2}, David Henault^{3,4}, Nouredin Messaoudi^{3,4,5}, Rolando Rebolledo^{3,4}, Emmanuel Montagnon², Geneviève Soucy⁶, John Stagg³, An Tang², Simon Turcotte^{3,4}, Samuel Kadoury^{1,2}

¹MedICAL Laboratory, Polytechnique Montreal, Montreal, Canada

²Imaging and Engineering Axis, Centre de recherche du Centre Hospitalier de l'Université de Montréal/Institut du cancer de Montréal, 900 rue Saint-Denis R10.430, Montréal, Québec, Canada, H2X 0A9

³Cancer Axis, Centre de recherche du Centre Hospitalier de l'Université de Montréal/Institut du cancer de Montréal, 900 rue Saint-Denis R10.430, Montréal, Québec, Canada, H2X 0A9

⁴Hepato-Pancreato-Biliary Surgery and Liver Transplantation Service, Centre hospitalier de l'Université de Montréal, 1000, rue Saint-Denis, Montréal, Québec, Canada, H2X 0C1

⁵Department of Surgery, Vrije Universiteit Brussel (VUB), Universitair Ziekenhuis Brussel (UZ Brussel) and Europe Hospitals, Brussels, Belgium

⁶Pahology department, Centre hospitalier de l'Université de Montréal, 1000, rue Saint-Denis, Montréal, Québec, Canada, H2X 0C1

5.1 Abstract

5.1.1 Background

Finding a noninvasive radiomic surrogate of tumor immune features could help identify patients more likely to respond to novel immune checkpoint inhibitors. Particularly, CD73 is an ectonucleotidase that catalyzes the breakdown of extracellular AMP into immunosuppressive adenosine, which can be blocked by therapeutic antibodies. High CD73 expression in colorectal cancer liver metastasis (CRLM) resected with curative intent is associated with early recurrence and shorter patient survival. The aim of this study was hence to evaluate whether machine learning analysis of preoperative liver CT-scan could estimate high vs low CD73 expression in CRLM and whether such radiomic score would have a prognostic significance.

5.1.2 Methods

We trained an Attentive Interpretable Tabular Learning (TabNet) model to predict, from preoperative CT images, stratified expression levels of CD73 (CD73^{High} vs. CD73^{Low}) assessed by immunofluorescence (IF) on tissue microarrays. Radiomic features were extracted from 160 segmented CRLM of 122 patients with matched IF data, preprocessed and used to train the predictive model. We applied a five-fold cross-validation and validated the performance on a hold-out test set.

5.1.3 Results

TabNet provided areas under the receiver operating characteristic curve of 0.95 (95% CI 0.87 to 1.0) and 0.79 (0.65 to 0.92) on the training and hold-out test sets respectively, and outperformed other machine learning models. The TabNet-derived score, termed rad-CD73, was positively correlated with CD73 histological expression in matched CRLM (Spearman's $\rho = 0.6004$; $P < 0.0001$). The median time to recurrence (TTR) and disease-specific survival (DSS) after CRLM resection in rad-CD73^{High} vs rad-CD73^{Low} patients was 13.0 vs 23.6 months ($P = 0.0098$) and 53.4 vs 126.0 months ($P = 0.0222$), respectively. The prognostic value of rad-CD73 was independent of the standard clinical risk score, for both TTR (HR = 2.11, 95% CI 1.30 to 3.45, $P < 0.005$) and DSS (HR = 1.88, 95% CI 1.11 to 3.18, $P = 0.020$).

5.1.4 Conclusions

Our findings reveal promising results for non-invasive CT-scan-based prediction of CD73 expression in CRLM and warrant further validation as to whether rad-CD73 could assist

oncologists as a biomarker of prognosis and response to immunotherapies targeting the adenosine pathway.

5.1.5 Keywords

cancer, immune checkpoint, adenosine pathway, CD73, radiomic biomarker, interpretable machine learning

5.2 Background

Liver metastases is the most common site of colorectal cancer progression, a malignancy that remains within the top leading causes of cancer-related deaths [178]. Complete resection of colorectal liver metastases (CRLM) combined with systemic chemotherapy, has a curative potential, but approximately 80% of patients recur [179]. Currently, no biomarkers can help identify patients at high risk of early recurrence after CRLM resection, for whom surgery may be futile, who may benefit from adjuvant chemotherapy [180], and who should be followed up more closely based on the risk of recurrence. Consequently, a significant number of patients are burdened by potential complications, side-effects and toxicities associated with treatments not informed by expected prognosis, without overall survival benefits.

Accumulating data support that the immune features of CRLM may be more significantly associated with recurrence and survival after resection, independently from clinical, pathological and tumor genomic features [29]. Extracellular adenosine in the tumor microenvironment (TME) appears as an immunosuppressive mechanism of particular relevance to CRLM microenvironment. The Ecto-5'-nucleotidase, or CD73, expressed by cancer and stromal cells within CRLM, is a rate-limiting enzyme that enhances the breakdown of ATP-derived extracellular adenosine monophosphate into immunosuppressive adenosine [181]. The latter binds to the A2A and A2B receptors expressed by T cells, Natural Killer cells, and other immune cells and inhibits their anti-tumoral cytotoxic functions [182]. High intratumoral CD73 expression is strongly associated with poor prognosis in both primary CRC [47] and CRLM [44].

The current histologic assessment of CD73 expression in CRLM, like other histologically-derived biomarkers, is however obtained only after resection of CRLM. Histological-based markers are also not ideal to assess pre-operatively via biopsy given their intratumoral heterogeneous expression [183]. The field still lacks noninvasive CRLM biomarkers of immune features to guide prognostication, which may also, at term, help identify patients most likely to benefit from immunotherapy. Features of computed tomography (CT) images are widely available and part of routine clinical practice, both at baseline and on follow-up examinations

for assessment of treatment response, and radiomics biomarkers obtained from these images could help clinical decisions early in the treatment course. Current developments in artificial intelligence span almost all conventional medical image analysis tasks, including the detection and segmentation of anatomical structures, classification and registration of medical images [184]. Radiomics, consisting in extracting quantifiable features from medical images, could detect not only macroscopic characteristics, but also hidden genomic and proteomic properties involved in biological processes [58, 59]. Radiomics studies have shown promising results in the diagnosis of liver diseases, including benign diseases and primary and secondary malignancies, cancer staging and grading and the prediction of patient clinical outcomes such as response to therapy [185, 186]. Concerning the liver, contrast-enhanced CT-scan features have been shown able to detect non-alcoholic steatohepatitis [187], to identify the histological growth pattern of CRLM [188], and to predict response to FOLFOX-based chemotherapy in untreated CRLM patients [78]. To our knowledge, however, CT-scan image features have not been associated with immune features in CRLM [189].

The goal of this study was to test whether we could train a deep learning model with radiomic features extracted from preoperative CT images, to predict postoperative CD73 expression in resected CRLM, and to test whether a probabilistic score predicted by the deep learning model, termed rad-CD73, was associated with patient oncological prognosis.

5.3 Methods

5.3.1 Study population

This study received the approval of the Institutional Review Board (No. 19.185 for patient consent for biobanking and database; No. 18.023 for project specific with radiomics). We retrospectively studied a cohort of 215 patients who underwent complete resection of CRLM at the Centre Hospitalier de l'Université de Montréal between 2011 and 2014 and were prospectively followed beyond recurrence and until death for clinicopathological and imaging data in a registry. Clinical annotations included demographics, dates and types of all treatments received, and the Clinical Risk Score, calculated by the addition of one point for each of the following features : disease-free interval between the diagnosis of primary tumor and liver metastases of <12 months ; number of metastases >1 ; pre-operative carcinoembryonic antigen (CEA) level >200 ng/mL ; largest metastasis >5 cm ; and lymph node positive primary tumor [20]. A liver pathologist reviewed hematoxylin and eosin whole slides of all cases to assess resection margins, the presence of tertiary lymphoid structures, the degree of necrosis

and pathological response to preoperative chemotherapy [18].

The inclusion criteria were as follows : (1) CRLM confirmed by histopathological analysis ; (2) complete resection of CRLM with curative intent ; (3) preoperative, intravenous contrast-enhanced abdominal CT scan available ; and (4) CD73 immunofluorescent quantification staining performed. Patients were excluded from the analysis if : (1) preoperative CT images were insufficient to perform feature analysis ; (2) CRLM was not visible on the preoperative CT scan or was calcified ; (3) histopathological results could not be associated with CT image of a given CRLM on the basis of its description in pathology report and review of the CT-scan. After applying the exclusion criteria, 160 CRLM resected in 122 patients were available for further analysis.

5.3.2 Evaluation of CD73 expression

CD73 expression in CRLM quantified by immunofluorescence were previously generated in this cohort [44]. Briefly, we built tissue microarrays (TMA) using six 0.6 μm TMA cores per CRLM, with up to three CRLM per patient, using FFPE blocks after hematoxylin and eosin review of viable tumor areas by a pathologist and trained resident. We optimized a multiplex immunofluorescence panel to concurrently detect CD73, cytokeratins to compartmentalize stromal and cancer cell expression patterns, and DAPI for nuclear staining of viable cells. Standard deparaffinization and rehydration protocols were used, followed by antigen retrieval (Dako S1699) in sub-boiling conditions for 40min, and protein-block (Dako X0909), specific staining with primary antibodies against CD73 (Abcam ab91086, 1 :300 dilution) and cytokeratins 8/18 (Dako IR094, 1 :2 dilution). We used an anti-mouse IgG1 Alexa-Fluor 647 (Life technology, A21240 ; 1/800) and anti-rabbit Alexa-Fluor 488 (Life technology, A21206 ; 1/400) as secondary antibodies, DAPI, and mounted the slides with ProLong Gold (Thermo-Fisher). Slides were digitalized at 20x with NanoZoomer-XR (Hamamatsu) and core images imported with TMA maps and identifiers into Visiomorph v.6 software (Visiopharm) for automated quantification. For each core, the percent surface area containing CD73⁺ cells (expression above background) over the surface area containing all viable cells was calculated, as well as the mean fluorescence intensity in each core. For each CRLM, a mean value was calculated from its corresponding six cores. Patients were stratified into two classes (CD73^{High} and CD73^{Low}) using the median CD73⁺ positive area score across all CRLM evaluated in this cohort as a cutoff value (3.8%). In patients with more than one CRLM, the mean CD73 was used to classify patients as low or high.

5.3.3 Image preprocessing and radiomics workflow

We analyzed the last contrast-enhanced CT images obtained prior to surgery for CRLM resection and acquired in portal venous phase. The images had a cross-sectional volume size of 512×512 , a mean in-plane resolution of 0.72 mm^2 (range = $[0.56, 0.98] \text{ mm}^2$) and a mean in-depth resolution of 2.26 mm (range = $[0.80, 5.0] \text{ mm}$). Images were resampled to isotropic resolution ($1 \times 1 \times 1 \text{ mm}^3$) so as to obtain a uniform pixel spacing within the dataset. An automated segmentation algorithm [71] was used to segment CRLM lesions. The segmentation model consists of two convolutional neural networks trained end-to-end in order to jointly segment the liver and the lesions within. A manual examination and refinement of the ensuing 3D segmentations was subsequently carried out by liver radiologists to verify the quality of the volumes of interest to be included in the analysis.

In a subsequent step, radiomic features were extracted from the volumes of interest using PyRadiomics v3.0.1 toolbox [190]. From each image/tumor mask pair, we extracted 107 radiomics features consisting of 18 first-order statistics, 14 shape features and 75 textural features. First-order statistics are histogram-derived features characterizing the distribution of voxel values within the tumor. Shape features encode the 3D shape and size of the region of interest. These features are calculated from approximated shapes, inferred using triangle meshes generated from binary masks using the marching cube technique. Details on the marching cubes algorithm used to build meshes are presented in the Pyradiomics documentation [190]. Thirdly, textural features are derived from predefined matrices and aim to construe the spatial arrangement of voxel intensity values within the lesion. The textural matrices included are the gray-level co-occurrence matrix (GLCM), the gray level dependence matrix (GLDM), the gray-level run length matrix (GLRLM), the gray-level size zone matrix (GLSZM) and the neighboring gray tone difference matrix (NGTDM).

The resulting feature set was standardized for the sake of obtaining a null mean and a unit standard deviation across instances. The least absolute shrinkage and selection operator (LASSO) [191] was used to select the most salient features. The rationale behind this step is to apply an initial coarse dimensionality reduction in order to discard irrelevant features that would otherwise introduce noise into the training process and mislead the model. The LASSO λ hyperparameter ($\lambda = 0.249$) was determined by applying five-fold cross validation using the mean squared error as an objective function. Each lesion segmented on CT was matched to the corresponding CRLM in the constructed TMA, using their pathology report block number.

5.3.4 TabNet training and evaluation

In this study, we trained an Attentive Interpretable Tabular Learning (TabNet) model [149] to predict the stratified CD73 expression levels ($CD73^{High}$ vs. $CD73^{Low}$). TabNet is a deep learning model that incorporates multiple stages of attention modules within its architecture (Additional File 1, Fig. S1). The attention mechanism is based on two transformer blocks : an attentive transformer and a feature transformer. The attentive transformer allows to compute learnable masks, which are used to select the most salient set of features at each step. It also incorporates a prior scale which encodes the degree to which features have been used in the previous steps. The feature transformer processes the filtered features through shared and decision step-specific layers. The outputs of the different decision steps are linearly combined to form the model’s encoder output. Finally, a fully connected layer processes the encoder’s output to obtain the overall output of the model. In order to compare TabNet’s performance with other machine learning models, we also trained an XGBoost, a random forest (RF), a support vector machine (SVM) with a linear kernel function, and a logistic regression (LR) model to perform the same task.

For TabNet and the other baseline models, we first divided the dataset into a training set (125 lesions) and a hold-out test set (35 lesions) for independent validation using data from our center. We then performed a five-fold cross-validation on the first subset. The dataset was hence divided into five separate folds on five iterations and in each iteration, four folds were used for training and the fifth was used for validation (never seen at training). Once cross-validation was completed, an external validation was performed on the hold-out test set. All splits were applied randomly and on a patient-level. In other words, lesions belonging to the same patient were assigned to the same subset in order to ensure that no overlap or information leakage occur between the training and the testing sets. TabNet was trained for 100 epochs on an NVIDIA GeForce GTX TITAN Xp 12GB with a batch size of 64 and a binary cross-entropy loss function. Adam optimizer [192] was used with a learning rate of 0.02 decaying by 10% after 50 epochs. The area under the receiver operating characteristic curve (AUC) was used to compare the performance of the models.

5.3.5 Model interpretability analysis

Two types of interpretability analyses were carried out in this study : global and local interpretability. In the former case, we attempted to holistically describe the model’s behavior by running a subsequent analysis on the model’s predictions. To this end, we adopted the Shapley Additive Explanations (SHAP) technique [193]. SHAP is an additive feature attri-

bution algorithm that intends to compute, for every feature, a Shapley value which mirrors the contribution of that feature to the model’s final predictions. In this work, we utilized the Kernel SHAP method, which approximates the Shapley values as being the coefficients of a weighted linear regression model, built from a set of sample coalitions. After computing an average Shapley value for each feature over all instances, the features were ranked according to their average Shapley values. Additionally, we made use of the inherent interpretability of the TabNet architecture to provide instance-level explanations of the model’s predictions. To do so, visualizations of the selected feature masks were generated and analyzed, in order to identify the most salient ones for each instance used by the model. The visualization of TabNet’s selected feature importance served two main purposes : (1) to acquire a local explanation of the model’s predictions for each instance and (2) to examine whether the predictive behavior projected by the model’s architecture, was concordant with SHAP results.

5.3.6 Statistical analysis

We used the Wilcoxon rank sum test and Fisher’s exact test for numerical and categorical variables, respectively. Spearman’s correlation coefficient was applied to evaluate correlations. The decision curve analysis [194] was performed using the Python dcurves v0.0.3 package by plotting the net benefit of the predicted biomarker for different threshold probability values (range = [0-1]), given as the minimum probability for which additional testing is recommended. The line representing the “all lesions are CD73^{Low}” hypothesis and the curve representing the “all lesions are CD73^{High}” hypotheses were also plotted for comparison purposes. Survival curves were generated with the Kaplan-Meier method and compared with the log-rank test. Disease-specific survival, for which deaths due to causes other than cancer progression were censored observations, and time to recurrence, were computed from the date of the first hepatectomy. Patients with missing data on the first surgery were excluded from the survival analysis. For survival analysis purposes, TabNet class probabilities, termed rad-CD73, were derived and stratified into rad-CD73^{High} and rad-CD73^{Low} by setting a cut-off value equal to the lower tertile (rad-CD73 > 0.362) based on the distribution of the rad-CD73 score (Additional File 1, Fig. S2), which was consistent with the optimal p-value cut-off proposed by the X-tile software (0.383). For patients with multiple lesions, predicted rad-CD73 scores were averaged across all lesions. Univariate and multivariate Cox proportional hazards regression models were used to generate hazards ratios (HR) with 95% confidence intervals. A two-sided p-value < 0.05 was considered statistically significant. Statistical analyses were conducted using Python Scipy v1.5.3, Python Lifelines v0.27.1 and R Survival v3.4 packages.

5.4 Results

The general workflow of the planned analysis is depicted in Fig. 5.1.

5.4.1 Patient cohort

The clinicopathological characteristics of the 122 patients treated for resectable CRLM are summarized in Table 5.1. Mean patient age was 63.4 years (35 to 84), and male patients predominated (63.9%). Preoperative chemotherapy was administered in 78.7% of the patients, consisting of four to six cycles of folfox-based regimen for the vast majority of patients (not shown). Most patients (81.1%) also received chemotherapy after resection of CRLM. Approximately half of the patients were treated for multiple metastasis with a mean number of two (range 1 to 10) and a mean diameter of 4.1 cm (range 1.0 to 20 cm). Based on the composite Clinical Risk Score [20], 38.5% of patients were at high risk of recurrence and death from cancer progression. At time of analysis, the median follow-up was of 57.0 months, during which time 76.2% of patients had recurred and 64.8% died of disease progression.

5.4.2 Prediction of CD73 expression from preoperative CT images

In the training cohort, the ability of the TabNet model to classify $CD73^{High}$ vs. $CD73^{Low}$ lesions was shown to have an AUC of 0.95 (95% confidence interval : 0.87- 1.0). The accuracy, sensitivity and specificity were 0.85, 0.91 and 0.79, respectively. Table 5.2 summarizes the performance of different models on the hold-out test set. Moreover, TabNet exhibited a high predictive performance on the hold-out test set with an AUC of 0.79 (0.65-0.92). The test set accuracy, sensitivity and specificity were 0.71, 0.63 and 0.79, respectively. Fig. 5.2 depicts the receiver operating characteristic curve and the confusion matrix of the model on the hold-out test set. The model exhibited balanced sensitivity and specificity values, and no class bias was observed. This is reflected by the distribution of the true positives and true negatives in the matrix. We also compared TabNet generalization capability with other machine learning models, which outperformed XGBoost, RF, SVM and LR models. Even though the XGBoost model achieved a satisfactory AUC of 0.61 (0.45-0.77), TabNet outclassed it with a high margin.

A significant difference was observed in the predicted TabNet rad-CD73 scores between $CD73^{High}$ and $CD73^{Low}$ lesions (Wilcoxon signed rank test, $P < 0.0001$) (Fig. 5.3A). Rad-CD73 scores were positively correlated with CD73 histological expression measured by IF (Spearman's $\rho = 0.600$, $P < 0.0001$) (Fig. 5.3B). With the goal of assessing the clinical uti-

TABLEAU 5.1 Patient clinicopathological characteristics and correlation with rad-CD73

Variable	All dataset (N=122)	rad-CD73 ^{High} (n=79)	rad-CD73 ^{Low} (n=43)	p-value
Age at hepatectomy (Mean, range)	(63.4, 35 - 84)	(63.6, 35 - 84)	(63.2, 44 - 81)	0.894
≤65 years	66 (54.1)	42 (53.2)	24 (55.8)	
>65 years	56 (45.9)	37 (46.8)	19 (44.2)	
Gender				0.43
Male	78 (63.9)	48 (60.8)	30 (69.8)	
Female	44 (36.1)	31 (39.2)	13 (30.2)	
Number of metastases (Mean, range)	(2.0, 1 - 10)	(2.1, 1 - 10)	(1.9, 1 - 5)	0.643
Single	59 (48.4)	41 (51.9)	18 (41.9)	
Multiple	63 (51.6)	38 (48.1)	25 (58.1)	
Diameter of largest metastasis (Mean, range)	(4.1, 1.0 - 20.0)	(4.3, 1.0 - 20.0)	(3.8, 1.3 - 11.0)	0.83
≤5 cm	93 (76.2)	59 (74.7)	34 (79.1)	
>5 cm	29 (23.8)	20 (25.3)	9 (20.9)	
KRAS status				0.735
Wild-type	27 (61.4)	18 (58.1)	9 (69.2)	
Mutated	17 (38.6)	13 (41.9)	4 (30.8)	
CEA level				0.287
≤200 ng/mL	118 (97.5)	77 (98.7)	41 (95.3)	
>200 ng/mL	3 (2.5)	1 (1.3)	2 (4.7)	
Clinical Risk Score*				0.017
Low risk (0-2)	72 (61.5)	53 (69.7)	19 (46.3)	
High risk (3-5)	45 (38.5)	23 (30.3)	22 (53.7)	
Margin liver resection				0.325
Negative	111 (91.0)	70 (88.6)	41 (95.3)	
Positive	11 (9.0)	9 (11.4)	2 (4.7)	
Tertiary lymphoid structure				0.775
No	107 (87.7)	70 (88.6)	37 (86.0)	
Yes	15 (12.3)	9 (11.4)	6 (14.0)	
Necrosis (Mean, range)	(12.3, 0.0 - 81.5)	(13.2, 0 - 81.5)	(10.7, 0.0 - 61.8)	0.358
≤ 25 %	104 (85.2)	67 (84.8)	37 (86.0)	
>25 %	18 (14.8)	12 (15.2)	6 (14.0)	
Primary tumor				0.14
Left sided	82 (69.5)	58 (74.4)	24 (60.0)	
Right sided	36 (30.5)	20 (25.6)	16 (40.0)	
pT category				0.19
pT1-pT3	93 (83.0)	56 (78.9)	37 (90.2)	
pT4	19 (17.0)	15 (21.1)	4 (9.8)	
pN category				0.116
pN0	43 (35.8)	32 (41.0)	11 (26.2)	
pN+	77 (64.2)	46 (59.0)	31 (73.8)	
Preop chemotherapy				1.000
No	26 (21.3)	17 (21.5)	9 (20.9)	
Yes	96 (78.7)	62 (78.5)	34 (79.1)	
Tumor regression grade				0.575
1-2	16 (13.1)	9 (11.4)	7 (16.3)	
3-4-5	106 (86.9)	70 (88.6)	36 (83.7)	
Extrahepatic recurrence				0.806
No	32 (36.0)	24 (37.5)	8 (32.0)	
Yes	57 (64.0)	40 (62.5)	17 (68.0)	
Disease-free interval between primary and liver met				0.037
<12 months	88 (72.1)	52 (65.8)	36 (83.7)	
≥12 months	34 (27.9)	27 (34.2)	7 (16.3)	

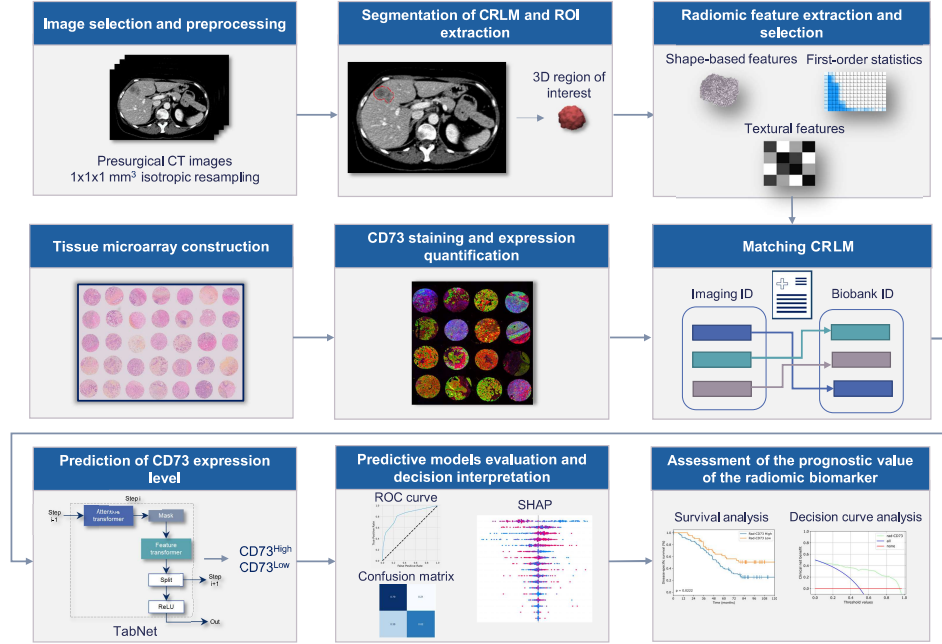


FIGURE 5.1 Isotropic spatial resampling was applied on preoperative CT-scan images to segment 160 colorectal liver metastases (CRLM) resected in 122 colorectal cancer (CRC) patients who underwent partial hepatectomy. Matching CRLM, identified by their pathology report block numbers and anatomical description, were included in tissue microarrays for automated quantification of CD73 intra-tumoral expression by immunofluorescence. Radiomic features were extracted from the resulting three-dimensional regions of interest (ROI) (18 first-order statistics, 14 shape features and 75 textural features) and preprocessed. Subsequently, an Attentive Interpretable Tabular Learning (TabNet) model was trained to predict CD73 expression dichotomized as CD73^{High} vs. CD73^{Low} based on the surface area expressing CD73 over the total surface of assessable tissue. The model was then evaluated, and its predictions were interpreted using ROC curve and the Shapley Additive Explanations technique (SHAP). The association between radiomic CD73 (rad-CD73) and oncological survival outcomes was analyzed by Kaplan-Meier and logrank test.

lity of the predicted rad-CD73 score, we performed a decision curve analysis and compared the net benefit of using rad-CD73 with the “treat all as CD73^{High}” and the “treat none as CD73^{High}” strategies. For probability thresholds higher than 0.08, rad-CD73 had a higher net benefit than both the “treat all” and “treat none” approaches (Fig. 5.4).

5.4.3 Interpretability of the predictive model

Fig. 5.5 shows representative CT-scan images and corresponding histological images and CD73 IF expression of two CRLM cases with high and low CD73 expression, respectively. Case 1 (left) represents a CRLM with a high CD73 expression (high red IF signal, % CD73+

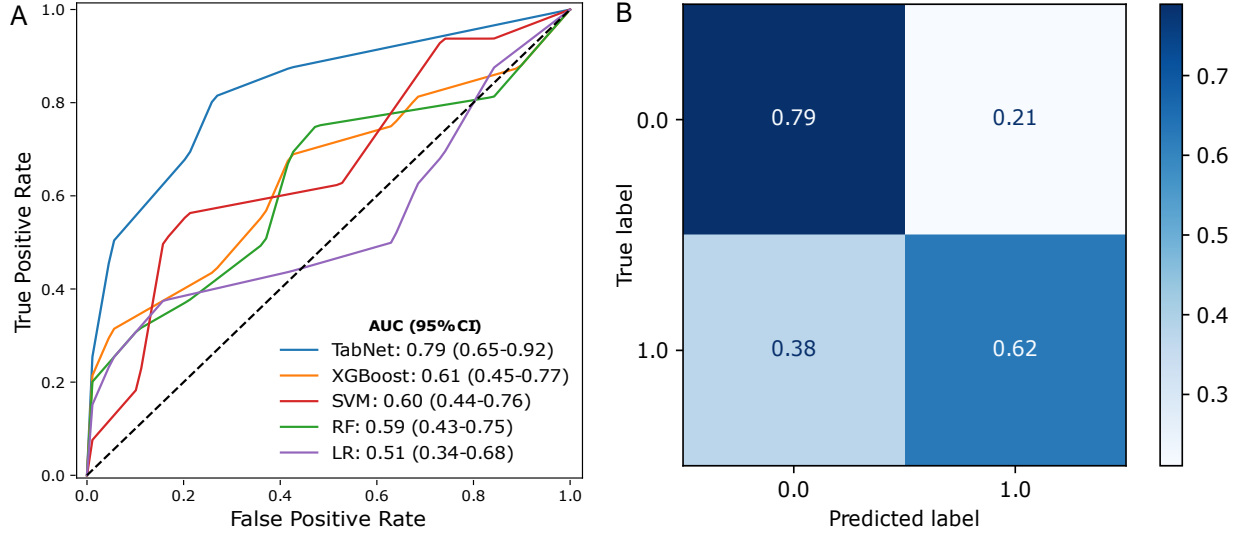


FIGURE 5.2 Performance of TabNet on the hold-out test set. **A** Comparison of the ROC curves of TabNet with the other baseline models. TabNet outperformed all the trained baseline models as mirrored by the area under the ROC curves (AUC). **B** Confusion matrix of TabNet for CD73 classification. LR, Logistic Regression; RF, Random Forest; SVM, Support Vector Machine; TabNet, Attentive Interpretable Tabular Learning; XGBoost, Extreme Gradient Boosting.

TABLEAU 5.2 Performance of the different models on the hold-out test set. AUC, Area Under the Curve; TabNet, Attentive Interpretable Tabular Learning; XGBoost, Extreme Gradient Boosting; SVM, Support Vector Machine; RF, Random Forest; LR, Logistic Regression.

Model	AUC	Accuracy	Sensitivity	Specificity	F1 score
TabNet	0.79 (95% CI 0.65-0.92)	0.71	0.63	0.79	0.67
XGBoost	0.61 (95% CI 0.45-0.77)	0.6	0.69	0.53	0.61
SVM	0.60 (95% CI 0.44-0.76)	0.49	0.94	0.11	0.63
RF	0.59 (95% CI 0.43-0.75)	0.63	0.75	0.53	0.65
LR	0.51 (95% CI 0.34-0.68)	0.46	0.63	0.32	0.51

surface area = 19.24). Concordantly, the corresponding rad-CD73 probabilistic score was 0.69. On the other hand, a low rad-CD73 score (0.06) was attributed to the CRLM of case 2 (right) having a low CD73 IF expression (% CD73+ surface area = 0.37). This finding supported that different CT-scan features could be observed between $CD73^{High}$ and $CD73^{Low}$ CRLM, with homogeneity in CT tumor segmentation in case 1 and heterogeneity in the CT tumor segmentation in case 2. To improve the interpretation of TabNet's predictions, we applied the SHAP technique and studied the features that contributed the most to the model's outputs. Fig. 5.6 shows that distinct features had different impact on the model's output, mirrored by its average Shapley value. Amongst the top five features selected by SHAP, four were textural,

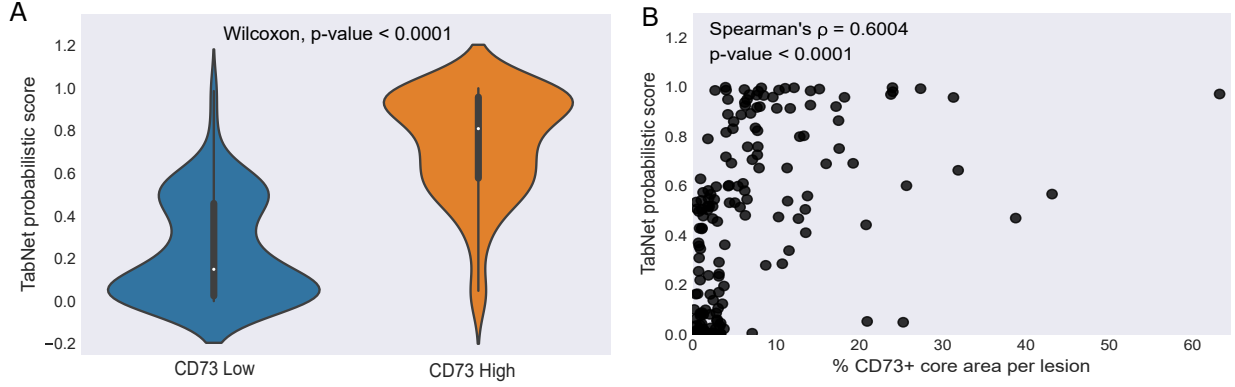


FIGURE 5.3 Assessment of the predicted TabNet probabilistic score, rad-CD73. **A** Violin plot depicting the distribution of rad-CD73 score by CD73 expression level. A statistically significant difference was observed in the radiomic score between the $CD73^{High}$ and $CD73^{Low}$ groups (Wilcoxon signed rank test; $P < 0.0001$). **B** Spearman's correlation between rad-CD73 and the actual CD73 expression (the percentage of CD73 positive area per metastasis).

reflecting the importance of texture-related characteristics of the lesions in predicting CD73 expression level.

The feature with the highest Shapley value was the “Dependence Non Uniformity Normalized” (DNUN), a textural feature computed using the GLDM matrix. The DNUN encodes the heterogeneity in terms of the dependence throughout the lesion : a high DNUN reveals that the image contains regions with disparate dependence levels. The dependence is a term reflecting whether the gray level of a given voxel is dependent on those of the neighboring ones. A region with a low dependence is hence formed of voxels with comparable gray levels whereas the voxels of a high dependence region exhibit discrepancies in their gray levels. SHAP results show that a low DNUN value had a positive impact on the model's output, prompting an increase in the predicted rad-CD73 score.

Similarly, the textural feature “Size Zone Non Uniformity Normalized” (SZNUN), which encodes the heterogeneity in the size zone volumes of a lesion defined as areas with a constant gray level, had an effect comparable the DNUN on TabNet predictions. The “Small Area Emphasis” (SAE) and the “Informational Measure of Correlation 1” (IMC1) came in third and fourth, respectively. The SAE reflects the prevalence of small zones while the IMC1 encodes the complexity of the textures. Both features had a positive impact on the model's output ; a higher feature value was corresponded to a higher Shapley value. These findings show that the model had the tendency to output high rad-CD73 scores for lesions exhibiting finer textures.

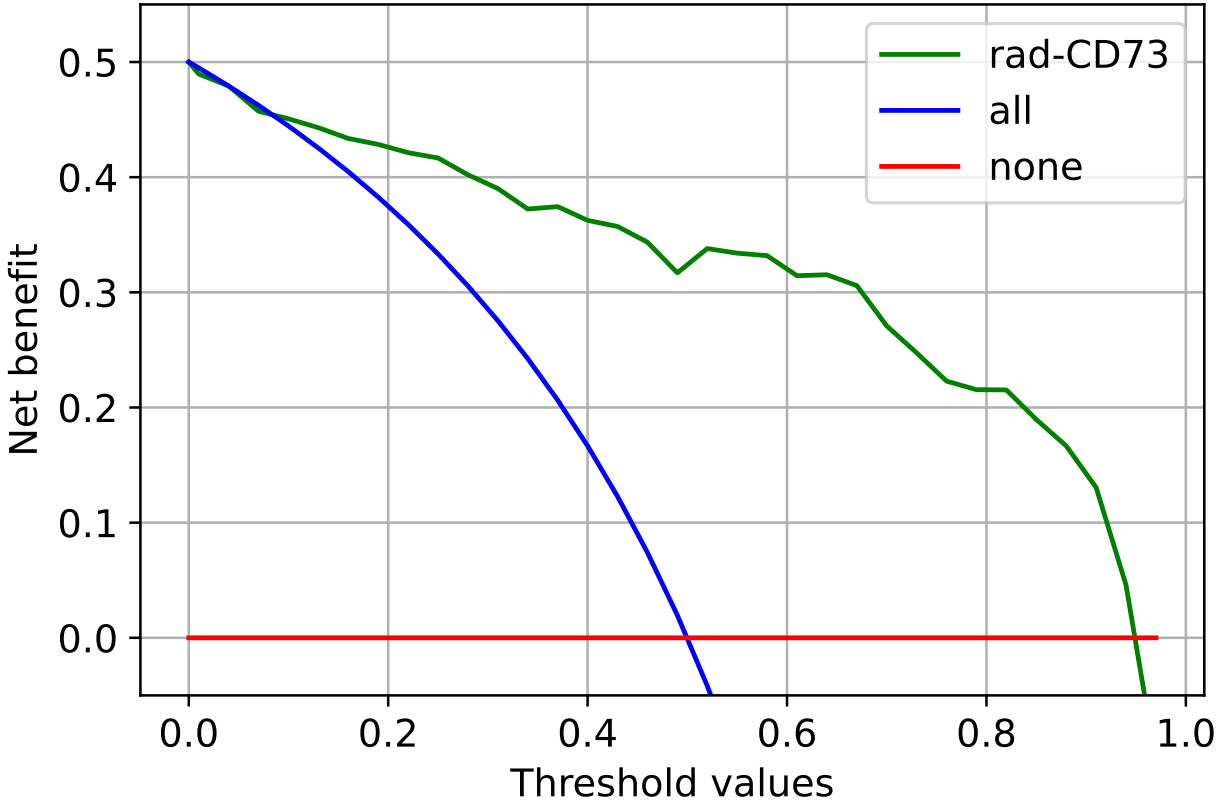


FIGURE 5.4 Decision curve analysis. The red line shows the treat none as $CD73^{High}$ approach while the blue line shows the treat all as $CD73^{High}$ approach. For threshold values greater than 0.08, rad-CD73 had a higher net benefit than both the “treat all” and “treat none” approaches.

Interestingly, the shape of the lesions also had an impact, albeit less prominent, on the model’s behavior. In fact, SHAP results reveal that spherical lesions were associated with high CD73 expression levels. Finally, Fig. 5.7 shows that the DNUN was selected for almost all test set instances among the most salient features in TabNet’s third feature selection stage. This finding was consistent with SHAP, as the DNUN had the highest average Shapley value.

5.4.4 Association of rad-CD73 with clinicopathological characteristics and oncological outcomes

We then analyzed the potential clinical significance of rad-CD73 high vs low status. As shown in Table 5.1, the proportions of rad-CD73^{High} and rad-CD73^{Low} patients were generally similar according to most clinicopathological characteristics, including CRLM size and number, having received chemotherapy or not prior to liver resection, and pathological response to

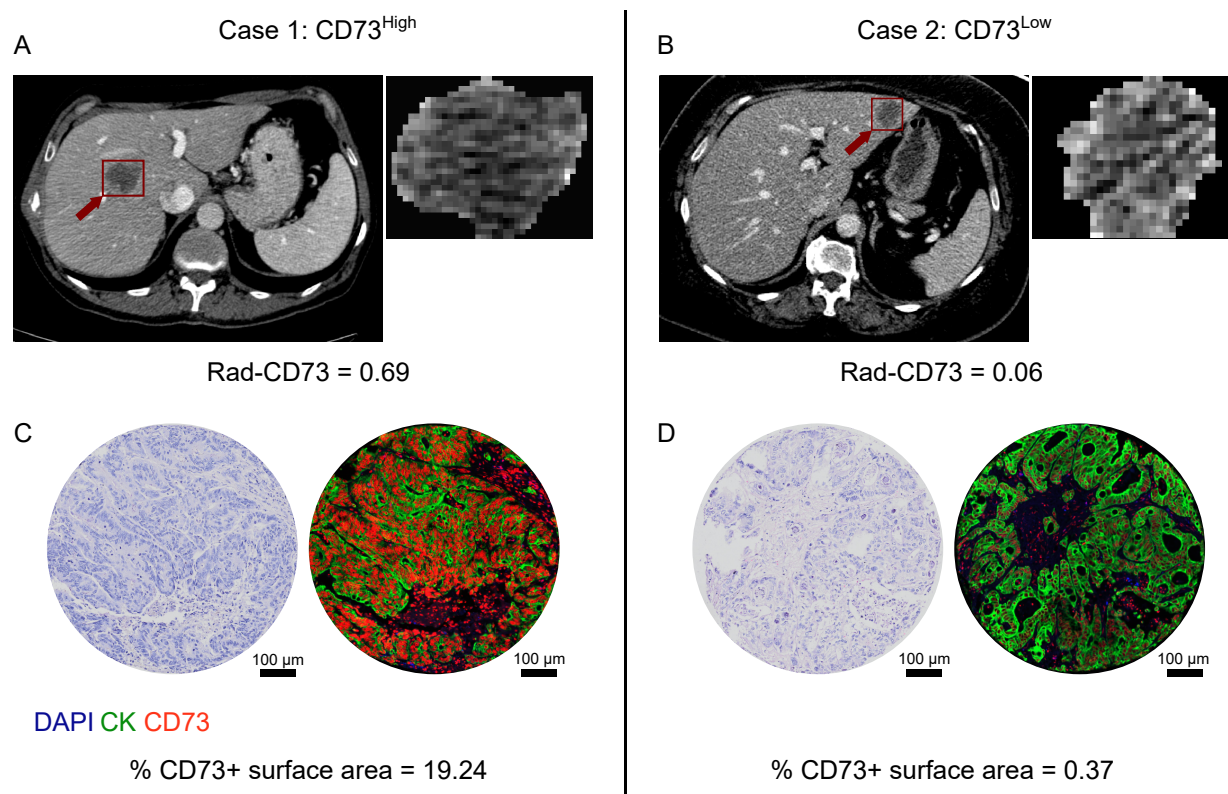


FIGURE 5.5 Representative cases with high and low CD73 expression. **A** A preoperative CT image of a CD73^{High} CRLM (arrow) in the right hemiliver. **B** A preoperative CT image of a CD73^{Low} CRLM (arrow) in the left hemiliver. **C-D** Representative tissue microarray cores of the high **C** and low **D** CD73 expression of the respective CRLM. Hematoxylin and eosin staining (left) shows the integrity and general architecture of the tissue. Immunofluorescent staining (right) shows cell nucleus (blue, DAPI), CD73 expression (red), and the cancer cells with the pan-cytokeratins epithelial cell marker (green). Scale bars represent 100 μ m.

pre-operative chemotherapy. There was hence no statistically significant difference in CRLM diameter ($P = 0.830$), whether the primary tumor is located in the right or left colon ($P = 0.140$), the KRAS mutation status ($P = 0.735$) and the CEA level ($P = 0.287$) between rad-CD73^{High} and rad-CD73^{Low} CRLM. More patients with rad-CD73^{High} CRLM were however found in those who were diagnosed with liver metastases less than 12 months after the diagnosis of primary CRC ($P = 0.037$). Although this criterion is one of those constituting the composite clinical risk score (CRS), there were less rad-CD73^{High} patients in those classified as higher risk of recurrence (CRS score 3, 4 or 5).

By univariate analysis, rad-CD73 high vs low status was significantly associated with TTR and DSS, as well as expected clinicopathological features such as the primary tumor depth of invasion (pT stage), high pre-operative CEA, and the composite CRS (Table 5.3). In the

TABLEAU 5.3 Univariate analyses of outcomes according to clinicopathological variables and rad-CD73. HR, Hazard Ratio; CI, Confidence interval.

	Disease-specific survival		Time-to-recurrence	
	HR (95% CI)	p-value	HR (95% CI)	p-value
Age at hepatectomy (≤ 65 vs > 65 years)	0.88 (0.56 - 1.39)	0.58	0.70 (0.46 - 1.07)	0.10
Gender (male vs female)	1.16 (0.72 - 1.84)	0.54	0.97 (0.63 - 1.51)	0.91
Number of metastases (single vs multiple)	1.37 (0.87 - 2.16)	0.17	1.50 (0.98 - 2.29)	0.06
Diameter of largest metastasis (≤ 5 vs > 5 cm)	1.43 (0.87 - 2.37)	0.16	1.44 (0.90 - 2.30)	0.13
Disease-free interval 12 months (no vs yes)	1.52 (0.90 - 2.58)	0.12	1.43 (0.89 - 2.28)	0.14
CEA level (≤ 200 vs > 200 ng/mL)	4.38 (1.34 - 14.38)	0.01	5.12 (1.57 - 16.62)	0.01
Tertiary lymphoid structure (absent vs present)	0.86 (0.43 - 1.73)	0.68	0.95 (0.49 - 1.84)	0.88
Liver resection margin (negative vs positive for cancer cells)	1.06 (0.50 - 2.22)	0.89	1.10 (0.55 - 2.18)	0.79
Necrosis (≤ 25 vs > 25 %)	1.18 (0.63 - 2.18)	0.61	1.02 (0.57 - 1.80)	0.96
Primary tumor (left vs right)	0.79 (0.48 - 1.28)	0.33	0.87 (0.55 - 1.37)	0.54
pT category (T1-T2-T3 vs T4)	2.05 (1.16 - 3.59)	0.01	2.51 (1.45 - 4.34)	< 0.005
pN category (N0 vs N+)	1.25 (0.77 - 2.02)	0.37	1.30 (0.84 - 2.01)	0.25
Pre-operative chemotherapy (no vs yes)	1.73 (0.91 - 3.27)	0.09	1.73 (0.99 - 3.03)	0.05
Tumor Regression Grade (1-2 vs 3-4-5)	0.84 (0.44 - 1.61)	0.61	1.31 (0.68 - 2.52)	0.43
KRAS status (wild-type vs mutated)	1.18 (0.61 - 2.28)	0.63	2.23 (1.14 - 4.37)	0.02
Clinical Risk Score (0-1-2 vs 3-4-5)	1.60 (1.01 - 2.51)	0.04	1.97 (1.29 - 3.03)	< 0.005
rad-CD73 (low vs high)	1.80 (1.08 - 3.00)	0.02	1.84 (1.15 - 2.95)	0.01

initial immunohistochemical study of CD73 expression in this patient cohort [44], patients with high intratumoral CD73 expression had significantly shorter median TTR compared to low CD73 (11.0 vs 46.4 months) and DSS (19.0 vs. 61.5 months), independent of conventional clinicopathological variables by multivariate analyses. Consistent with the worse prognosis observed in patients bearing tumor with high intratumoral CD73 expression assessed by immunohistochemistry [44, 195–198], patients with rad-CD73^{High} CRLM had a shorter median TTR of 13.0 months compared to 23.6 months in rad-CD73^{Low} CRLM patients ($P = 0.0098$). Consistently, the median DSS of rad-CD73^{High} CRLM patients was 53.4 months compared to 126.0 months in rad-CD73^{Low} CRLM patients ($P = 0.0222$) (Fig. 5.8). Consistent with the lack of positive association observed between the CRS and rad-CD73, multivariate modeling supported that the prognostic value of rad-CD73 was independent of the CRS for both TTR and DSS (Table 5.4).

5.5 Discussion

In this study, we developed a noninvasive imaging surrogate of CD73 by leveraging state-of-the-art deep learning techniques trained with radiomic features. Despite recent progress in prognostication based on immune features of CRLM resected with curative intent [29], there are no noninvasive immune biomarkers that ultimately may guide clinical decision making. To our knowledge, this is the first work developing and testing a machine learning tool to predict immunosuppressive CD73 expression from diagnostic CT images. The proposed model exhibited good performance in classifying CRLM lesions into CD73^{High} and CD73^{Low} groups. We also demonstrated that the predicted rad-CD73 score was highly correlated with the actual expression as measured *in vitro* by immunohistochemistry. Moreover, the clinical significance of rad-CD73 was supported by its association with patient prognosis.

Radiomics has achieved major breakthroughs in recent years in precision oncology and has

TABLEAU 5.4 Multivariate analyses of outcomes according to clinical risk score and rad-CD73. HR, Hazard Ratio; CI, Confidence interval.

	Disease-specific survival		Time-to-recurrence	
	HR (95% CI)	p-value	HR (95% CI)	p-value
Clinical Risk Score (0-1-2 vs 3-4-5)	1.83 (1.15 - 2.92)	0.01	2.40 (1.54 - 3.74)	<0.005
rad-CD73 (low vs high)	1.88 (1.11 - 3.18)	0.02	2.11 (1.30 - 3.45)	<0.005

paved the way for individualized patient care. Its clinical applications include disease diagnosis, prognosis and treatment planning [199,200]. It could hence be applied for cancer detection, allowing for a noninvasive differentiation between benign and malignant neoplasms and consequently minimizing the unwarranted collection of tissue samples. Moreover, radiomics was coupled with conventional diagnostic tools in order to augment their sensitivity to detect diseases early in their development [201]. Finally, it was shown to be able to forecast oncological outcomes of patients such as survival, recurrence, response to adjuvant therapy and metastatic progression [202]. The main advantage of radiomics over conventional techniques is that it provides a holistic and noninvasive assessment of the tissues, as opposed to more invasive histopathological tissue analysis methods and RNA sequencing which require biopsies taken from tumor regions. Its performance is hence less affected by the tumoral heterogeneity associated with the biopsy site. Moreover, it is less labor intensive and allows a quicker profiling of the patients than existing diagnostic tools. Nevertheless, the translation of radiomics to the clinic has been hindered by some challenges including the current retrospective design of the majority of the radiomics studies and the “black box” nature of the predictive models. Therefore, radiomics should be considered as a promising complementary tool for personalized treatment to be validated prospectively.

In the past few years, several efforts have been made for the interpretation of the results produced by machine learning models, which are still considered as black boxes. Model interpretability is particularly important in the medical oncology field to ensure traceability and informed clinical decision-making. Interpretability techniques could be divided into two major categories : model-specific and model-agnostic [203]. The former interpretability is acquired in models that are conceived to be inherently explainable through attention mechanisms for instance, which provides feedback on regions of focus in deep learning [204,205]. On the other hand, model-agnostic interpretability techniques are usually implemented separately and do not depend on the model’s architecture. While the latter involves the implementation of an additional step, it presents the advantage of being applicable to a variety of models. In this work, we tested both techniques. We first trained a TabNet model and sought to understand its instance-level decisions by visualizing its feature selection masks. In a post-hoc analysis, we applied SHAP technique to decipher its overall behavior. The study revealed that the most salient features for the prediction of CD73 expression were texture-related. Moreover, textural heterogeneity was associated with a lower CD73 expression and inversely applicable. This was mirrored by the impact of the DNUN, the SZNUN and the IMC1 on the model’s predictions. These findings are in line with several previous studies associating textural features with response to immunotherapy [206]. Tang et al. [207] found a cluster of non-small cell lung cancer patients exhibiting concurrent tumoral heterogeneity and high CD3 T cell

infiltration. Yoon et al. [107] showed that type 2 helper T cells were associated with high variance and IMC, mirroring lesion heterogeneity. On the other hand, Sun et al [104] found, in a study combining several cancer sites, that a high CD8 score, indicative of inflammatory infiltrate, was associated with homogeneous lesion appearances. They also attributed heterogeneity in pixel intensities to convoluted underlying processes such as excessive disorderly tumor vascularization. CT-scan image features may also be enhanced by the use of nanoparticles, characterized by a high permeability and retention in tumors. Several nanoparticles have been tested [208], including silver nanoparticles [209]. In this context, Devkota et al [210] demonstrated that radiomics features, extracted from nanoparticle contrast-enhanced CT rather than conventional imaging, were better suited for the prediction of response to cellular immunotherapy.

Immunotherapy has become a mainstay in the treatment of several advanced malignancies. This has motivated several researchers to leverage radiomics to forecast response to immunotherapeutic agents [60, 211, 212] either by predicting established biomarkers within the tumor microenvironment or by attempting to directly associate imaging features with patient outcomes, such as radiological response, survival and time-to-recurrence. However, the vast majority of the conducted studies focus on non-small cell lung cancer given the universal availability of chest images and the proven effectiveness of immune checkpoint inhibitors (ICI) in advanced lung cancer. The application of radiomics in CRLM immuno-oncology remains vastly unexplored. While prior work on radiomics of CRLM have focused on response to chemotherapy as measured by the tumor regression grade [213] or the RECIST criteria [214, 215], and tumor histological features such as the histological growth patterns [188], we aimed to develop and validate a radiomic immune marker for CRLM.

While first generation ICI such as anti-programmed cell death protein 1 (PD1), anti-PD1 ligand and anti-cytotoxic T lymphocyte antigen 4 have proven to be effective in some cancers [216], 95% of metastatic CRC are refractory to these immunotherapies [13, 52]. The underlying mechanisms driving these poor outcomes include tumor heterogeneity and the coexistence of complex immune escape mechanisms within the hepatic tumor microenvironment. Because of the immunosuppressive nature of the adenosine pathway, adenosinergic molecules are now being explored for the development of novel therapeutic agents [217]. In particular, CD73 ectonucleotidase plays a major role in the generation of immunosuppressive adenosine and has recently emerged as a novel immunotherapeutic target that can be blocked by monoclonal antibodies, while adenosine receptor inhibitors are also being tested in early phase trials [218, 219].

Overall, the clinical use of high rad-CD73 on CT-scan imaging of patients with resectable

CRLM, as it may identify a subset of patients with earlier recurrence, death, and higher intra-tumoral CD73 expression, could be tested prospectively in many ways to determine whether : a) closer follow-up after CRLM resection could lead to earlier treatment of recurrence and survival benefits ; b) adjuvant systemic therapy after CRLM resection could reduce the risk of recurrence and death ; and c) it can predict the efficacy of patients more likely to respond to anti-CD73 or adenosine receptor inhibitors.

Our work has some limitations notwithstanding. First, because this is a single-center investigation, it will be important to verify the reproducibility of the results by testing the model on a cohort of patients recruited in different institutions. Leveraging epidemiologically diversified databases is of equal importance in order to minimize any bias that could be introduced by unrepresentative datasets. Second, the sample size is relatively limited to deploy deep learning models. Nevertheless, radiomic pipelines have the advantage of being more transparent than end-to-end black box models due to the interpretability of radiomic features. Third, our work does not take into account the dynamic aspect of the tumor microenvironment, with varying delays between the preoperative CT and the histopathological analysis from one patient to another. Future studies should take into account the temporal fluctuations when training artificial intelligence tools to characterize tumor biology. This is markedly true in radiomics-based analyses since medical images could be leveraged in longitudinal studies as a result of their omnipresence.

5.6 Conclusions

In this study, we introduced a deep learning pipeline for the prediction of CD73 expression in curatively resected CRLM from preoperative CT images. The conceived rad-CD73 biomarker could serve as a noninvasive, fast and low cost tool to identify candidates for targeted immunotherapy. Due to its association with patient prognosis, it could also be leveraged to assist oncologists to personalize the need for adjuvant treatments and the intensity of follow-up strategies. The generalizability of the model needs to be validated on independent, large and epidemiologically diverse cohorts, and its impact on clinical decision making will need to be tested prospectively.

5.7 Abbreviations

AUC : Area under the receiver operating characteristic curve ; CEA : Carcinoembryonic antigen ; CRLM : Colorectal cancer liver metastases ; CRC : Colorectal cancer ; CRS : Clinical

risk score; CT : Computed tomography; DNUN : Dependence non uniformity normalized; GLCM : Gray-level co-occurrence matrix; GLDM : Gray level dependence matrix; GLRLM : Gray-level run length matrix; GLSZM : Gray-level size zone matrix; HR : Hazards ratios; ICI : Immune checkpoint inhibitors; IMC : Informational measure of correlation; LASSO : Least absolute shrinkage and selection operator; LR; Logistic regression; NGTDM : Neighboring gray tone difference matrix; PD1 : Programmed cell death protein 1; RF : random forest; SVM : Support vector machine; SAE : Small area emphasis; SHAP : Shapley additive explanations; SZNUN : Size zone non uniformity normalized; TabNet : Attentive Interpretable Tabular Learning; TMA : Tissue microarrays.

5.8 Acknowledgments

The authors would like to thank L. Rousseau, S. Langevin and J. Bilodeau for the recruitment of patients and the maintenance of clinicopathological data; L. Meunier and V. Barès for building the tissue microarray and scanning; A. Cleret-Bohot from the CRCHUM for assisting in the automated quantification of cells and markers; M. Cerny, V. Hamilton, T. Derennes and A. Ilinca for editing segmentations and distinguishing liver metastases from benign coexisting liver lesions.

5.9 Authors' contributions

SK conceived the initial project of predicting immunological profiles from CT images and ST provided domain expertise in CRLM. NM, DH, ST and JS generated CD73 expression data. GS supervised block selection for TMA construction. ST, RS, DH and RR performed CRLM matching. AT and EM supervised CRLM segmentation on CT. RS developed the framework, trained the models, performed statistical analyses and wrote the original paper draft. All authors provided critical revision of the manuscript according to their domain knowledge.

5.10 Funding

This work was supported by the Canada Research Chairs, by the National Science and Engineering Research Council of Canada (NSERC) Discovery grant RGPIN-2020-06558 and the Université de Montréal Roger Des Groseillers Research Chair in Hepatopancreatobiliary Surgical Oncology. ST and SK are scientists of the Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM) supported by the Fonds de recherche du Québec -

Santé (FRQ-S). ST was supported by the FRQ-S Young Clinician Scientist Seed Grant (No. 32633), the FRQS Clinician Scientist Junior-1&2 Salary Award (No. 30861, No. 298832), and the Institut du Cancer de Montréal establishment award. DH was supported by the FRQ-S phase 1 award for medical resident engaged in clinician-scientist training. NM was supported by the International Hepato-Pancreato-Biliary Association (IHPBA) Kenneth Warren Research Fellowship and Ethicon Inc. (Johnson & Johnson).

5.11 Availability of data and materials

All raw tissue microarray images and associated marker quantification and patient clinico-pathological characteristics are available upon request from the study authors.

5.12 Ethics approval and consent to participate

This study was approved by the Centre Hospitalier de l'Université de Montréal (CHUM) institutional review board (No. 16.262) and was performed in accordance with the Declaration of Helsinki. All patients provided informed written consent to participate to the CHUM hepatopancreatobiliary cancer biobank and prospective registry associated with this study (No. 09.237).

5.13 Consent for publication

Does not apply.

5.14 Competing interests

Outside the submitted work, ST has received consultant fees from Bristol Myers Squibb and Turnstone Biologics, speaking fees from Celgene and Astra Zeneca, and has research funding from Lovance Biotherapeutics and Turnstone Biologics. JS is a permanent member of the scientific advisory board of Surface Oncology and holds stocks of Surface Oncology.

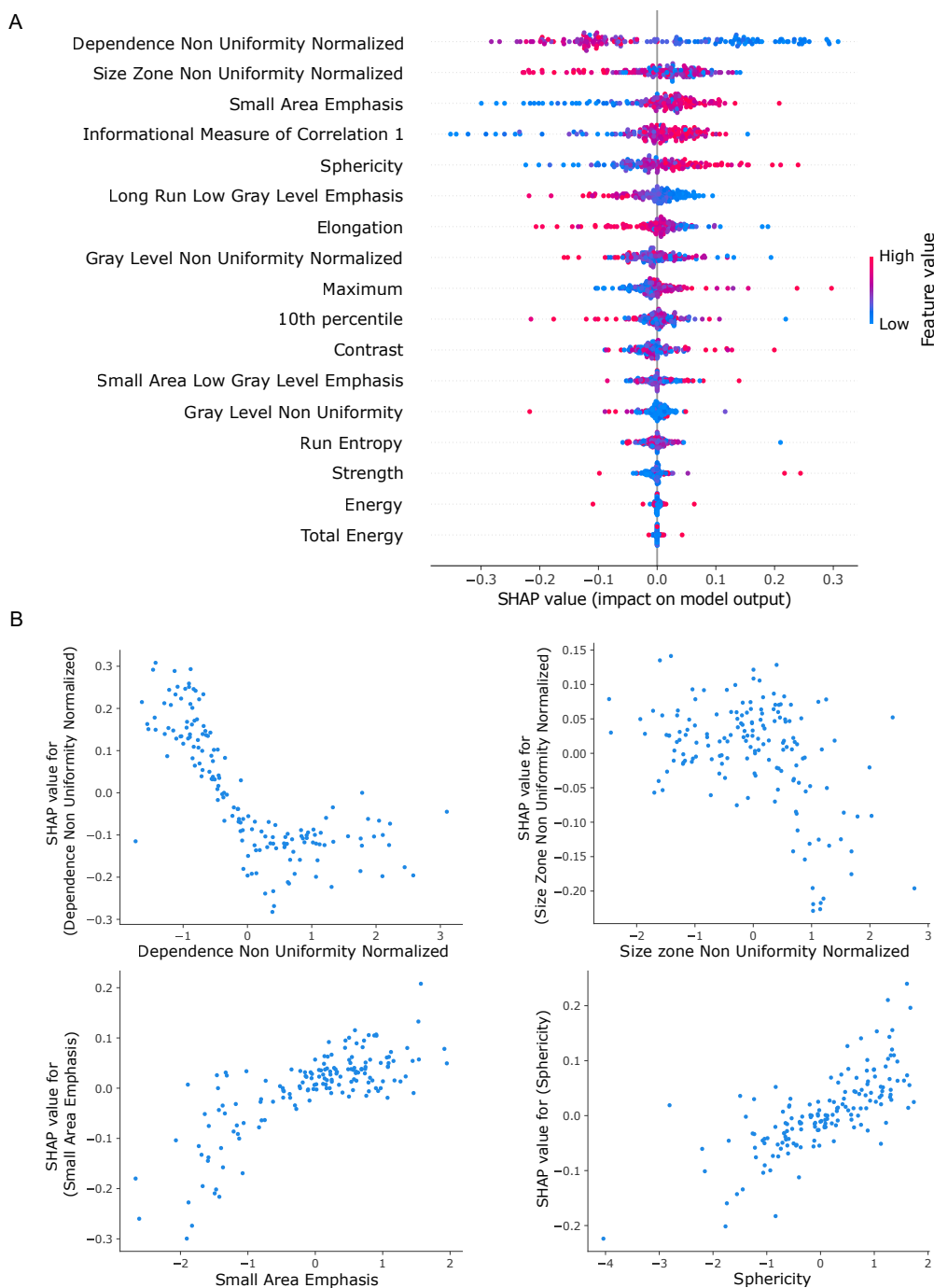


FIGURE 5.6 TabNet global interpretability analysis using the Shapley Additive Explanations (SHAP) technique. **A** Summary plot listing radiomics features from top to bottom in a decreasing order of their impact on the model's decision. The top four most significant features were texture-related features. **B** Variation of the SHAP value of four selected features with respect to the actual feature value. Homogeneous textures were associated with a higher CD73 expression, as mirrored by the effect of the top-ranking features.

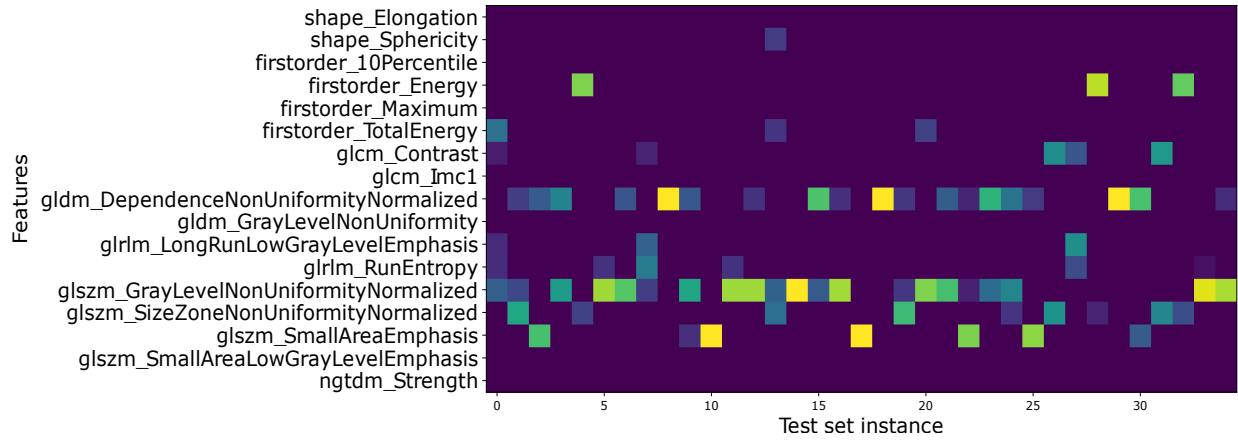


FIGURE 5.7 Visualization of TabNet learnable mask from the third decision step. The dependence non uniformity normalized (DNUN) feature, designated as the top 1 feature in SHAP analysis, was selected by the learnable mask among the most salient features for almost all test set instances. Brighter colors indicate higher feature importance.

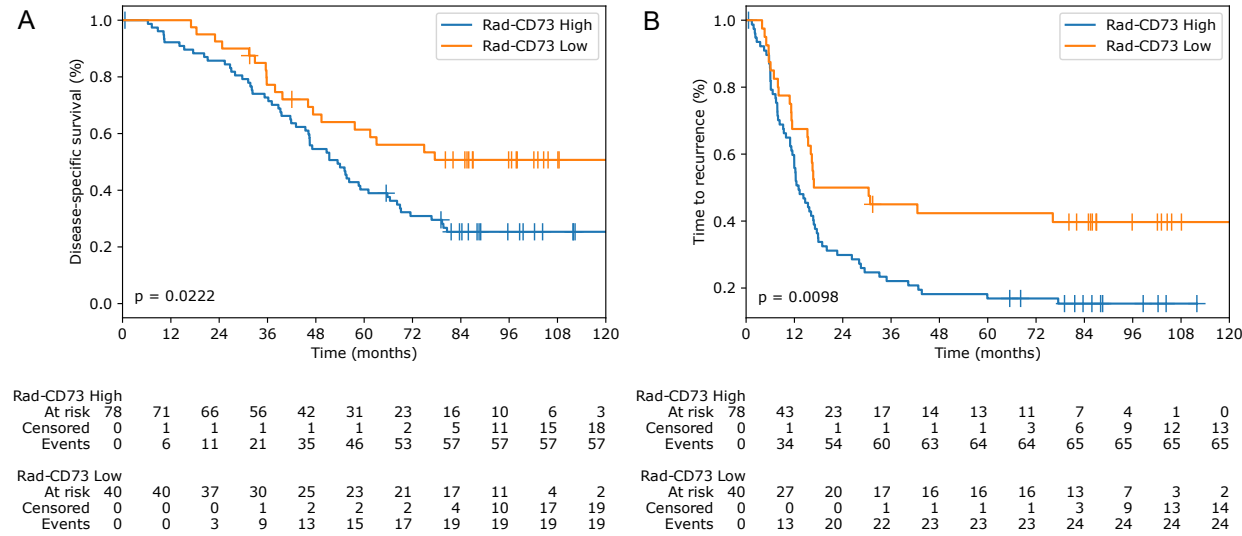


FIGURE 5.8 Prognostic value of rad-CD73 in colorectal liver metastases (CRLM). **A** Disease-specific survival according to rad-CD73 and **B** time-to-recurrence after the initial complete surgical resection of CRLM. In patients with more than one CRLM, the mean rad-CD73 was used to classify patients as low or high. The lower tertile was used as a cut-off value (rad-CD73 > 0.362).

CHAPITRE 6 ARTICLE 2 : ENSEMBLE TABNET PREDICTING A T-CELL/MHC-I-BASED IMMUNE PROFILE BIOMARKER FOR COLORECTAL LIVER METASTASES FROM CT IMAGES

Cet article a été publié dans le congrès 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI) en septembre 2023.

Remarque : L'étudiant a produit la contribution majeure à l'article : Conceptualisation, Prétraitement des données, Logiciel, Analyse formelle, Validation, Enquête, Visualisation, Revue de la littérature, Méthodologie, Rédaction (version originale), Rédaction (relecture). La contribution de l'étudiant constitue 90% du travail.

Ensemble TabNet predicting a T-cell/MHC-I-based immune profile biomarker for colorectal liver metastases from CT images

Ralph Saber^{1,2}, David Henault^{2,3}, Rolando Rebolledo^{2,3}, Simon Turcotte^{2,3}, Samuel Kadoury^{1,2}

¹Polytechnique Montréal, Montreal, Quebec, Canada

²Centre de recherche du CHUM, Montreal, Quebec, Canada

³Université de Montréal, Montreal, Quebec, Canada

6.1 Abstract

Colorectal cancer liver metastases (CLM) are the most common type of distant metastases originating from the abdomen and are characterized by a high recurrence rate after curative resection. It has been previously reported that CLM presenting a low cluster of differentiation 3 (CD3) positive T-cell infiltration density concurrent with a high major histocompatibility complex class I (MHC-I) expression were associated with poor clinical outcomes. In this study, we attempt to noninvasively predict whether a CLM exhibits the $CD3^{Low}MHC^{High}$ immunological profile using preoperative CT images. To this end, we propose an ensemble network combining multiple Attentive Interpretable Tabular learning (TabNet) models, trained using CT-derived radiomic features. A total of 160 CLM were included in this study and randomly divided between a training set (n=130) and a hold-out test set (n=30). The proposed model yielded good prediction performance on the test set with an accuracy of 70.0% [95% confidence interval 53.6%-86.4%] and an area under the curve of 69.4% [52.9%-85.9%].

It also outperformed other off-the-shelf machine learning models. We finally demonstrated that the predicted immune profile was associated with a shorter disease-specific survival ($p = .023$) and time-to-recurrence ($p = .020$), showing the value of assessing the immune response.

Colorectal cancer liver metastases, noninvasive prognostic biomarker, immune profiling, attentive interpretable tabular learning, ensemble learning.

6.2 Introduction

Colorectal cancer (CRC) ranks second in terms of cancer-related mortality worldwide [220]. Colorectal liver metastases (CLM) are the most common distant metastases of colorectal origin. In 30 to 40% of patients with metastatic CRC, CLM develop as the exclusive metastatic site [221]. Amongst CLM patients who meet the resectability criteria, 20 to 50% survive for more than five years after curative resection [222]. Nevertheless, postsurgical recurrence rates are close to 70% [223]. Several efforts have been made to understand the underlying biological processes involved in tumor genesis and identify prognostic biomarkers.

Evidence shows that cancer outcomes do not solely depend on the tumoral cells, but also on the microenvironment surrounding the tumor [12]. Particularly, tumor infiltrating lymphocytes (TIL) have been associated with patient outcomes and metastatic progression [29, 224], and emerged as a novel cancer biomarker. Studies showed a cluster of differentiation 3 (CD3) positive TIL could stratify CRC patients' survival [225]. Moreover, findings suggest that adding the major histocompatibility complex class I (MHC-I) enhances the prognostic value of CD3+ TIL in CLM [42]. A specific group of patients with CLM characterized by a low CD3+ T-cell infiltration density, concordant with a high MHC-I expression ($CD3^{Low}MHC^{High}$) was shown to have shorter disease-specific survival (DSS) and time-to-recurrence (TTR) and poorer response to chemotherapy [174].

Several machine learning methods have been proposed for disease stratification [226], as well as for the prediction of response to therapy [227]. Recent radiomic studies have shown that imaging features could be leveraged to decipher not only macroscopic patterns, but also hidden biological processes within the tumor microenvironment [58]. In this context, several studies have tested the association between radiomic features and the TIL density. Tang et al. [207] separated non-small cell lung cancer (NSCLC) patients into four clusters using CT derived radiomic features and distinguished a class of patients characterized by good outcomes and concurrent low PD-L1 expression and high CD3+ TIL density. Yoon et al. [107] proposed to predict stratified densities of type 2 helper T-cells in NSCLC patients by training a linear discriminant analysis model on radiomic features extracted from CT images and reached an

area under the curve (AUC) of 68.4% on the test set. Nevertheless, applications on lung cancer are vastly predominant. The potential of radiomic features extracted from CT images for the prediction of immunological phenotypes of CLM remains largely uninvestigated.

Ensemble learning is a technique that consists in aggregating the outputs of several predictors, hence creating a single composite model. Ensemble learning addresses limitations in small-sized datasets, as they tend to overfit on the training data. Ensembling multiple models allows to mitigate this issue by combining several learned models, leading to higher generalization capability than a single predictor. Furthermore, given that the output of an ensemble model depends on different predictors, it is less likely to get trapped in local minima during training. Moreover, having multiple models working in tandem allows to expand the search space and reduces the risk that the solution falls outside the model’s space. Finally, some ensemble learning algorithms are particularly useful when dealing with high dimensional feature sets or class imbalance [228].

In this work, we evaluated ensemble feature learning to predict if CLM exhibit the $CD3^{Low}MHC^{High}$ immune phenotype from diagnostic CT. To this end, we proposed to train an ensemble of the Attentive Interpretable Tabular Learning (TabNet) model. In a post-hoc analysis, we assessed whether the predicted immunological phenotype could stratify patient outcomes, namely the TTR and the DSS.

6.3 Materials and Methods

6.3.1 Patient cohort

In this study, 160 pathologically confirmed CLM from 122 consenting patients who underwent curative resection at the Centre Hospitalier de l’Université de Montréal, were used from our prospectively maintained dataset. The cohort had a mean age of 63.4 (35 – 84) and was predominantly constituted of male patients (63.9%). The associated clinical outcomes, namely the DSS and the TTR were also provided.

6.3.2 Data preparation

Imaging data

Presurgical CT images of CLM patients were included. The images had a transverse volume size of 512×512 , a mean axial resolution of 0.72 mm^2 (range = $[0.56, 0.98] \text{ mm}^2$) and a mean slice thickness of 2.26 mm (range = $[0.80, 5.0] \text{ mm}$). An isotropic spatial resampling was applied in order to have a uniform voxel spacing across samples ($1 \times 1 \times 1 \text{ mm}^3$). A joint

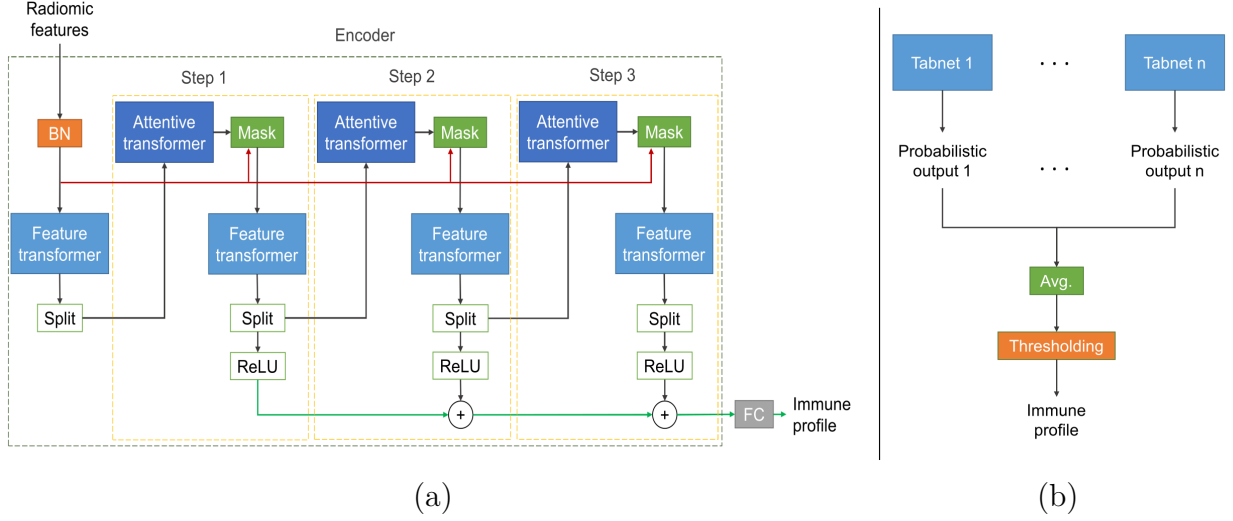


FIGURE 6.1 (a) Attentive Interpretable Tabular Learning (TabNet) model architecture. The attentive transformer outputs a feature selection mask which aim at discarding irrelevant features at a given decision step. The feature transformer processes the selected features through shared and decision step-specific layers. The model's output is obtained by adding a fully connected layer on top of the encoder's output. (b) Proposed Multi-TabNet ensemble. The multi-TabNet ensemble aggregates separately trained TabNet models through voting. Avg, Averaging; BN, Batch Normalization; FC, Fully Connected layer; ReLU, Rectified Linear Unit.

liver/hepatic automatic lesion segmentation algorithm [71] was utilized to segment the CLM. The segmentation model consists of two convolutional networks : the former segments the liver and the latter the lesions within. Segmentation masks were then validated by an experienced radiologist.

Immunological data

In this study, we used previously produced CLM immunological data [174]. For every CLM, tissue microarrays (TMA) were prepared as described in [229] using six 0.6 mm diameter intratumoral core biopsies so as to take into account tumor heterogeneity and ensure sufficient immunological representation of the entire metastatic lesion. The TMA was immunohistochemically stained for CD3 (F7.2.38, Dako, Carpinteria, CA) and MHC-I (HC10, mouse monoclonal pan-anti-human HLA class I, provided by H. Ploegh, Whitehead Institute, Cambridge, MA, USA), digitized at $\times 20$ and automated counting of CD3 and MHC-I was performed, yielding the number of CD3+ cells per mm^2 and the ratio of MHC-I positive surface area over total assessable area, respectively. Subsequently, CD3 and MHC-I values

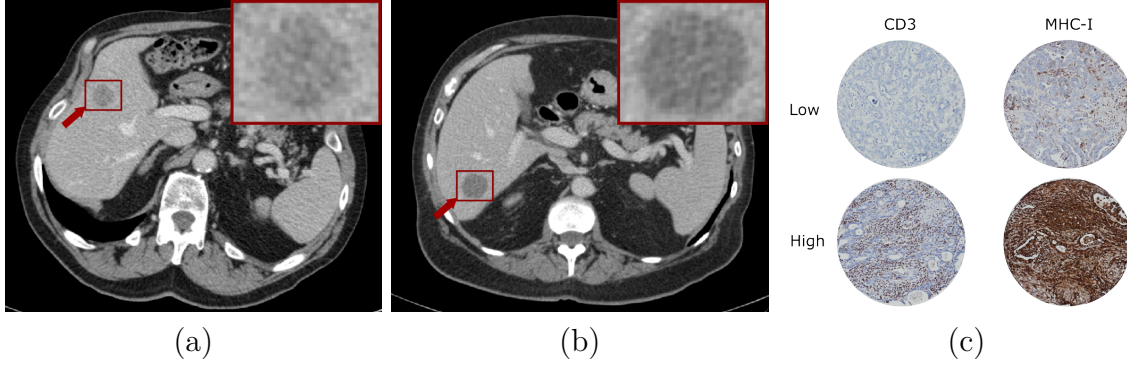


FIGURE 6.2 Representative CT images with (a) $CD3^{Low}MHC-I^{High}$ CLM and (b) $CD3^{Low}MHC-I^{Low}$ CLM. (c) Representative TMA cores with CD3 and MHC-I staining.

were stratified into low and high groups, by setting a cutoff of 932.1 cells/mm² and 45.0% for CD3 and MHC-I, respectively. The cutoff was chosen by the X-tile software [230] with the technique adopted for colorectal cancer immune scoring using the minimal p value approach with outcome data [229]. Lesions were finally divided into 2 classes according to whether they concurrently exhibit $CD3^{Low}$ and $MHC-I^{High}$ profiles (64 lesions) or not (96 lesions).

6.3.3 Radiomic feature extraction

From each segmented CLM, 107 radiomic features were extracted using the PyRadiomics v3.0.1 toolbox [190]. The rationale behind using radiomic features is twofold : (1) radiomic pipelines tend to yield better results on small-sized datasets than end-to-end models which are inherently data-hungry and prone to overfitting when trained on limited data ; (2) radiomic features have the distinctive advantage of being interpretable, which is particularly useful in the medical field where transparency is mandatory for the deployment in routine clinical practice. Features consisted of 18 first-order statistics, 14 shape, and 75 textural features. All features were standardized in order to obtain a feature set with a mean of zero and a standard deviation of one.

6.3.4 Prediction framework for T-cell/MHC-I profile

In this study, we propose a multi-TabNet ensemble model to predict whether a given CLM exhibits the $CD3^{Low}MHC^{High}$ immunological profile (Fig. 6.1).

Attentive Interpretable Tabular Learning

The Attentive Interpretable Tabular Learning (TabNet) model is a multi-stage deep learning model introduced by Google Cloud AI [149] and applies a sequential instance-wise attention mechanism allowing it to inherently select the most salient set of radiomic features at different decision steps within its architecture (Fig. 6.1). TabNet was shown to outperform the Extreme Gradient Boosting (XGBoost) classifier, previously considered as the state-of-the-art model for tabular data.

The model incorporates two key components : an attentive transformer and a feature transformer. The former aims at generating feature selection masks, which are used to discard irrelevant features at each step. For this, the attentive transformer processes features from a previous step through a single fully connected layer. It then generates sparse probabilities using a sparsemax activation while taking into account the extent to which each feature has contributed to the previous step. The feature transformer processes the ensuing filtered features using two subsequent blocks of layers : shared and decision step-dependent layers. Each block consists of two successive sets of fully connected layers, batch normalization layers (BN) and gated linear units (GLU).

The radiomic features are first sent to a BN layer and processed using a feature transformer block. Then at each step, a feature selection mask is generated by the attentive transformer and multiplied with the full feature set in order to obtain the optimal set of features in the given step. The masking operation is decision step-specific, meaning that a different set of features is selected at each step. The filtered features are subsequently processed by the feature transformer and divided into two parts : the first part is sent to the subsequent decision step while the second one is used as the output of the current decision step.

The encoder output is a linear combination of the decision step output and Tabnet's output is computed by applying a fully connected layer to the encoder's output.

Multi-TabNet ensemble model

In this work, we propose to compute the probabilistic outputs of multiple TabNet models trained separately for the prediction of the CD3/MHC-I immunological profile and then aggregate the outputs using soft voting [231] (Fig. 6.1).

In hard voting, the class with the highest number of votes is returned by the ensemble model whereas in soft voting, the class probabilities P of each TabNet model T_j , trained on the input $\mathbf{x}$, are averaged and the class i with the highest ensuing probability is picked by the model (Eq. 1) :

$$\hat{y} = \underset{i}{\operatorname{argmax}} \left(\frac{1}{n} \sum_{j=1}^n P(T_j(\mathbf{x}) = i) \right) \quad (6.1)$$

where $\hat{y}$ is the predicted class and n is the number of models in the ensemble.

6.4 Results and discussion

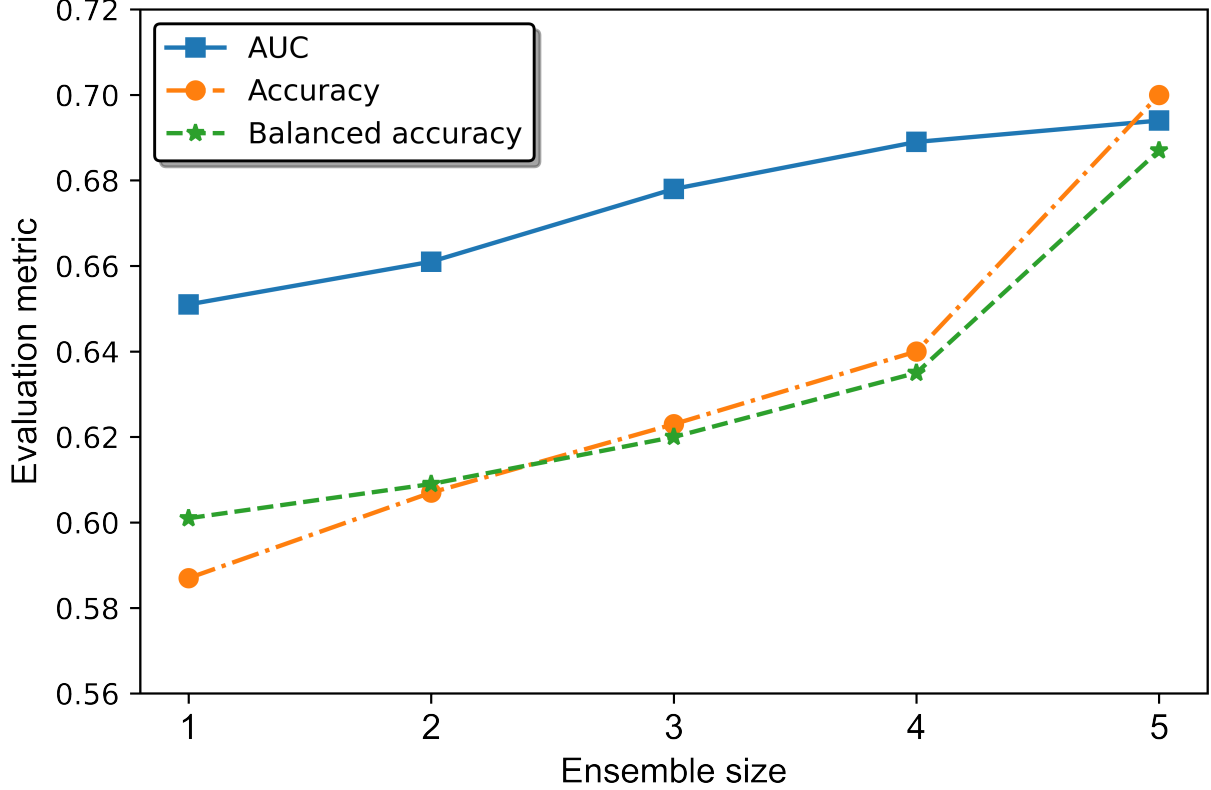


FIGURE 6.3 Evolution of Multi-TabNet ensemble performance in terms of AUC, accuracy and balanced accuracy, by ensemble size (n). For $n < 5$, the average performance across all possible combinations of n -sized ensembles is reported.

6.4.1 Evaluation methodology

The dataset was randomly divided between a training subset (130 lesions) and a hold-out test subset (30 lesions). To avoid any information crossing between the training and the hold-out test sets, all lesions of a given patient were always included in the same subset.

For benchmarking purposes, we also trained a logistic regression model (LR), a linear support vector machine (SVM) and an XGBoost classifier. For all trained models, including

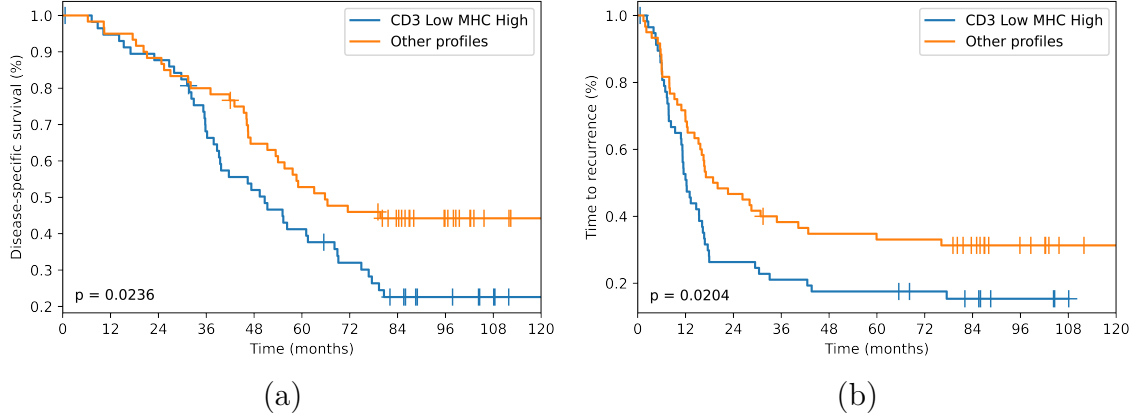


FIGURE 6.4 Prognostic value of the predicted immune profile : (a) DSS and (b) TTR following the initial CLM resection.

multi-TabNet, a 5-fold cross-validation scheme was applied on the training set for hyperparameter tuning. To handle class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) [232] was used on the training set prior to training. All models were trained to predict whether CLM exhibit the $CD3^{Low}MHC^{High}$ profile or not (Fig. 6.2).

The TabNet ensemble model was trained using an effective batch size of 64 and the Adam optimizer with a binary cross-entropy loss function for 100 epochs. Training was performed on an NVIDIA GeForce GTX TITAN Xp 12GB. The performance metrics adopted were the AUC, sensitivity, specificity, accuracy and balanced accuracy which takes into account the number of instances in each class.

In order to study the ensembling effect on TabNet models performance, we tested multi-TabNet ensembles with sizes going from one to five.

6.4.2 Prediction of the CD3/MHC-I-based immune profile

TABLEAU 6.1 Performance on the hold-out test set of the different models on the prediction of the CD3/MHC-I-based immune profile. Top results are shown in bold. 95% confidence intervals are represented between brackets.

Model	AUC (%)	Acc (%)	Sens (%)	Spec (%)
LR	48.3 [30.4-66.2]	53.3	54.5	52.6
SVM	64.1 [47.0-81.3]	60.0	63.6	57.9
XGBoost	68.9 [52.3-85.5]	63.3	54.5	68.4
Multi-TabNet	69.4 [52.9-85.9]	70.0	63.6	73.7

Table 6.1 depicts the results of the different models trained for the prediction of the immunological status : $CD3^{Low}MHC^{High}$ versus other profiles. The results show that the proposed multi-TabNet ensemble combining 5 TabNet models outperforms the other machine learning models with an AUC of 69.4% [95% confidence interval 52.9%-85.9%], an accuracy of 70.0% [53.6%-86.4%] and a balanced accuracy of 68.7% [52.1%-85.3%]. It is to be noted that the proposed model exhibits a satisfying detection of both classes as mirrored by the sensitivity and specificity values.

As shown in Fig. 6.3, when the number of aggregated TabNet models in the proposed architecture increases from 1 to 5, the accuracy, the balanced accuracy and the AUC increase from 58.7% to 70.0%, 60.1% to 68.7% and 65.1% to 69.4%, respectively. This demonstrates the benefit of ensembling on the prediction performance.

6.4.3 Prognostic value of the predicted immune profile

With the goal of assessing whether the predicted immunological profile is associated with patient outcomes, we used the Kaplan-Meier technique to draw the survival curves (DSS and TTR) up to ten years post-surgery of patients having at least one $CD3^{Low}MHC^{High}$ lesion versus the rest of the cohort. Patient stratification into the two groups was performed by setting a cut-off value on the probabilistic output of the ensemble model, chosen using the X-tile software [230]. Survival curves were compared using the log-rank test.

The Kaplan-Meier curves of Fig. 6.4 show that patients with at least one CLM with a $CD3^{Low}MHC^{High}$ profile had statistically significantly shorter DSS ($p = .023$) and TTR ($p = .020$) than the rest of the patients. This reveals that the CT-based immune profiling of the CLM holds prognostic value and could serve as a preoperative biomarker.

6.5 Conclusion

In this study, we demonstrated that a transformer-based ensemble network trained with extracted radiomic features from preoperative CT images, hold relevant biological information that could be associated with CLM immunological profiling. The proposed multi-TabNet ensemble network was capable of noninvasively identifying lesions exhibiting a low $CD3+$ TIL density and a high MHC-I expression. Moreover, experimental results showed that the proposed model yielded good results on the hold-out test set and outperformed the other machine learning models trained on the same task. Finally, we showed that the predicted profile was associated with patients' outcomes and could potentially be used as a noninvasive prognostic indicator of CLM. Prospectively, one could focus on training the model on a

larger sample size and performing an external validation to evaluate its generalizability across different centers and patient ethnicities. Furthermore, with more data available, one could attempt to make use of the CD3/MHC-I data to perform a four-class classification and assess whether each of the four classes translates into a different clinical phenotype.

6.6 Compliance with ethical standards

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Institutional Review Board.

6.7 Acknowledgments

This work was supported by the Canada Research Chairs and by the National Science and Engineering Research Council of Canada and the Université de Montréal Roger Des Groseillers Research Chair in Hepatopancreatobiliary Surgical Oncology. ST and SK are scientists of the Centre de recherche du Centre hospitalier de l'Université de Montréal (CRCHUM) supported by the Fonds de recherche du Québec - Santé (FRQ-S). ST was supported by the FRQ-S Young Clinician Scientist Seed Grant (No. 32633), the FRQS Clinician Scientist Junior-1&2 Salary Award (No. 30861, No. 298832), and the Institut du Cancer de Montréal establishment award. DH was supported by the FRQ-S phase 1 award for medical resident engaged in clinician-scientist training. The authors have no financial or non-financial interests to disclose.

CHAPITRE 7 ARTICLE 3 : MULTI-INSTITUTIONAL CT SCAN-BASED RADIOMICS FOR PREDICTING TUMOR PD-L1 EXPRESSION IN PATIENTS WITH ADVANCED AND LIMITED NON-SMALL CELL LUNG CANCER

Cet article a été soumis au journal OncoImmunology le 18 mars 2025.

Remarque : L'étudiant a produit la contribution majeure à l'article : Conceptualisation, Prétraitement des données, Logiciel, Analyse formelle, Validation, Enquête, Visualisation, Revue de la littérature, Méthodologie, Rédaction (version originale), Rédaction (relecture). La contribution de l'étudiant constitue 90% du travail.

Multi-institutional CT scan-based radiomics for predicting tumor PD-L1 expression in patients with advanced and limited non-small cell lung cancer

Ralph Saber^{1,2}, Marion Tonneau³, Olivier Salko², Moishe Liberman², Julie Malo², Arielle Elkrief², Simon Turcotte², Nicole Bouchard⁴, Philippe Joubert⁵, Samuel Kadoury^{1,2}, Bertrand Routy^{2*}

¹MedICAL Laboratory, Polytechnique Montréal, Montréal, Québec, Canada

²Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada

³Université de Lille, France

⁴Centre Hospitalier Universitaire de Sherbrooke, Québec, Canada

⁵Institut Universitaire de Cardiologie et de Pneumologie de Québec-Université Laval, Québec, Canada

7.1 Abstract

Immune checkpoint inhibitors (ICIs) have revolutionized the treatment landscape of advanced non-small cell lung cancer (NSCLC), yet 70% of patients experience disease progression, underscoring the critical need for predictive biomarkers. Programmed death-ligand 1 (PD-L1) expression remains the most adopted biomarker for ICIs. With the emergence of machine learning, the development of radiomics algorithms based on CT scan images have demonstrated potential as a novel addition to the biomarker landscape in oncology. In this

study, we aimed to develop a non-invasive surrogate of PD-L1 expression (rad-PDL1) derived from computed tomography (CT) scan imaging and compare its predictive value to pathological assessments. Furthermore, we evaluated its generalizability across advanced and limited-stage NSCLC. Radiomics features extracted from pretreatment CT were analyzed using a self-training pipeline that incorporated the feature tokenizer Transformer model to classify tumors as high vs low PD-L1 expression. We trained and internally validated our AI-based pipeline on 482 advanced NSCLC patients treated with ICIs across three medical centers. The algorithm was then validated in an independent cohort of 51 patients with limited NSCLC treated with neoadjuvant ICI and chemotherapy. Our pipeline demonstrated strong predictive performance in primary and independent validation ($AUC = 0.75$ and 0.68 , $accuracy = 0.73$ and 0.69 , respectively), highlighting its generalizability and adaptability to various disease stages. Kaplan-Meier curves revealed a longer progression-free survival for patients in the high rad-PDL1. These findings suggest that rad-PDL1 has the potential to serve as a non-invasive biomarker for predicting patient prognosis and response to ICIs.

7.1.1 Keywords

immune checkpoint blockade ; programmed death-ligand 1 ; noninvasive biomarker ; non-small cell lung cancer ; radiomics.

7.2 Introduction

Despite the paradigm shift instigated by the approval of immune checkpoint inhibitors (ICIs) for the treatment of multiple malignancies, lung cancer continues to be the leading cause of cancer-related mortality in Canada, accounting for 23% of cancer deaths in 2024 [10]. Inhibitors of programmed death-1 (PD-1) or its ligand (PD-L1), either as monotherapy or in combination with platinum-doublet represent the cornerstone of treatment for patients with advanced non-small cell lung cancer (NSCLC) and have recently expanded in the neoadjuvant setting [233–235]. Nevertheless, less than 50% of patients experience sustained clinical benefit from ICIs [236]. As a result, a considerable number of patients experience potential risks, and toxicities associated with treatments that are not guided by the projected outcome, failing to halt or impede disease progression.

Currently no single biomarker fully captures tumor biology and forecast potential resistance to ICIs [237]. Nevertheless, PD-L1 expression remains the most reliable and dictate treatment strategy for patients with advanced NSCLC without driver mutations [238,239]. The phase III KEYNOTE-024 study demonstrated that in advanced NSCLC patients with high PD-

L1 expression ($\geq 50\%$), pembrolizumab significantly extended the median progression-free survival (PFS) compared to chemotherapy (7.1 vs. 6.4 months). Moreover, pembrolizumab demonstrated superior efficacy in terms of 5-year overall survival (OS) compared to platinum-based chemotherapy, with OS rates of 31.9% and 16.3%, respectively [240]. Conversely, in the KEYNOTE 042 trial, patients with PD-L1 $< 50\%$ did not benefit from single agent anti-PD-1. Therefore, following several positive trials with combination strategies involving anti-PD-1 +/- CTLA-4 with platinum doublet have become the standard of care for patients with PD-L1 expression $< 50\%$ [241–243]. Additionally, the integration of a combination of anti-PD-(L)1 and chemotherapy has become the standard of care for patients in the neo-adjuvant/perioperative settings for patients with limited-stage disease who are eligible for surgery [244–247]. These recent studies also suggest that patients with PD-L1 $> 50\%$ derive greater benefit from neo-adjuvant chemo-IO [246]. The rapid and accurate evaluation of PD-L1 expression remains paramount in patients with NSCLC. This is currently conducted on tumor biopsies using immunohistochemistry (IHC), which is associated with procedural limitations as well as several constraints, including intratumoral heterogeneity, variability between primary tumors and metastases, fluctuations in sample quality, and differences in staining protocols [248].

Conversely, computed tomography (CT) images are widely accessible in the context of NSCLC, as they are routinely acquired for initial diagnosis and ongoing monitoring [249]. The field of radiomics, which involves the extraction of hand-engineered features from medical images, has the potential to identify underlying genomic and proteomic properties that are imperceptible to the naked eye [58]. This approach enables the detection of subtle patterns within image voxels that may reflect molecular-level processes and tumor biology. Several studies have leveraged radiomics features and machine learning models to predict tumor infiltrating lymphocytes, namely cytotoxic T cells [104] and helper T cells [107] and the intratumoral expression of immune biomarkers [172] including PD-L1 expression [86, 87]. Other works have focused on the prediction of genetic mutations for the KRAS [96] and EGFR [92, 93, 95] genes, as well as the tumor mutational burden [250]. Moreover, additional studies have sought to classify patients' PFS at 6 months after treatment [251], since this clinical outcome was identified as one of the most reliable [252]. However, most radiomics studies are limited by single-center datasets, which lack variability in patients' characteristics, data acquisition parameters, and the types of treatments administered. Moreover, these studies often lack subgroup analyses to validate the model's external validity across diverse patient subpopulations, particularly in relation to disease stage and treatment strategy. Notably, there is a paucity of data regarding the model's generalizability to NSCLC patients receiving neoadjuvant ICIs in combination with chemotherapy. They also provide minimal

insight into the machine learning models employed, which are typically treated as "black boxes." Ensuring that prediction models are generalizable across NSCLC patients receiving different treatments is essential for their clinical applicability and utility in diverse therapeutic settings. Consequently, there is a significant gap in multi-institutional research focused on developing reliable, noninvasive, and interpretable biomarkers of response to ICIs in NSCLC that can be generalized across various stages.

In this study, we present an AI-based pipeline trained on a multi-institutional dataset to predict PD-L1 expression in NSCLC patients treated with ICI. The pipeline leverages advanced deep learning techniques, including self-training to enhance its predictive accuracy and generalizability. We validate the pipeline in an independent cohort of limited-stage NSCLC patients treated with neoadjuvant ICI and chemotherapy. Additionally, we provide an interpretability analysis demonstrating that the probabilistic rad-PDL1 scores generated by the pipeline were associated with patient PFS. Our findings highlight the potential of rad-PDL1 as a non-invasive, robust biomarker for predicting response to ICIs and guiding treatment decisions in NSCLC.

7.3 Materials and methods

7.3.1 Study population and data collection

We conducted a retrospective analysis on 720 patients with advanced NSCLC. All subjects were treated between 2015 and 2022 using anti-PD-1 immunotherapy either in first-line or chemotherapy-refractory settings across three academic institutions : Centre Hospitalier de l'Université de Montréal (CHUM), Centre Hospitalier Universitaire de Sherbrooke (CHUS), and Institut Universitaire de Cardiologie et de Pneumologie de Québec–Université Laval (IUCPQ-UL). Moreover, we validated the AI-based pipeline on an independent cohort of 65 limited patients treated for NSCLC at CHUM using a neoadjuvant regimen combining chemotherapy with anti-PD-1 monoclonal antibody between 2019 and 2024. The inclusion criteria were as follows : (1) histopathologically confirmed NSCLC ; (2) PD-L1 expression assessed via immunohistochemical assays ; (3) pretreatment CT imaging acquired within three months prior to the start of treatment ; and (4) clear delineation of the primary tumor lesion on the CT image. The application of these inclusion criteria resulted in the selection of 482 patients in the primary dataset and 51 in the neoadjuvant cohort for further analysis. The "CHUM Institutional Review Board", through the Human Ethics Committee granted ethical approval for this study. Patients were either enrolled in the CHUM Lung Biobank and signed written informed consent forms or were included in project 18.255 (MP-02-2019-8091), which

was approved by the CHUM Human Ethics Committee for retrospective analysis without direct patient contact.

For self-training purposes (refer to section : Self-training-based PD-L1 predictive pipeline), we included an additional 422 NSCLC patients from the NSCLC radiomics dataset [253], publicly available on The Cancer Imaging Archive [254]. The dataset consists of publicly available CT images, accompanied by segmentations delineating the primary lesions. Clinical annotations included demographic information, treatment dates and types, the ECOG status and the smoking status as well as the clinical outcome of PFS at 6 months. Patient follow-up was conducted to document disease progression or death events for all study participants. The PFS was calculated starting from the surgical intervention date. The analysis excluded patients lacking data on their initial surgery. All patients involved in the study had a PFS greater than 6 months or experienced progression prior to this time point. PD-L1 expression in NSCLC tissue samples was measured using the Tumor Proportion Score (TPS) through IHC. PD-L1 are reported by the pathologists either $<1\%$, $1-49\%$ and $\geq 50\%$. We compared our radiological PD-L1 reading to pathological PD-L1 high ($>50\%$) vs PD-L1 low ($<1\%$ and $1-49\%$).

7.3.2 CT image preparation and pre-processing

We pre-processed the pre-treatment CT images of NSCLC patients to ensure uniform data characteristics. Spatial resampling to isotropic $1 \times 1 \times 1 \text{ mm}^3$ voxels was conducted to ensure uniform spacing for all images in the dataset, followed by truncation of Hounsfield Units (HU) to the lung window (window level = -600, window width = 1500). Subsequently, the primary lung tumor in each CT scan was segmented in the primary dataset and testing cohort of patients who underwent neoadjuvant treatment. For the primary dataset (used for training and primary testing), a semi-automatic segmentation of NSCLC was carried out as described in [251]. Primary lung tumors were annotated, and the lesions' longest axes were identified by a radiation oncologist or radiologist. The segmentation process comprised three stages : (1) morphology-based chest isolation followed by chest segmentation using connected regions [255], (2) automatic lung segmentation using the detected boundary, and (3) intersection of the lung segmentation with a clustering-based lesion mask to identify lesion segmentation, taking lung symmetry into consideration. For the more recently recruited neoadjuvant cohort (used as a hold-out test set), a recent fully automated deep learning segmentation algorithm was used to segment the NSCLC lesions [256] to account for the increased challenges in segmenting lesions with different stages. In total, 56408 slices (from 1092 images) were used for training/internal validation and 20843 slices (from 4823 images)

were used for external validation. The pipeline performs lung isolation then segments the lesions within using an adapted U-Net architecture [257]. All segmentations were reviewed by an experienced radiation oncologist and were subject to manual adjustment when needed. This pre-processing pipeline ensured standardized image quality and accurate tumor delineation for subsequent analysis.

7.3.3 Radiomics workflow

Radiomics features were extracted from each NSCLC lesion using the PyRadiomics v3.0.1 toolbox [190]. Prior to feature extraction, image filtering was performed using either a Laplacian of Gaussian (LoG) or Wavelet transformation. The LoG filter was applied to the original images with varying sigma values (0.1, 0.5, 1, 2, 3, 4, and 5), resulting in the generation of 651 LoG-derived features. In parallel, eight Wavelet decompositions were performed by applying high-pass or low-pass filters in each of the three dimensions, yielding a total of 744 Wavelet-derived features. The process resulted in a total of 1,409 radiomics features, including 14 shape-based features, 270 first-order statistics, and 1,125 texture-based features. To ensure consistency across features, standardization was applied to achieve a zero mean and unit standard deviation. Subsequently, dimensionality reduction was performed using the least absolute shrinkage and selection operator (LASSO) [191] to retain the most discriminative features for subsequent analysis.

7.3.4 Self-training-based PD-L1 predictive pipeline

The proposed AI-based pipeline combines the Feature Tokenizer Transformer (FT-Transformer) [151] with a self-training mechanism to predict whether NSCLC demonstrates a PD-L1 expression $\geq 50\%$ (classified as PD-L1^{High}) or less than 50% (classified as PD-L1^{Low}), using radiomics features extracted from the segmented primary lesion on CT as input (Figure 7.1). The FT-Transformer represents a specialized adaptation of the basic Transformer architecture, designed specifically for tabular data such as radiomics features. Its structure comprises three primary modules : the feature tokenizer, the Transformer encoder, and the classifier (Supplementary Figure 1).

To enhance the performance of the predictive pipeline and compensate for limited data, a self-training approach was designed for the FT-Transformer. First, the FT-Transformer was trained in a supervised manner utilizing solely the labeled dataset. This step produced initial predictions of marker expression, allowing for an evaluation of the pipeline’s performance. Subsequently, the trained pipeline was employed to generate pseudolabels for CT images lacking marker expression annotations, including both the public dataset and unlabeled in-

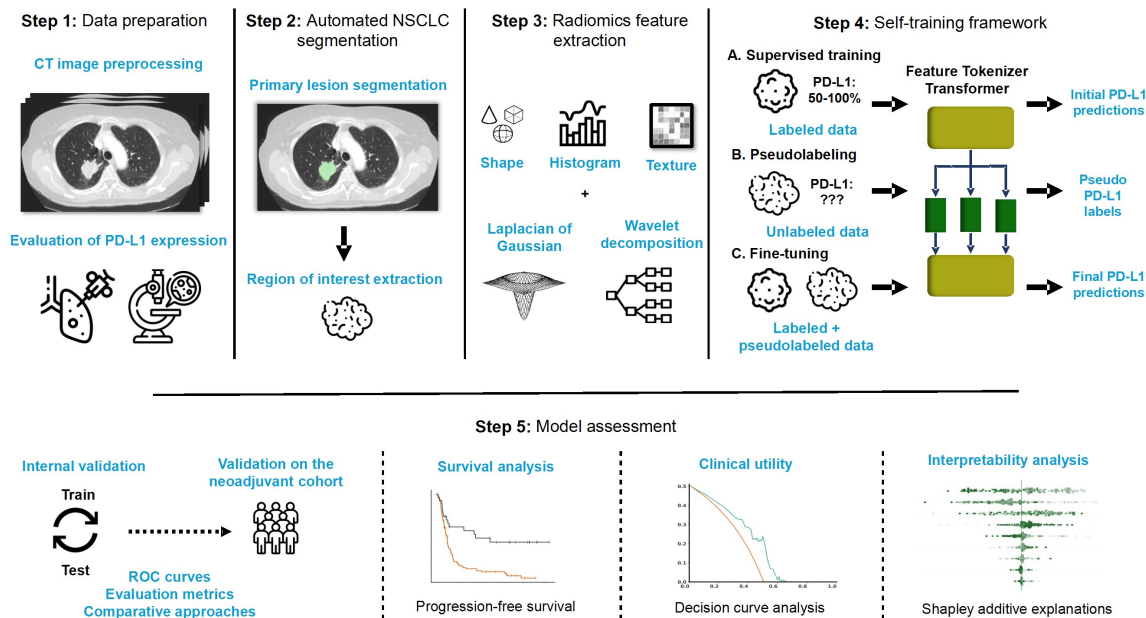


FIGURE 7.1 Overview of the AI-based PD-L1 classification pipeline. Pre-treatment CT images of patients with advanced NSCLC treated with ICIs are collected and pre-processed. The corresponding PD-L1 expression is associated and dichotomized into two classes : $<50\%$ vs $\geq 50\%$. The ensuing images are segmented and radiomics features (shape, texture and histogram-based features) are extracted from the resulting lesions after the application of a Laplacian of Gaussian filter or a wavelet decomposition to the image. A self-training framework is then adopted in which the deep learning model, the Feature Tokenizer Transformer, is first trained using labeled data and then leveraged to generate pseudolabels for instances with no PD-L1 expression information. Labeled and pseudolabeled data are subsequently combined to fine tune the model and improve its predictive performance. The model is finally validated on the internal test set and a separate cohort of patients who underwent neoadjuvant treatment combining ICIs and chemotherapy.

house instances. Pseudolabeling was exclusively applied to instances for which the pipeline exhibited the highest confidence. This approach was motivated by the fact that the majority of erroneous predictions made by classifiers tend to occur in proximity to the decision boundary [258]. Consequently, pseudolabels were assigned only to instances for which the pipeline produced probabilistic outputs that fell outside the decision boundary. Following the pseudolabeling process, the pipeline underwent fine-tuning using both the originally labeled data and the newly pseudolabeled instances.

For training and internal validation, the primary dataset ($N = 482$) was randomly partitioned into a training subset (80%) and a test subset (20%), while ensuring that all splits were patient-based. Pipeline training was conducted on an NVIDIA Titan RTX GPU equipped

with 64 GB of RAM. For benchmarking purposes, the proposed pipeline’s performance was compared to the following machine learning and deep learning models, trained using the same data to perform the same PD-L1 classification task ($\geq 50\%$ vs $< 50\%$) : (1) extreme gradient boosting classifier (XGB), (2) Support Vector Machine (SVM), (3) random forest (RF), (4) Self-Attention and Intersample Attention Transformer (SAINT) [259]. Additionally, we evaluated the models aforementioned, which were trained using radiomics features, against two end-to-end models trained using voxel intensities directly : (1) the standard Vision Transformer (ViT) and (2) the Compact Convolutional Transformer (CCT) [260], designed specifically for small datasets. For these two models, 3D volumes of NSCLC tumors were resized to $50 \times 50 \times 50$, and voxel intensities were normalized to achieve a zero mean and unit standard deviation. To mitigate overfitting, data augmentation was applied through random horizontal and vertical flipping during training. The binary cross-entropy loss function and the Adam optimizer were utilized to train all deep learning models.

7.3.5 Interpretability of the AI-based pipeline

We conducted a comprehensive analysis of the pipeline’s behavior by applying the Shapley Additive Explanations (SHAP) technique to the produced predictions [261], for both the primary and neoadjuvant treatment regimen cohort. SHAP is a technique that calculates Shapley values for each feature, representing their individual contributions to the pipeline’s output. We then ranked features based on their average Shapley values across all instances. Positive Shapley values suggest that a feature influences the pipeline towards predicting high scores, while negative values indicate a tendency towards low score predictions.

7.3.6 Statistical analysis

Statistical analysis was performed using non-parametric methods. For comparisons of numerical variables between groups, the Wilcoxon rank sum test was employed. For categorical variables, Fisher’s exact test was utilized. The decision curve analysis was plotted using Python dcurves v1.1.0 package. For each threshold probability the net benefit was calculated. Net benefit was computed as a weighted combination of true positives and false positives, where the weight is derived from the threshold probability. Specifically, it represents the proportion of patients correctly identified for treatment minus the proportion incorrectly identified, adjusted by the threshold probability. Kaplan-Meier survival curves were produced using the Python Lifelines v0.27.8 library, with statistical comparisons conducted using the log-rank test. To perform survival analysis, we derived the FT-Transformer class probabilities, denoted as the rad-PDL1 score. These probabilities were then categorized into rad-PDL1^{High} and

rad-PDL1^{Low} groups. The threshold for this stratification was determined using the X-tile software, which proposed a cut-off value based on the rad-PDL1 score distribution. To assess the independent prognostic value of the rad-PDL1 score and other clinical characteristics simultaneously, we carried out a multivariate Cox proportional hazards regression analysis. This method generated hazard ratios (HR) along with their corresponding 95% confidence intervals. Statistical significance was defined as a two-sided p-value less than 0.05.

7.4 Results

7.4.1 Patient cohort

The clinicopathological properties of the NSCLC patients for both the primary cohort (N = 482 patients) and the cohort of patients who underwent neoadjuvant treatment (N = 51 patients) are detailed in Table 7.1. In the primary cohort, the mean age was 66.7 (range : 39-89), with a roughly equal representation of men (51%) and women (49%). In the neoadjuvant cohort, the mean age was 65.1 (range : 37-80), with a predominance of male patients (55%). Active smokers comprised 29% and 27% of patients in the primary testing and neoadjuvant cohorts, respectively. Approximately half of patients had a PD-L1 expression $\geq 50\%$ in the primary testing set, compared to 59% in the neoadjuvant cohort. The proportion of patients with a PFS greater than 6 months was 54% and 82% in the primary testing set and the neoadjuvant cohorts, respectively. The median PFS was 6.6 months in the primary cohort and 15.2 months in the neoadjuvant cohort. Upon statistical evaluation, the aforementioned features showed no significant differences between the two cohorts, except for the PFS (p = 0.0001) and the ECOG status (0.020).

7.4.2 Prediction of PD-L1 expression and validation on the neoadjuvant cohort

We assessed the prediction performance of the AI-based pipeline on the primary testing set of the primary cohort and its generalizability on the cohort of patients who underwent the neoadjuvant treatment regimen, by comparing the pipeline predictions with the expression values obtained by IHC. In the primary testing set of the primary cohort, the proposed pipeline's ability to distinguish between rad-PDL1^{High} and rad-PDL1^{Low} was demonstrated by achieving an AUC of 0.75 (95% confidence interval : 0.66-0.83). The accuracy, precision, recall and F1-score were 0.73 (0.64-0.82), 0.73 (0.64-0.81), 0.76 (0.67-0.84) and 0.74 (0.65-0.83), respectively. Furthermore, the pipeline exhibited a good generalization capability on the external testing cohort, comprised of patients receiving neoadjuvant treatment, with an AUC of 0.68 (0.55-0.81). The accuracy, precision, recall and F1-score were 0.69 (0.56-0.81), 0.72

TABLEAU 7.1 Clinicopathological characteristics of the primary and neoadjuvant cohorts.

	Primary cohort (N=482)	Neoadjuvant cohort (N = 56)	p-value
Age	66.7 (39-89)	65.1 (37-80)	0.244
Sex - n			0.659
Male	244 (51)	28 (55)	
Female	238 (49)	23 (45)	
Smoking status - n			0.869
Current	139 (29)	13 (27)	
Former or never	338 (71)	36 (73)	
ECOG status - n			0.02
>1	67 (14)	1 (2)	
<=1	415 (86)	46 (98)	
PD-L1 status - n			0.303
>=50	243 (50)	30 (59)	
<50	239 (50)	21 (41)	
Progression -free survival - n			0.0001
>=6 months	257 (54)	41 (82)	
<6 months	220 (46)	9 (18)	

(0.60-0.84), 0.77 (0.65-0.88) and 0.74 (0.62-0.86), respectively. In both internal and external validation scenarios, the pipeline’s performance was characterized by balanced precision and recall, indicating an absence of class bias (Figure 7.2).

There was a statistically significant difference in the predicted rad-PDL1 score between the PDL1^{High} and PDL1^{Low} tumors in both the primary testing cohort (Mann-Whitney U test ; $p < 0.0001$) and the neoadjuvant cohort ($p = 0.0298$). Moreover, the predictive performance of the self-training pipeline leveraging the FT-Transformer was compared to other deep and conventional machine learning models. The FT-Transformer trained using a self-training framework outperformed deep and machine learning models in PD-L1 classification. The models trained using radiomics features outperformed the end-to-end deep learning models, namely the ViT and the CCT (Supplementary figure 2).

7.4.3 Interpretability analysis

After training and evaluating the PD-L1 prediction pipeline, we sought to interpret the pipeline’s predictions by evaluating the relative contribution of each radiomics feature using the SHAP technique. Figure 7.3 shows the top ranking radiomics features according to their average Shapley values, which reflect their impact on the pipeline’s output. The top six features were exclusively textural, emphasizing the significance of the textural properties of

NSCLC lesions in classifying PD-L1 expression.

The top feature returned by the SHAP analysis was the Size Zone Non-Uniformity Normalized (SZNUN) extracted after the application of a wavelet decomposition using high-pass filters along the x and y-axes and a low pass filter along the z-axis (Wavelet-HHL). The Size Zone Non-Uniformity Normalized (SZNUN) measures the variability of size zone volumes in the lung lesion, normalized by the square of the total number of zones. A lower SZNUN value indicates greater homogeneity among the size zone volumes in the image. The SHAP analysis revealed that higher SZNUN values corresponded to positive Shapley values, influencing the pipeline to predict elevated PD-L1 expression levels. This suggests that increased heterogeneity in size zone volumes within the image is associated with higher predicted marker expression. The second top ranking radiomics feature was ClusterShade, extracted after the application of a LoG filter (sigma value = 0.5 mm) to the image. The Clustershade feature encodes the skewness or asymmetry in the image texture. Low ClusterShade values suggest more symmetry and uniformity in the texture. The latter feature demonstrated a pattern comparable to the SZNUN, where elevated feature values corresponded to positive Shapley values. Consequently, this relationship led to predictions of high PD-L1 expression scores by the pipeline. The results indicate that the pipeline was inclined to predict elevated rad-PDL1 scores for lesions characterized by more intricate or heterogeneous textural patterns.

In line with these findings, Figure 7.4 shows representative cases of NSCLC tumors with low and high PD-L1 expressions, respectively. The pretreatment CT image in subfigure (a) demonstrates an NSCLC in the right lung with a high PD-L1 expression (TPS : 50-100%), as evidenced by IHC. Correspondingly, the pipeline classified it as a rad-PDL1^{High} NSCLC. In contrast, the NSCLC observed in the CT image of subfigure (b) has a low PD-L1 expression (TPS <50%) and was classified as rad-PDL1^{Low}. These cases confirm the hypothesis that distinct CT features can be exhibited between PD-L1^{High} and PD-L1^{Low} NSCLC. This observation suggests a potential correlation between PD-L1 expression levels and radiographic characteristics of NSCLC, which has potential implications for non-invasive tumor assessment and treatment planning.

7.4.4 Decision curve analysis

Figure 7.5 shows that the net benefit provided by the AI-based pipeline is higher than the treat all as PD-L1 high (orange curve) and the treat none as PD-L1 high (black line) approaches, for all probability thresholds in primary testing and for probability thresholds greater or equal to 0.25 in the neoadjuvant external validation cohort. Therefore, our proposed pipeline demonstrates superior clinical utility by precisely identifying patients who will

truly benefit from treatment, effectively distinguishing between genuine positive candidates and those unlikely to respond. Unlike broad treat all or treat none approaches, our method optimizes patient outcomes by targeting interventions more accurately, reducing unnecessary treatments and potential side effects.

7.4.5 Association of rad-PDL1 with clinical outcomes

We finally evaluated the potential clinical impact of categorizing patients based on high versus low expression levels of rad-PDL1 (Figure 7.6). Consistent with the improved prognosis observed in patients with high PD-L1 expression treated with ICIs [262–264], patients with rad-PDL1^{High} NSCLC had a longer median PFS than patients with rad-PDL1^{High} NSCLC in the primary testing set (8.08 vs 3.12 months ; $p = 0.02$). In the neoadjuvant cohort, the median PFS in the rad-PDL1^{High} vs rad-PDL1^{Low} groups was 18.6 vs 13.8 months ($p = 0.09$). The incidence of pathologic complete response in the neoadjuvant cohort was 45.5% and 17.9% in the rad-PDL1^{High} and the rad-PDL1^{Low} groups, respectively ($p = 0.10$). The median OS was 15.2 vs 12.5 months and 18.6 vs 14.8 months in rad-PDL1^{High} vs rad-PDL1^{Low} groups, in the primary and neoadjuvant cohort respectively ($p > 0.05$). The multivariate analysis (Table 7.2) confirmed that the prognostic significance of the rad-PDL1 marker for PFS was independent of the clinicopathological characteristics ($p = 0.01$).

TABLEAU 7.2 Multivariate analysis of the progression-free survival (PFS) according to the rad-PDL1 score and clinical features. The rad-PDL1 score was a predictor of the PFS, independently of the clinical features.

	Progression-free survival	
	HR (95% CI)	p-value
rad-PDL1 (<50% vs $\geq$ 50%)	0.45 (0.24-0.81)	0.01
Age (<60 vs $\geq$ 60)	0.88 (0.52-1.47)	0.62
Sex (female vs male)	0.95 (0.62-1.45)	0.8
Current smoker (no vs yes)	0.92 (0.58-1.46)	0.73
ECOG (<2 vs $\geq$ 2)	1.6 (0.75-3.38)	0.22

7.5 Discussion

The PD-L1 expression, assessed via IHC, is currently the only approved biomarker by the Food and Drug Administration (FDA) for guiding ICI therapy in advanced NSCLC [251]. It is also the most well-established and widely utilized biomarker for ICIs in routine clinical practice [265]. Positive PD-L1 expression was linked to better outcomes and higher clinical

response rates compared to patients with negative PD-L1 expression [266]. In this study, we proposed an AI-based pipeline based on self-training, Transformer networks and radiomics features extracted from pretreatment CT, to classify PD-L1 expression in NSCLC amenable to ICIs. We also validated the pipeline on an independent cohort of patients with limited NSCLC treated with a neoadjuvant regimen combining ICIs and chemotherapy. Our pipeline exhibited strong predictive performance on both datasets, highlighting its high generalization capability and versatility across different treatment regimens. The PD-L1 prediction performance of machine learning pipelines in NSCLC should be generalizable across patients receiving different treatments because oncological response and outcomes can vary significantly based on these combinations. The CheckMate 816 trial showed that neoadjuvant nivolumab plus chemotherapy significantly improved event-free survival and pathological complete response rates compared to chemotherapy alone in resectable NSCLC, providing long-term clinical benefits [244]. Therefore, a pipeline that non-invasively predicts PD-L1 expression accurately across various treatments ensures its utility in diverse clinical scenarios.

Furthermore, the clinical relevance of the rad-PDL1 score derived from the pipeline was corroborated by its significant association with oncological outcomes, demonstrating its potential prognostic value. Its utility in the clinical setting was demonstrated by the decision curve analysis results obtained in both cohorts. By strategically selecting patients with the highest probability of positive response using the rad-PDL1 score, the pipeline not only improves individual net benefit but also enhances overall healthcare resource allocation, striking a critical balance between medical effectiveness and efficient resource management.

The noninvasive assessment of PD-L1 expression using CT images provides several advantages over traditional IHC-based methods, based on biopsies. Notably, the high accessibility and cost-effectiveness of this approach make it a valuable tool for dynamically monitoring treatment responses. This is particularly relevant in NSCLC, as it enables longitudinal studies to repeatedly assess PD-L1 expression through follow-up scans, facilitating the detection of acquired resistance to ICIs during treatment. Such insights could guide timely adjustments to therapeutic strategies, improving patient outcomes [236, 267]. Moreover, the noninvasive approach can be leveraged to evaluate spatial heterogeneity between the primary tumor and metastases. Metastatic lesions may exhibit disparate PD-L1 positivity rates, which can impact treatment decisions and response to ICIs [268, 269]. Additionally, understanding the variation in PD-L1 expression profiles may assist physicians in selecting optimal biopsy locations [270], thus reducing false-negatives.

This work leverages advanced deep learning techniques, including self-training and the feature tokenizer Transformer, which improves the classification accuracy and generalizability

of the AI-based pipeline. In computer vision applications, Transformers excel at focusing on key image regions, but their data requirements and susceptibility to overfitting make them challenging with limited data, especially in the high-dimensional space of CT images. Radiomics provides a valuable alternative to end-to-end pipelines by reducing the feature space and extracting more well-established interpretable features. Furthermore, this allowed to conduct a post-hoc analysis by employing the SHAP method, allowing for a detailed investigation into the pipeline’s decision-making process. High rad-PDL1 scores were associated with intratumoral textural heterogeneity in NSCLC. These findings are concordant with previous work showing that high gray level co-occurrence matrix entropy, a feature encoding textural randomness and irregularity, were correlated with high PD-L1 expression scores [271].

Nevertheless, this work presents some limitations. First, while the multi-institutional dataset includes patients treated at three distinct medical centers and utilizes a variety of scanners, including those from Siemens, GE Medical Systems, Philips, and Toshiba, all hospitals are located within the same province in Canada with a relatively limited number of patients. To achieve a more epidemiologically diverse dataset, large scale prospective multi-center studies involving patients from different regions in the world may be required. Second, the work relies solely on CT scans. Multimodal approaches involving different imaging modalities and data sources might boost the performance of our predictive pipeline further. In fact, recent studies have demonstrated the association between positron emission tomography (PET) features and PD-L1 expression in NSCLC [272,273]. Consequently, integrating PET and CT features using efficient fusion techniques may prove beneficial in this context. Moreover, the two imaging modalities could be combined with other data types to develop an integrative and reliable survival pipeline. Third, the interval between pretreatment CT imaging and histopathological analysis varied among patients, potentially allowing for temporal fluctuations in PD-L1 expression due to its dynamic nature. To mitigate this issue, we restricted our study to CT scans acquired within a three-month window prior to treatment initiation. This approach aims to minimize the impact of temporal heterogeneity on the assessment of PD-L1 expression and its relationship to imaging features.

7.6 Conclusion

In summary, this study introduced a reliable, noninvasive and interpretable imaging surrogate of PD-L1 expression in NSCLC. The proposed CT-based immune marker, termed rad-PDL1, exhibited a strong generalization capability across patients treated with advanced and limited NSCLC. The versatility of the proposed biomarker and its observed association with oncological outcomes suggest its potential utility as an adjunctive tool for oncologists

to : 1) select suitable candidates for ICIs; 2) noninvasively assess PD-L1 expression in patients with limited NSCLC amenable to neoadjuvant chemo-ICIs; 3) evaluate primary and acquired resistance to ICIs; 4) identify patients who might benefit from a closer follow-up or a potential treatment modification; and 5) understand spatial and temporal heterogeneity of the tumor microenvironment. In future investigations, our proposed AI-based pipeline could be evaluated and potentially adapted for the prediction of PD-L1 expression in other malignancies where it has received FDA approval for guiding ICIs such as bladder, cervical and triple-negative breast cancer. In addition, a comprehensive international database could be leveraged to develop a multimodal approach to reliably predict PD-L1 expression and prognosis of patients with disparate epidemiological characteristics worldwide.

7.7 Acknowledgments

The authors acknowledge Imagia for providing semi-automatic segmentations based on the lung nodule annotations.

7.8 Declaration of interest statement

The authors state that there are no conflicts of interest regarding this work.

7.9 Funding

This research was supported by the Natural Sciences and Engineering Research Council of Canada (SK), the Institut du Cancer du Poumon (BR), and the TransMedTech Institute (BR). BR and PJ received funding from the Fonds de recherche du Québec - Santé (FRQ-S). RS is supported by the Fonds de recherche du Québec – Nature et Technologie (FRQ-NT). The recruitment of patients and the annotation of images at IUCPQ-UL were facilitated by the Quebec Respiratory Health Research Network-IUCPQ site and Oncotech grant (PJ).

7.10 Data availability statement

Imaging data and associated clinicopathological characteristics and marker quantification are available upon request from the authors.

7.11 Authors' contributions

BR and SK performed project administration and conceptualization. MT, OS, JM and RS performed data curation. PJ, NB, AE, ML, BR worked on resource gathering. SK, ST, BR performed supervision tasks. BR, SK and PJ reviewed and edited the manuscript and acquired funding. RS implemented the methodology, performed formal analysis, investigation, data curation, software analysis, validation, visualization, funding acquisition, writing original draft and editing.

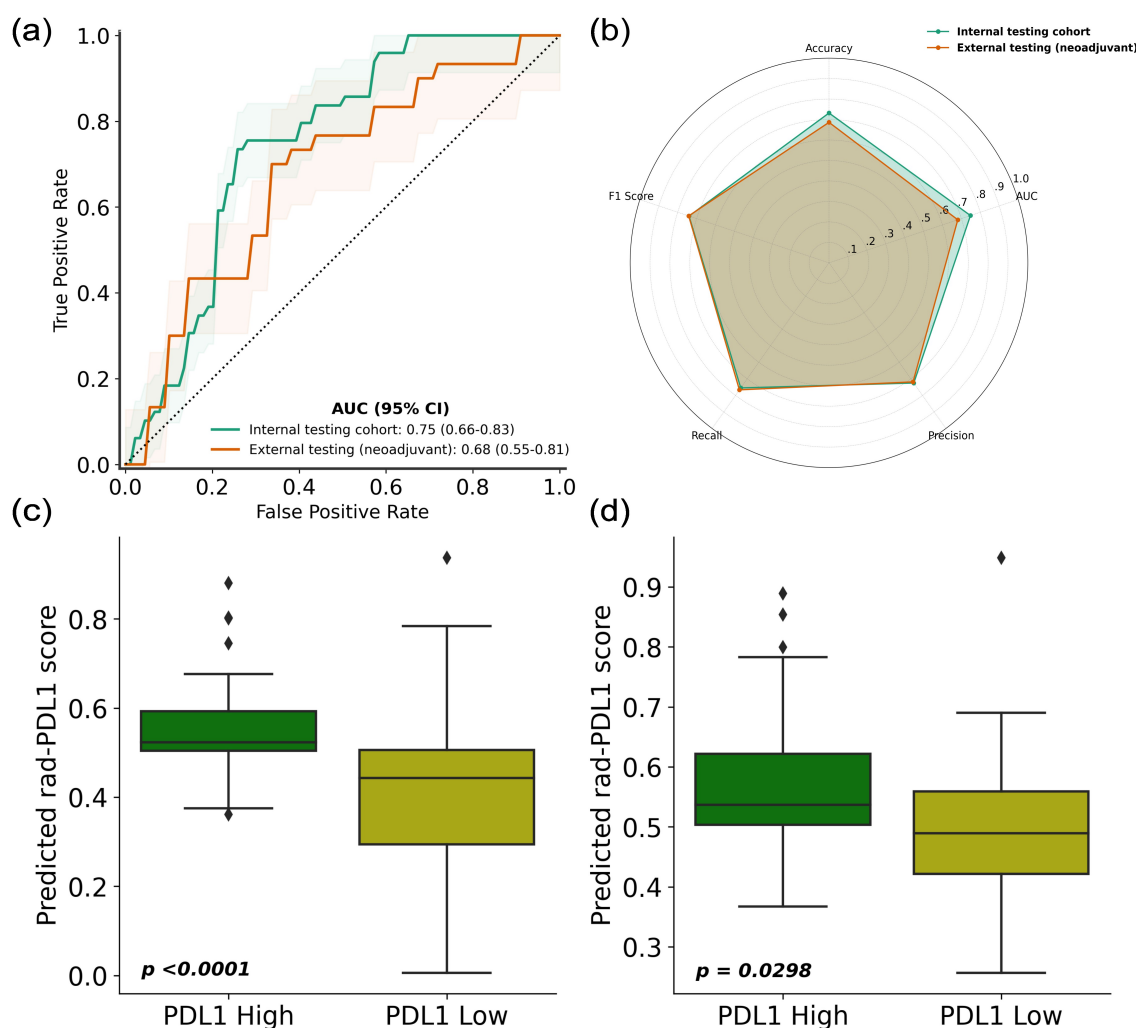


FIGURE 7.2 PD-L1 classification results of the predictive pipeline. (a) Receiver operating characteristics curve on the primary testing dataset (green) and the external validation cohort consisting of patients who underwent neoadjuvant treatment (orange). (b) Spider plot showing the performance metrics of the predictive pipeline on the internal and external validation cohorts. AUC; area under the receiver operating characteristic curve. (c-d) Evaluation of the probabilistic rad-PDL1 score obtained using the predictive pipeline. The box plots present the distribution of rad-PDL1 by PD-L1 expression level (high vs low). The data indicated a statistically significant difference in rad-PDL1 between the PD-L1High and PD-L1Low groups in both (c) the primary testing cohort (Mann-Whitney U test; $p < 0.0001$) and (d) the external neoadjuvant cohort ($p = 0.0298$).

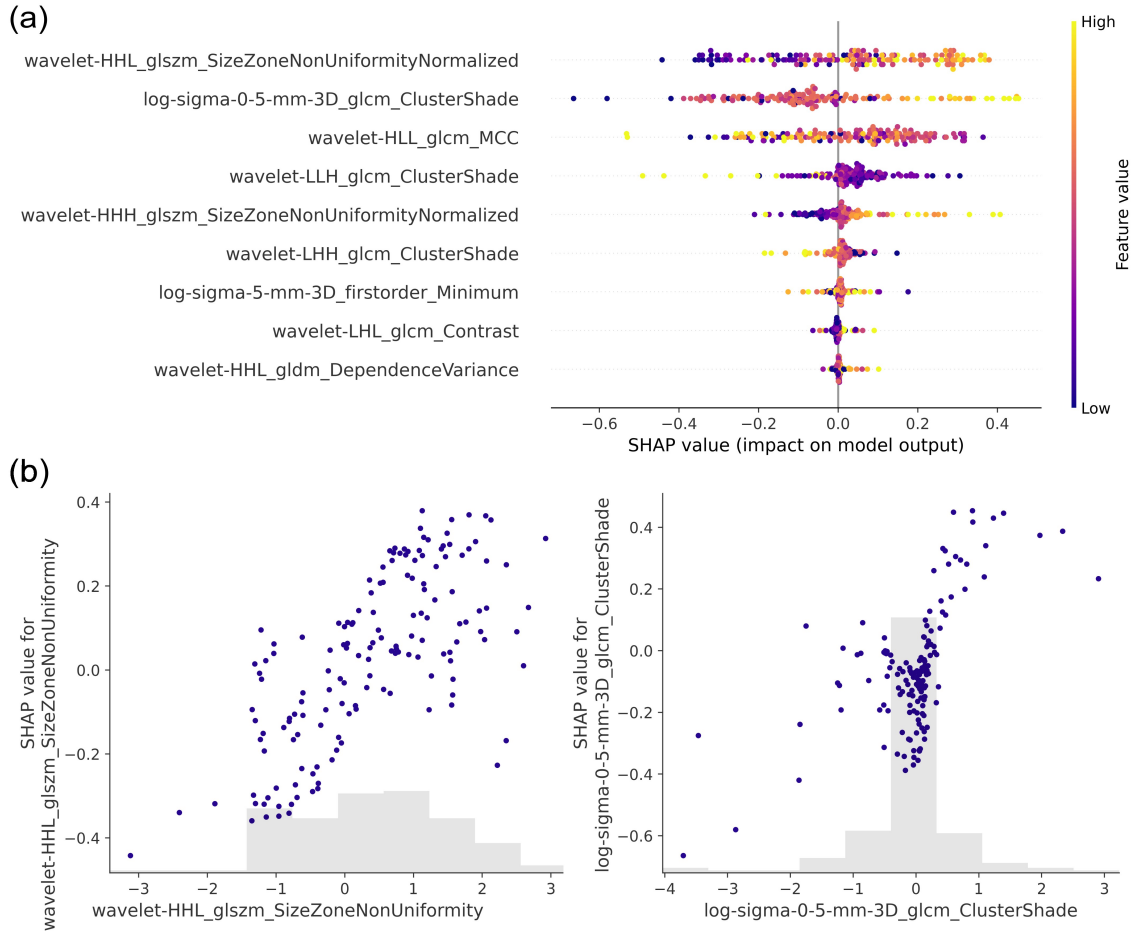


FIGURE 7.3 Interpretability analysis of the predictive pipeline. (a) Summary plot showing the top features obtained using the Shapley Additive Explanations (SHAP) technique, according to their average Shapley values, which reflect their impact on the model's output. The top six features were found to be textural features, which highlights the importance of texture-related properties of NSCLC lesions in classifying NSCLC. (b) Variation in SHAP values of the top two radiomics features as a function of their actual values. The results show that heterogeneous NSCLC lesions were associated with a high PD-L1 expression.

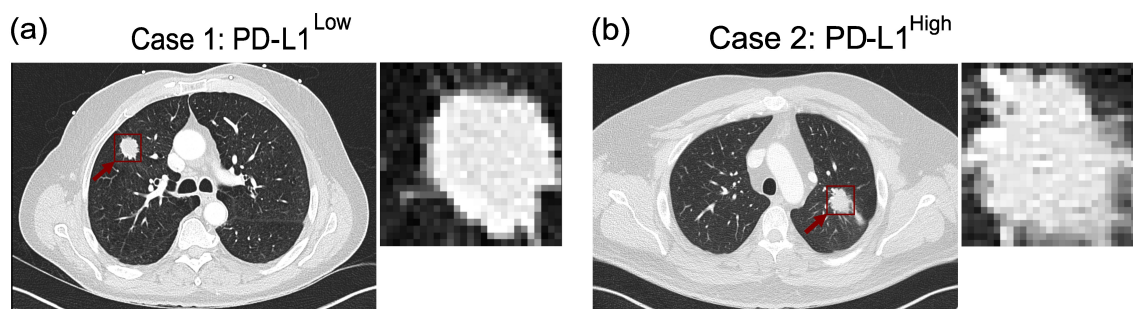


FIGURE 7.4 Representative cases with low and high PD-L1 expression. (a) A pretreatment CT image of a PD-L1^{High} NSCLC (arrow) in the right lung. (b) A pretreatment CT image of a PD-L1^{Low} NSCLC (arrow) in the left lung.

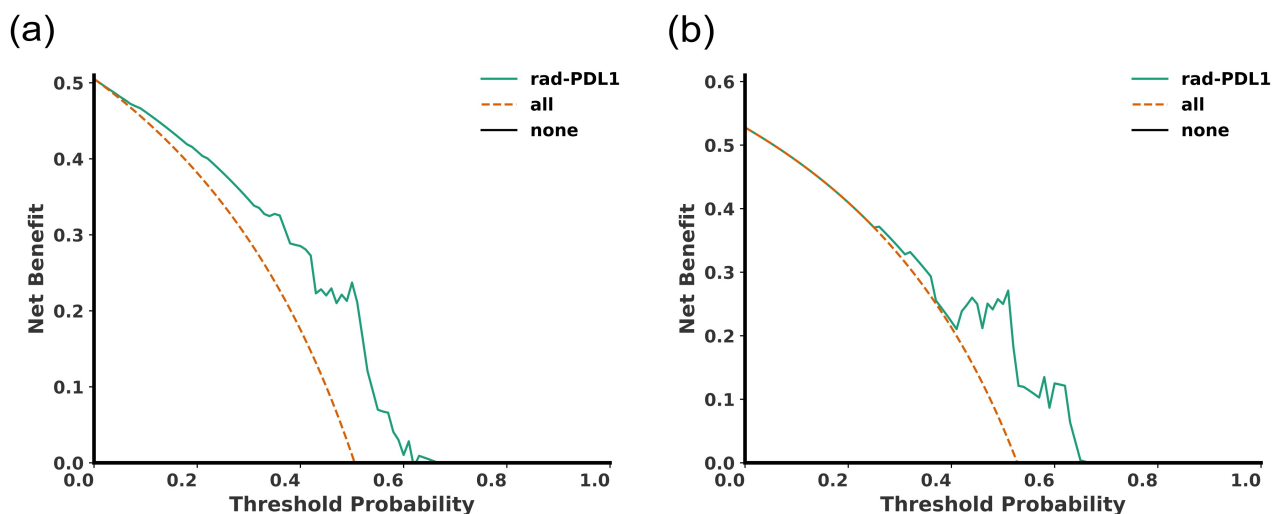


FIGURE 7.5 Decision curve analysis by modality on (a) the primary and (b) the external (neoadjuvant chemotherapy) testing cohorts. The net benefit is computed for each threshold probability, representing the acceptable risk beyond which acting (treatment, further testing, etc.) is advised. The dashed orange curve represents the approach where all patients are considered to have a high expression of PD-L1 and the black x-axis represents the approach where no patient is considered to have a high PD-L1 expression. As shown in the plots, the net benefit of the proposed pipeline (green curve) is higher than the “treat all patients as high PD-L1” (orange curve) and “treat no patient as high PD-L1” (black line) approaches in both the primary and external (neoadjuvant) testing cohorts.

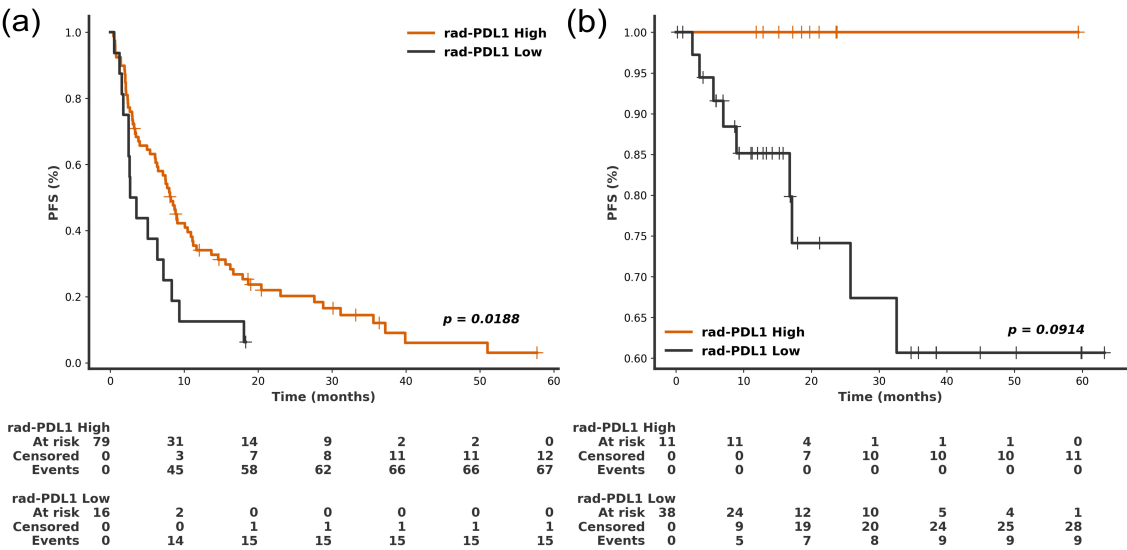


FIGURE 7.6 Prognostic significance of the predicted rad-PDL1 score in (a) the primary testing cohort and (b) the external testing cohort composed of patients receiving neoadjuvant treatment. The progression-free survival was significantly shorter in the rad-PDL1Low group compared to the rad-PDL1High in the primary testing cohort ($p = 0.02$).

CHAPITRE 8 ARTICLE 4 : MULTI-OMICS FUSION NETWORK FOR PREDICTION OF EARLY RECURRENCE IN COLORECTAL LIVER METASTASES

Cet article a été soumis au Journal of Biomedical and Health Informatics le 21 janvier 2025. L'étude a été financée par les Fonds de Recherche du Québec - Nature et Technologie (FR-QNT).

Remarque : L'étudiant a produit la contribution majeure à l'article : Conceptualisation, Prétraitement des données, Logiciel, Analyse formelle, Validation, Enquête, Visualisation, Revue de la littérature, Méthodologie, Rédaction (version originale), Rédaction (relecture). La contribution de l'étudiant constitue 90% du travail.

Multi-omics fusion network for prediction of early recurrence in colorectal liver metastases

Ralph Saber, Mayra Carneiro, Emmanuel Montagnon, An Tang, Simon Turcotte, and Samuel Kadoury

Ralph Saber and Samuel Kadoury are with the MedICAL laboratory at Polytechnique Montréal, Montreal, QC, Canada. Simon Turcotte is with the Hepatopancreatobiliary Surgery Service, Centre hospitalier de l'Université de Montréal, Department of Surgery, Université de Montréal. Simon Turcotte, Samuel Kadoury, An Tang, Mayra Carneiro, Emmanuel Montagnon, Samuel Kadoury and Ralph Saber are with the Research Center of the University of Montreal Hospital (CRCHUM), Montreal, QC, Canada.

8.1 Abstract

Liver metastases develop in nearly half of patients with colorectal cancer, a disease that ranks second in cancer-related mortality worldwide. While surgical resection of colorectal liver metastases (CRLM) remains the only path to a potential cure, recurrences occur in more than 70% of the patients post-hepatectomy, with most relapses occurring within two years of surgery. Currently, there is a lack of accurate prognostic tools to inform patients of their risk. In this work, we propose a multi-omics framework to predict the two-year recurrence likelihood after CRLM resection. Our pipeline integrates computed tomography (CT) images, RNA sequencing and the Clinical Risk Score (CRS), shown to be associated with patient

prognosis, and commonly used by oncologists. Furthermore, we address the challenge of undetected RNA gene transcripts by including a generative adversarial imputation mechanism. We also leverage generative learning and Transformers to circumvent the high dimensionality of gene expression datasets. The imaging component incorporates a foundation model as a feature extraction module which complements the interpretable radiomics features. To predict two-year recurrence risk on a prospectively maintained dataset of 129 CRLM patients, our multi-omics pipeline achieved an AUC of 0.75, outperforming its single and dual-modal counterparts, and improved over the CRS (AUC = 7.14%; $p = 0.01$). It also provided significant risk stratification and clinical utility. Our proposed pipeline could serve as a tool to better inform patients on their expected outcome, guide the frequency of follow-up to treat early recurrence and inform the design of trials testing novel post-operative treatments.

8.1.1 Keywords

Multimodal prognostication, colorectal liver metastases, generative learning, foundation models, RNA sequencing, computed tomography

8.2 Introduction

Colorectal cancer (CRC) is the second most lethal cancer in the United States [178]. Approximately half of patients with CRC develop liver metastases [274], the most prevalent progression site. Although complete surgical removal of colorectal liver metastases (CRLM) alongside systemic chemotherapy can potentially cure the disease [275], nearly 70% of patients experience a recurrence [223]. Most CRLM recurrences occur during the first two years following curative resection [179]. Thus, prognostic biomarkers are key to identify high risk patients that would typically benefit from adjuvant chemotherapy and more frequent follow-ups after surgery to detect and treat recurrence early. Such biomarkers could also help develop new adjuvant treatments for patients at high risk of recurrence. On the other hand, patients at low risk of recurrence may be spared the burden of intense follow-up and adjuvant chemotherapy.

Several prognostic scores have been proposed in the literature to stratify CRLM patients based on their survival outcomes such as the Clinical Risk Score (CRS), proposed by Fong et al. [20], assesses the pathological response to pre-operative chemotherapy and the TNM stage [18,19]. The most established presurgical metric is the Fong CRS, a pre-operative score ranging from 0 to 5, which allows to stratify patients according to their recurrence risk. Nevertheless, despite being widely accepted by oncologists, the Fong score remains a generic

metric that fails to provide comprehensive insights into individual patient prognosis, which is heavily influenced by tumor biology.

Recent efforts have been made to assess the impact of genomic tumor mutation on CRLM recurrence and overall survival. For instance, the incorporation of KRAS mutations has been shown to improve survival prediction in patients with resected CRLM [276]. Furthermore, liver-specific transcription factors FOXA2 and HNF1A were associated with metastatic progression of CRC [277]. RNA sequencing has been recently assessed for prognostication purposes [278]. Genes with significant expression level alterations have been used to predict the survival of stage IV CRC patients with liver metastases [279]. Nevertheless, missing gene expression counts can occur due to low expression levels and the inherent stochasticity of RNA sequencing. Discarding these missing counts could negatively impact downstream analyses.

On the other hand, recent findings suggest that imaging characteristics may provide prognostic insights into CRLM [280]. Machine learning models have been trained using imaging features to predict recurrence-free survival, overall survival, as well as the response to chemotherapy [281–283]. Different imaging modalities were leveraged in these studies, including computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography and histopathological images. Particularly, CT images have the unique advantage of being readily accessible and routinely used in clinical practice, both during initial diagnosis and follow-up evaluations of treatment response. Radiomics features, which represent measurable quantifications of shape, texture and intensity of CT tumors, have been shown to exhibit significant prognostic capabilities [284]. Moreover, recently emerging foundation models, trained on extensive datasets, also show considerable potential for specialized tasks as predicting clinical outcomes. [285].

Despite recent progress in machine learning tools for the prognostication of CRLM, most models rely on individual modalities and fail to provide a holistic evaluation of the disease by considering all available modalities. While some deep learning approaches for cancer prognosis combine different data types [286], there is yet a multimodal pipeline for CRLM. When it comes to non-metastatic CRC, Ho et al. [287] combined tabular clinical information with time-series laboratory test results for recurrence prediction. Chaudhary et al. [288] merged RNA sequencing, miRNA sequencing and DNA methylation of primary liver cancer in an integrative autoencoder-based model which separates patients subgroups with statistically significant survival differences. These studies confirm the benefit of developing an integrative prognostic tool incorporating multiple medical data types.

In this work, we propose a novel multi-omics pipeline for predicting the two-year CRLM recurrence risk after surgical resection. The pipeline aims to disentangle intricate biological

processes involved in cancer progression using different modalities, namely diagnostic CT images taken pre-operatively, RNA sequencing data as well as highly prognostic clinical features. Generative deep learning techniques are leveraged in the framework to address the challenge of high dimensionality in RNA sequencing data. We also harness advances in versatile vision foundation models to derive advanced imaging features, probing metastatic progression.

The main contributions of this study can be summarized as follows :

1. The proposed multi-omics predictive pipeline uses multimodal data sources to assess the likelihood of early CRLM recurrence (two years post-hepatectomy) leveraging advances in foundation models [7], Transformers [151] and generative learning [159, 177]. To the best of our knowledge, this is the first work that presents a composite framework incorporating different feature types to predict recurrence in the context of liver metastases of colorectal origin.
2. Our proposed pipeline incorporates a RNA sequencing-based module, which addresses the challenge in manipulating transcriptomics datasets, namely undetected sample transcripts (technically known as dropout) and the high dimensionality of gene expression datasets.
3. We introduce a CT-based recurrence prediction module which employs cascaded models trained on complementary foundation model-derived representations and interpretable radiomics features, hence yielding a rich and versatile characterization of the tumoral microenvironment.
4. Our study provides interpretability of the model’s outputs through analyses associating genes with CRLM recurrence. Moreover, our multi-omics pipeline is confronted with the Fong CRS, a clinically established prognostic measure, on a prospectively maintained dataset of CRLM patients.

8.3 Background

8.3.1 CRLM outcome prediction from CT

Recent years have witnessed increased efforts in the application of machine learning techniques for the prognostication of CRLM using CT images, thereby complementing conventional staging techniques [78, 289]. Lubner et al. [282] found that radiomics features were associated with survival outcomes in CRLM patients. Ye et al. [283] computed delta-radiomics features using CT images taken before and after neoadjuvant therapy to predict progression-free survival. Granata et al. [290] used radiomics features derived from CT and MRI to

predict prognosis-linked patterns and recurrence. Luo et al. [291] trained a random survival forest using preselected radiomics features and found a statistically significant separation between the high and low risk groups on both the training and testing subsets. Montagnon et al. [281] predicted the time to recurrence and disease-specific survival using survival models trained on radiomics features alone or in combination with the Fong CRS. Keyl et al. [292] leveraged a pretrained nnUnet model [293] to automatically segment CRLM and used the ensuing segmentations to compute the surface area of the metastases. The latter was then combined with the primary lesion sidedness and muscle-to-bone ratio in a multivariate Cox proportional hazards model for the prediction of overall survival. Wang et al. [294] performed unsupervised clustering using radiomics features extracted from segmented CRLM and found a cluster with worse clinical outcomes, namely shorter disease free and overall survivals. This demonstrated that the radiomics-based model improved the prognostic capabilities to the CRS and the tumor burden score. However, these studies mainly rely on predefined radiomics features alone, which, while highly informative, may be insufficient to fully capture the complexity of tumor biology.

8.3.2 CRC outcome prediction from RNA sequencing

While transcriptomics-based prognostic pipelines are limited in the case of CRLM, several studies have been conducted on the primary CRC. Yang et al. [295] tested several machine learning models separately on mRNA, miRNA and microbiome of CRC tissue to classify patients into short versus long-term survivors. Lu et al. [296] trained various machine learning models to distinguish responders from non-responders to FOLFOX therapy using differentially expressed genes among the two groups. Similarly, Tian et al. [297] and Amniouel et al. [298] predicted FOLFOX response using gene expression data of CRC patients. Kim et al. [299] reported that an ElasticNet-based model trained on transcriptomes of normal tissue surrounding CRC outperformed the one trained on intratumoral transcriptomes, highlighting the potential prognostic value of the peritumoral tissue. When it comes to CRC liver metastases, Zheng et al. [300] trained four machine learning models to identify CRLM-specific genes and evaluated the prognostic value of the resulting genes using the Lasso-Cox technique. Although a strong association between the human transcriptome and the prognosis of CRC patients has been demonstrated, most of these studies focus primarily on primary lesions and do not address common challenges in gene expression data analysis, such as dropout and high dimensionality.

8.3.3 CRLM outcome prediction from clinical models

As for RNA and CT-based methods, several studies have been conducted to predict CRLM oncological outcomes using predefined clinical features. Nga Lam et al. [301] collected 65 features including demographics, pre- and post-hepatectomy, to train a Cox proportional hazards model predicting the overall and recurrence-free survivals. Paredes et al. [302] predicted CRLM recurrence using clinical and morphological data and compared the performance of the model to the Fong CRS. Zhai et al. [303] evaluated the value of the tumor size ratio, defined as the ratio of the respective diameters of the largest and smallest lesions, in predicting the recurrence-free survival of CRLM. Likewise, Moaven et al. [304] trained a logistic regression and gradient boosted trees using 128 features including demographic, lifestyle and clinico-pathological data to predict recurrence and overall survival in CRLM. Nevertheless, these studies use predefined clinical characteristics and fail to provide a holistic assessment of CRLM.

8.4 Materials and methods

The overall workflow of the proposed method is illustrated in Figure 8.1. RNA sequencing data was first processed separately using the following steps : (1) missing gene expression counts were generatively imputed using the Generative Adversarial Imputation Network (GAIN), (2) the complete gene expression matrix was augmented through conditional generative learning, (3) the original and synthetic RNA sequencing instances were combined and used to train a Feature Tokenizer Transformer (FT-Transformer) for the prediction of 2-year recurrence. In parallel, we extracted radiomics features and versatile deep characteristics obtained by leveraging a foundation model as a feature extractor from preprocessed CT images and merged the two feature types using cascaded models. Finally, the multimodal pipeline integrated RNA sequencing, CT and clinical features to predict the early recurrence of CRLM.

8.4.1 Patient cohort

We conducted a prospective study of surgically resected CRLM patients at the Centre Hospitalier de l'Université de Montréal (CHUM). Consenting patients were treated between 2010 and 2022 and a prospective follow-up was performed beyond recurrence and until death or 2023, whichever occurred first. Patients were included if : (1) CRLM were histopathologically confirmed, (2) CRLM were completely resected with curative intent, (3) RNA sequencing was performed on the resected CRLM. Based on these criteria, 129 patients were included

for further analysis. The average age of patients was 64.1 years (ranging from 38 to 83), with a majority of men (65.9%). Upon data collection, 65.1% of patients experienced recurrence within the first two years following surgery. The median follow-up duration was 45.5 months, during which 56.6% succumbed to the illness by the end of the follow-up period. Furthermore, 64.7% of patients had a Fong score between 0 and 2, while the remaining had a score between 3 and 5. Neoadjuvant, pre-operative chemotherapy was given to 84.5% of the patients. Additionally, 62.0% of patients received chemotherapy following resective surgery. The study was approved by the CHUM institutional review board (No. 16.262).

8.4.2 Data preparation

Imaging data preprocessing

In this study, we included the baseline contrast-enhanced CT images acquired before hepatectomy. Therefore, for patients who underwent neoadjuvant chemotherapy, the pre-chemotherapy image was selected. The images had an in-plane size of 512×512 , an average transaxial resolution of 0.75 mm^2 (range = $[0.55, 0.98] \text{ mm}^2$) and an axial slice thickness of 2.91 mm (range = $[0.80, 6.5] \text{ mm}$). All images were resampled to 1 mm isotropic pixel spacing to remove the discrepancies in the axial and in-plane resolutions across images. For images acquired before 2018, CRLM were automatically segmented using a joint liver/hepatic segmentation model consisting of two convolutional networks [305]. For the remaining CT scans obtained in or after 2018, the MedSAM Slicer plugin [171] was used to segment the lesions. In recent studies, MedSAM was shown to outperform the SAM and U-Net models in liver tumor segmentation with a dice score greater than 89% [171]. The resulting segmentations were validated by an experienced radiologist.

RNA sequencing data

Sequencing work was carried out on resected CRLM in three different genomics core facilities (Genome Quebec, Montreal, Canada; McGill Genomic center, Montreal, Canada and Turnstone, San Diego, United States). Reads were aligned to the human genome and expression calls were derived in each sample. The preprocessing of the three gene expression datasets was performed following three stages : (1) data filtering during which common genes were retained and genes with less than ten counts were filtered out; (2) batch correction was performed using the Combat-seq method [32]; and (3) the median of ratios technique was applied to normalized expression counts using R package DeSeq2 and data were transformed using the variance stabilizing transformation technique. The resulting dataset consisted of

normalized counts of 28,578 genes among which 12,806 were randomly undetected across samples.

Clinical features

Clinical data was collected in this study included demographics, oncological outcomes — namely the recurrence-free and disease-specific survival, the types and dates of administered therapies, and the Fong CRS. The latter was computed by adding a point whenever one of the following five conditions is satisfied : (1) the disease-free interval between the diagnosis of the primary tumor and liver metastases was less than a year ; (2) the number of metastases was greater than 1 ; (3) the pre-operative carcinoembryonic antigen (CEA) level was greater than 200 ng/mL ; (4) the largest metastasis was greater than 5 cm ; (5) the primary tumor was lymph node positive [20]. The Fong CRS, gender and age were subsequently chosen to be included in the pipeline as clinical features. Gender was one-hot encoded and age and the Fong score were Z-standardized before being used to train the model.

8.4.3 RNA sequencing module

Imputation of missing expression counts

The gene expression dataset was initially preprocessed by applying the Generative Adversarial Imputation Network (GAIN) [8] to impute missing expression counts, as shown in [177]. The architecture of the GAIN module is depicted in Figure 8.2.

The first major component is the generator which is trained to synthesize imputations of missing gene expression counts. Using a partial observation of a random variable $\tilde{\mathbf{X}}$, a binary mask $\mathbf{M}$, and a noise $\mathbf{Z}$, the generator G is trained to output the imputation matrix $\bar{\mathbf{X}}$ [8] :

$$\bar{\mathbf{X}} = G(\tilde{\mathbf{X}}, \mathbf{M}, (1 - \mathbf{M}) \cdot \mathbf{Z}). \quad (8.1)$$

The binary mask is of the same dimension as $\tilde{\mathbf{X}}$ and allows the generator to differentiate between real (1) and missing (0) counts. The completed data $\hat{\mathbf{X}}$ is composed of the observed instances and the imputations synthesized by the generator :

$$\hat{\mathbf{X}} = \mathbf{M} \cdot \tilde{\mathbf{X}} + (1 - \mathbf{M}) \cdot \bar{\mathbf{X}}. \quad (8.2)$$

The second main component of the GAIN module is the discriminator trained to predict $\mathbf{M}$, which is equivalent to predicting which instances are observed and which are generated. In

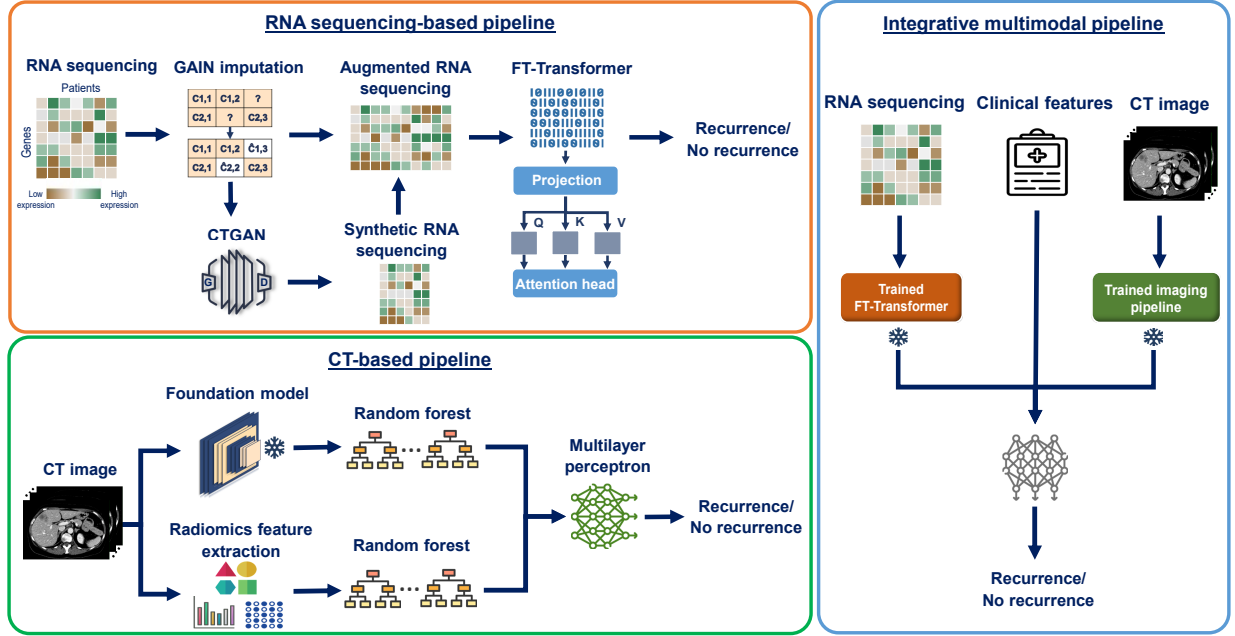


FIGURE 8.1 Overall workflow of the proposed approach which combines transcriptomics, imaging and clinical data to predict the early recurrence of resected CRLM (two years after complete resection). RNA sequencing and CT images were first processed separately and the ensuing transcriptomics and imaging scores were leveraged along with clinical features (Fong CRS, age and gender) to predict the likelihood of disease recurrence. The RNA sequencing module comprised a Generative Adversarial Imputation Network (GAIN) which performed generative imputation of missing gene expression counts, followed by a Conditional Tabular Generative Adversarial Network (CTGAN) which synthesized RNA sequencing instances that were combined with the original samples and used to train a Feature Tokenizer Transformer (FT-Transformer). The CT component incorporated a foundation model leveraged as a feature extractor, along with a radiomics feature extraction module. The resulting features were used to train cascaded random forest models. The scores of the respective models were combined in a multilayer perceptron to predict the 2-year recurrence.

addition to the complete matrix $\hat{\mathbf{X}}$, the discriminator is provided with partial information on the mask via a hint matrix $\mathbf{H}$. The latter reveals a limited number of the actual mask values. This is indispensable in the GAIN architecture since without sufficient knowledge on $\mathbf{M}$, multiple distributions generated by the generator could be optimal to the discriminator. As in the vanilla generative adversarial network (GAN), an iterative training is followed in GAIN using a minmax scheme :

$$\min_G \max_D \mathbb{E}[\mathcal{L}(\mathbf{M}, \widehat{\mathbf{M}})] \quad (8.3)$$

where G is the generator, D the discriminator, $\mathcal{L}$ the cross-entropy loss function, and $\widehat{\mathbf{M}}$ the

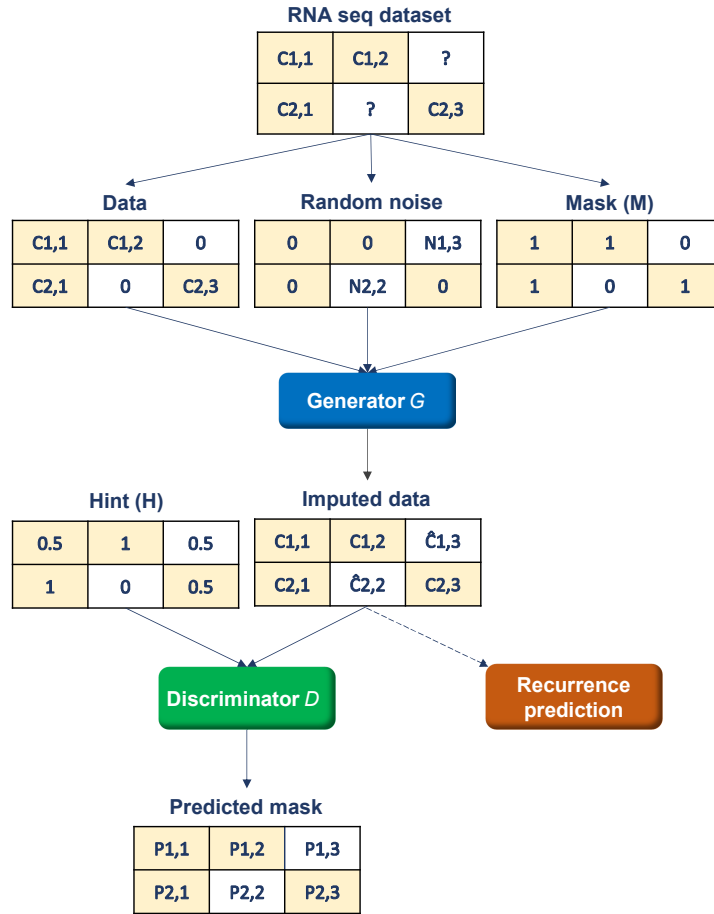


FIGURE 8.2 Generative imputation of missing gene expression counts. The GAIN model [8] was leveraged to impute missing gene expressions. A generator synthesized imputations of missing gene expression counts and a discriminator was trained to predict whether expression counts were real or synthetic imputations proposed by the generator. The imputed gene expression matrix was subsequently deployed for recurrence prediction.

mask predicted by the discriminator.

The generator and discriminator were fully connected neural networks with 3 layers each, of respective dimensions $N \times 2$, N and N , where N is the number of features. Due to memory constraints, the dataset features were divided into 3 subsets and generative data imputation was performed on each subset separately.

Synthetic instance generation

Given the large number of genes in the RNA sequencing dataset and the relatively limited number of instances, the transcriptomics dataset is intrinsically high-dimensional with a relatively low number of instances and a large number of genes. Therefore, we synthesized

artificial gene expression instances using the Conditional Tabular Generative Adversarial Network (CTGAN) [159] to help alleviate issues with data high-dimensionality. CTGAN is a variant of the vanilla GAN which generates synthetic tabular data by modeling the probability distribution of its rows. CTGAN tackles the issue of class imbalance in the discrete columns (the recurrence status in this case) by effectively resampling, during the training process, in a manner that allows all the classes of the discrete columns to be sampled evenly. This process is denoted as 'training-by-sampling' and is described in Algorithm 1. CTGAN leverages a conditional generator which aims to learn the real conditional distribution of the data. In other words, given a value v^* of the h^{*th} discrete column D_{h^*} , the generated instances $\hat{\mathbf{r}}$ follow a conditional distribution given the particular value v^* of D_{h^*} :

$$\hat{\mathbf{r}} \sim \mathbb{P}_G(row | D_{h^*} = v^*). \quad (8.4)$$

The generator and the discriminator were composed both of a fully-connected network with two hidden layers. Batch normalization and a ReLU activation function were used in every hidden layer of the generator while a leaky ReLU and 50% dropout were used in the discriminator. All hidden layers had 256 nodes.

Algorithm 1 CTGAN training-by-sampling approach

Inputs :

N_d : number of discrete columns

$\mathbf{r}_n$: row representation with discrete and continuous columns

$z \sim \mathcal{N}(0, 1)$

1. for $i \in 1, \dots, N_d$ **do**

 Create mask $\mathbf{m}_i \in \mathbb{R}^{D_i}$ $D_i \triangleleft$ number of categories in column i

 Fill $\mathbf{m}_i$ with zeroes

end for
2. Randomly select a discrete column D_h

$h^* \leftarrow$ index of D_h

3. Randomly select value v^* using probability mass function across values of D_h
4. $\mathbf{m}_{h^*}^{(v^*)} \leftarrow 1$
5. $cond \leftarrow \text{concat}[\mathbf{m}_1, \dots, \mathbf{m}_{N_d}]$
6. $\hat{\mathbf{r}} \leftarrow \text{Generator}(cond, z)$
7. Score $\leftarrow \text{Discriminator}(\mathbf{r}_1, \dots, \mathbf{r}_n, \dots, cond)$

FT-Transformer model

The synthetic RNA sequencing instances were combined with the original samples and used to train a FT-Transformer model for the prediction of the 2-year recurrence. The FT-Transformer incorporates a feature tokenizing step, allowing to convert features into rich task-specific representations, followed by a Transformer encoder and a classifier. The overall architecture of the model is shown in Figure 8.3.

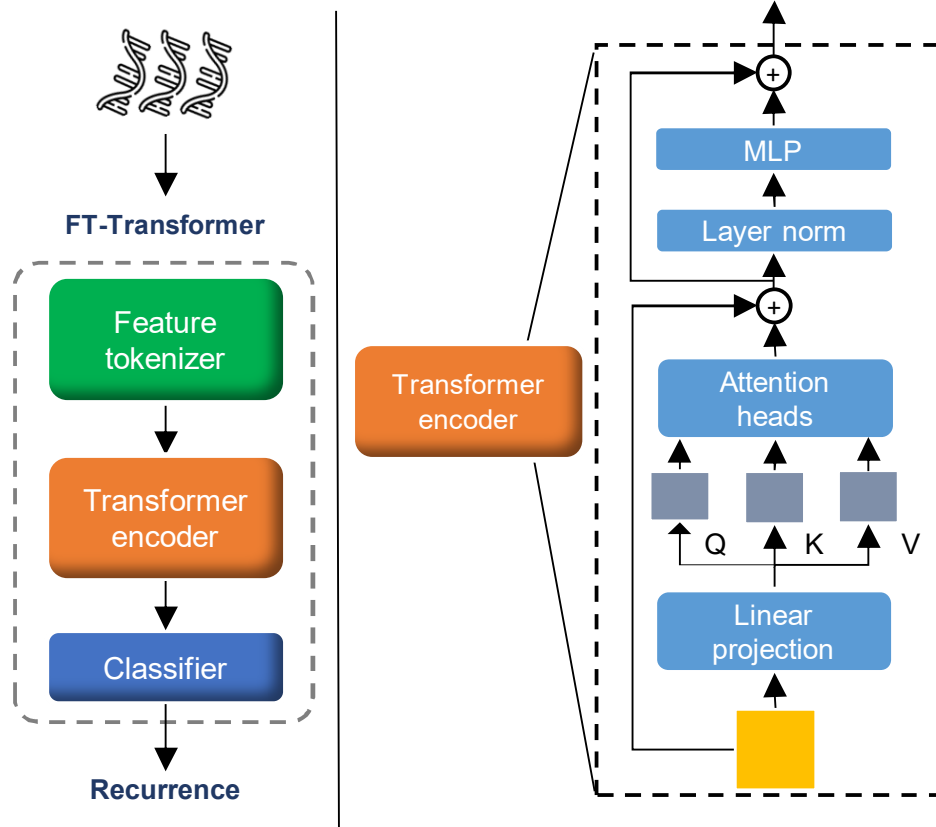


FIGURE 8.3 Architecture of the FT-Transformer model trained on the synthetically augmented gene expression dataset for the prediction of 2-year recurrence. Features are first embedded using the feature tokenizer module and the embeddings are fed into the Transformer encoder including self-attention and fully-connected layers.

Features are first converted into embeddings using the feature tokenizer module. For a feature x_i , the embedding H_i is computed as follows :

$$H_i = W_i \cdot x_i + b_i \in \mathbb{R}^k \quad (8.5)$$

where W_i is the weight vector, b_i is the bias and k is the embedding size. Once all features are tokenized, stacking of the feature representations is performed and the embedding of the

output classification token CLS is appended to the tokens :

$$H = \text{concat}[H_1, \dots, H_d] \in \mathbb{R}^{d \times k} \quad (8.6)$$

$$H_0 = \text{concat}[CLS, H] \in \mathbb{R}^{(d+1) \times k} \quad (8.7)$$

where d is the number of features, H is the set of feature embeddings and H_0 is the final set of embeddings.

The ensuing set of representations H_0 is fed into a Transformer layer composed of a self-attention mechanism and a feed-forward network. Finally, the embedding of the classification token is sent to a fully connected layer for the prediction of recurrence.

8.4.4 Imaging module

In addition to the RNA data, we derived two feature types from pre-operative CT images : (1) radiomics features and (2) deep representations obtained using a pretrained foundation model leveraged as a feature extractor. The rationale for combining these two feature types lies in the complementary strengths they offer. Radiomics features are well-established and interpretable, providing robust and reproducible quantitative information. In contrast, the foundation model can generate versatile features that capture additional information beyond the scope of predefined radiomics unseen to the human eye, thereby enriching the overall feature set with previously unexplored insights.

Radiomics workflow

From each segmented CRLM lesion, 1223 radiomics features were extracted using the Py-Radiomics v3.0.1 toolbox [190]. Feature extraction was customized by applying a Wavelet or a Laplacian of Gaussian (LoG) filter. In the first case, 744 Wavelet features were extracted after deriving eight Wavelet decompositions resulting from the application of a high or a low pass filter on the original images in each of the three dimensions. In the second case, 465 LoG features resulted from the application of a LoG with several sigma values (0.1, 0.5, 1, 2, 3). The ensuing features consisted of a total of 14 shape, 234 first-order, and 975 textural features. All extracted features were standardized to obtain a zero mean and unit standard deviation. The least absolute shrinkage and selection operator (LASSO) [191] was then applied on the standardized set to select the most salient features that were subsequently used for model training. A random forest model was then trained using the selected features to

predict the likelihood of recurrence two years after surgery.

Foundation model workflow

Additionally, we leveraged the foundation model introduced by [7] to capture the underlying intricate features in CRLM. The proposed model was trained using approximately 11.5K three-dimensional CT images and validated as both a diagnostic and prognostic biomarker. An unsupervised contrastive training approach was adopted based on an adjusted SimCLR framework [163]. In the proposed approach, transformations were applied to each 3D bounding box and a 3D ResNet50 [306] model was used to derive latent representations of each modified version. During training, a contrastive loss function [307] guided the model to generate comparable representations for instances corresponding to the same image and disparate ones for instances emanating from different images. External validation of the pretrained model as a prognostic biomarker was also performed on the LUNG1 and RADIO [308] public datasets. Results showed that using the model as a feature extractor led to better results compared to transfer learning or fine-tuning its weights on the downstream task [7]. In this work, we used the pretrained foundation model as a feature extractor for our recurrence prediction task. Using the center point coordinates of the previously generated segmentations, bounding boxes of size $50 \times 50 \times 50$ voxels were cropped around each CRLM. A min-max normalization was then performed and the ensuing regions of interest were fed into the foundation model resulting in a total of 4096 features extracted using the last layer of the ResNet50's convolutional encoder. The foundation model-derived features were subsequently used to train a separate random forest for the prediction of recurrence.

Integrative CT-based pipeline

In a subsequent step, a multi-layer perceptron (MLP) with two hidden layers (features : 12, 8, activation function : ReLU) was trained using the class probabilities of the two random forest models trained on the radiomics and foundation model-derived features respectively. This cascading technique allowed boosting the performance of the resulting classification pipeline as it incorporated different models working in tandem on complementary features extracted in a hierarchical fashion.

8.4.5 Multi-omics feature fusion

Finally, the imaging and transcriptomics score derived from the CT and RNA sequencing components respectively, were combined with the clinical features (Fong CRS, age and gen-

der) and fed to a final MLP. The latter was composed of two hidden layers of dimensions 12 and 8, respectively and ReLU activation functions. A sigmoid activation function was applied to the output layer.

8.4.6 Statistical and interpretability analyses

The Student's t test with Nadeau and Bengio correction [309] was conducted to evaluate whether a statistically significant difference in performance was present between the different trained models. Python dcurves v1.0.6.4 was used to plot the decision curve analysis [310]. For each threshold probability representing the acceptable risk beyond which further testing is advised, the net benefit was computed as follows :

$$Net\ benefit = \frac{TP}{N} - \frac{FP}{N} \left(\frac{p_t}{1 - p_t} \right) \quad (8.8)$$

where TP is the true positive count, TN the true negative count, N the total number of patients and p_t the probability threshold.

Kaplan-Meier curves were generated using the Python Lifelines v0.27.8 package and compared using the long-rank test. The X-tile software was used to stratify patients into high versus low risk groups by applying a cutoff to the probabilistic outputs of the trained models. The time to recurrence was derived for each patient from the date of the first CRLM resection. Statistical tests were performed using Python Scipy v1.10 and a threshold of 0.05 was chosen for statistical significance of p -values.

To identify the most important genes used by the model for predictions, we performed differential gene expression analysis on the two groups predicted by the multi-omics model in cross-validation (recurrent vs. non-recurrent) using pydeseq2 package v.0.4.0. Genes were considered differentially expressed if : (1) the adjusted p -value was less than 0.05, (2) the $\log_2$ fold change was greater than 1 and (3) the base mean was greater than 10. Moreover, we performed the Shapley Additive Explanations (SHAP) analysis to derive the top contributing imaging and clinical features [261].

8.4.7 Training and implementation details

Models were trained using a 64 GB RAM NVIDIA Titan RTX GPU. Evaluations were performed using a five-fold cross-validation scheme, with the same folds created for all models to ensure an equitable comparison. Pre-operative CT images were normalized in Hounsfield units across different scanners and vendors. The FT-Transformer was trained using a token

size of 8, a compression factor of 0.05, a single attention head and for 20 epochs. Training of all deep learning models was performed using Pytorch v2.0.1.

8.5 Results

8.5.1 Comparison across modalities

The classification performance of the proposed multi-omics pipeline was evaluated across the different combinations of modalities obtained using one, two or three modalities (Table 8.1, Figure 8.4). The results show that training the model using the synthetically augmented RNA sequencing data leads to an increase in the model’s area under the receiver operating characteristic curve (AUC) from 0.62 (95% CI : 0.54–0.71) (without synthetic augmentation) to 0.68 (95% CI : 0.60–0.76) (with synthetic augmentation). When trained using features derived from a single modality, the highest AUC was observed using the clinical features (0.70, 95% CI : 0.62–0.78). This correlates with the strong prognostic value of the Fong CRS. Models trained using combinations of two modalities consistently outperformed those trained using a single modality and reached an AUC greater than 0.70 : 0.71 (95% CI : 0.63–0.79), 0.71 (95% CI : 0.63–0.79) and 0.72 (95% CI : 0.64–0.79) with RNA sequencing + imaging, imaging + clinical and RNA sequencing + clinical features, respectively). The best performance was observed using the combination of RNA sequencing, clinical and imaging features with an AUC of 0.75 (95% CI : 0.68–0.82), an accuracy of 0.70, a precision of 0.72, a recall of 0.88 and an F1-score of 0.79.

Subsequently, we evaluated whether the improvement observed using the multi-omics pipeline over the Fong CRS, currently used in routine clinical practice, is statistically significant. The statistical analysis results show that a significant increase of 7.14% in terms of AUC was observed between the proposed multi-omics pipeline and the model trained on clinical features only ($p = 0.01$, corrected student t test) (Figure 8.5). The performance difference obtained with the other models, including the ones trained using a combination of two modalities, did not reach statistical significance ($p > 0.05$). This observation confirms the importance of integrating all three features types, providing complementary information and leading to an improved recurrence prediction performance compared to the Fong CRS alone.

8.5.2 Performance of the imaging component

Figure 8.6 shows the classification performance of the CT-based module of the multi-omics pipeline. Using radiomics and foundation model-derived features individually, an accuracy of 0.64 and 0.68 and an AUC of 0.61 (95% CI : 0.52–0.70) and 0.61 (95% CI : 0.52–0.70)

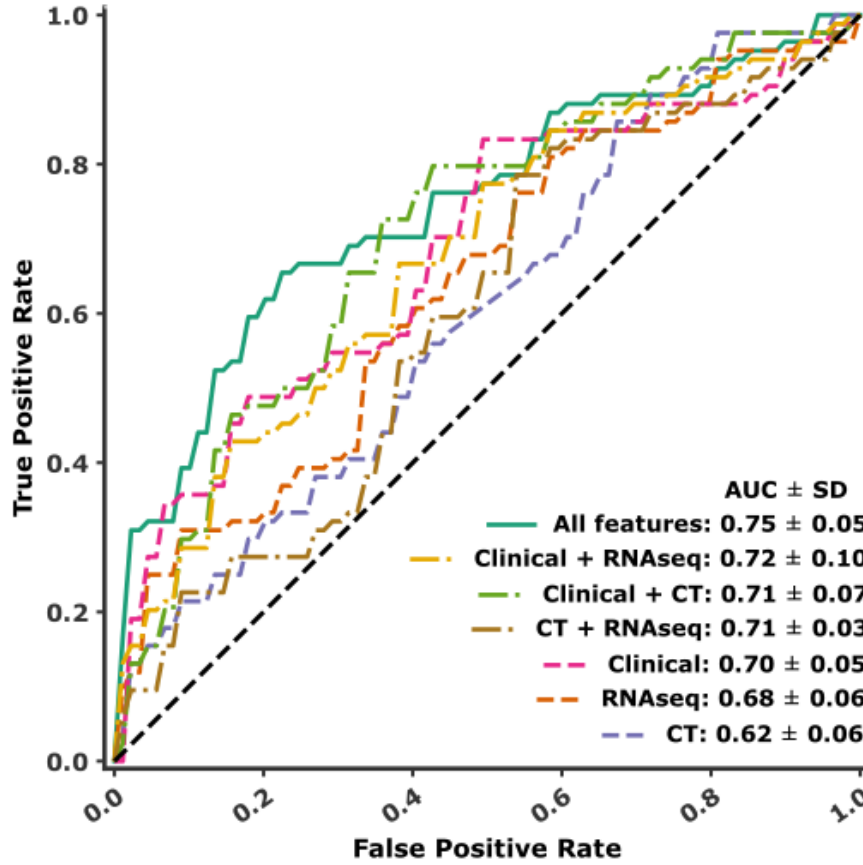


FIGURE 8.4 Receiver operating characteristic (ROC) curves by modality using the proposed deep learning pipeline for the prediction of 2-year recurrence in colorectal liver metastasis surgery cases ($n = 129$, 84 events). Figures show the AUC with the corresponding standard deviation across cross-validation fold ($k = 5$). The diagonal dashed line represents a performance equivalent to random chance.

were obtained, respectively. The integrative approach combining radiomics and foundation model-derived features exhibits a higher classification performance with an accuracy of 0.70 and an AUC of 0.62 (95% CI : 0.53–0.71). This highlights the complementary nature of the radiomics features and the features extracted using the foundation model.

8.5.3 Prognostic value of the proposed pipeline

Figure 8.8 illustrates the Kaplan-Meier curves obtained from the models trained using the different combinations of modalities. The multi-omics pipeline combining the CT, RNA sequencing and clinical features was able to yield the best separation of the disease-free survival curves, identifying the highest proportion of patients without recurrence. In high risk patients, 80% of recurrences were observed at two years of follow-up, compared to 40% in patients at

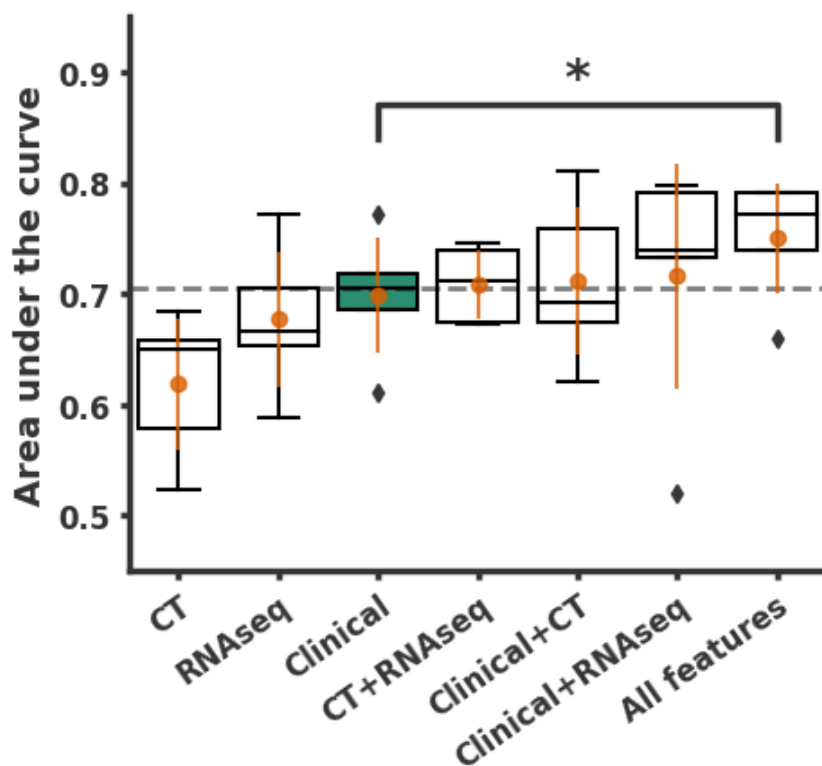


FIGURE 8.5 Comparison of the areas under the receiver operating characteristic curves by modality against the baseline clinical-only model. Box plots show the Q1, median and Q3, with box whisker's representing the 1.5 interquartile range. Statistically significant differences ($p < 0.05$) are highlighted with an asterisk. The dashed horizontal line depicts the median AUC of the clinical-only model. The brown dots indicate the mean AUC and the brown errors bars the standard deviations.

low risk of recurrence. The risk stratification was also statistically significant with the models trained using two modalities. The p -value obtained when the three modalities were included in the pipeline was 0.0008 (log-rank test).

8.5.4 Clinical utility of the proposed pipeline

As presented by the decision curve analysis (Figure 8.9), the net benefit of the proposed multi-omics pipeline is higher than the “treat all patients as high risk” (blue curve) and “treat no patient as high risk” (black line) approaches for risk thresholds greater than 0.45. For most probability thresholds, the net benefit of the multi-omics pipeline is higher than that of the other models, which shows that the clinical utility of our proposed pipeline is more significant than the models trained using one or two modalities. In other words, the proposed approach is more effective at correctly identifying patients who truly benefit from

TABLEAU 8.1 Ablation study results. The mean and standard deviation of the performance metrics in cross-validation are reported in the table. The performance observed with the proposed multi-omics approach was the only one that exhibited a statistically significant difference with the model trained using clinical features (marked with an asterisk). SA ; synthetic augmentation.

Features				AUC	Accuracy	Precision	Recall	F1-score
RNAseq w/o SA	Clinical	RNAseq	CT					
	✓			0.70 (0.05)	0.64 (0.04)	0.67 (0.02)	0.88 (0.16)	0.75 (0.07)
✓				0.62 (0.09)	0.65 (0.05)	0.70 (0.06)	0.83 (0.09)	0.76 (0.03)
		✓		0.68 (0.06)	0.67 (0.06)	0.73 (0.07)	0.81 (0.07)	0.76 (0.03)
			✓	0.62 (0.06)	0.70 (0.02)	0.70 (0.02)	0.95 (0.05)	0.81 (0.02)
		✓	✓	0.71 (0.03)	0.67 (0.07)	0.72 (0.08)	0.83 (0.07)	0.77 (0.04)
	✓		✓	0.71 (0.07)	0.68 (0.09)	0.70 (0.04)	0.88 (0.16)	0.77 (0.09)
	✓	✓		0.72 (0.10)	0.68 (0.09)	0.71 (0.07)	0.87 (0.09)	0.78 (0.06)
	✓	✓	✓	*0.75 (0.05)	0.70 (0.07)	0.72 (0.06)	0.88 (0.08)	0.79 (0.05)

treatment (true positives) while avoiding unnecessary treatments for those who don't (false positives), thus improving patient outcomes and resource use.

8.5.5 Comparison with state-of-the-art tabular learning models

The performance of the proposed RNA sequencing-based pipeline was compared to other tabular learning models including : (1) machine learning models (decision tree, random forest and support vector machine (SVM)), (2) gradient boosting algorithms (the Categorical Boosting (CatBoost), the Adaptive Boosting (AdaBoost), the Extreme Gradient Boosting (XGBoost)), and (3) deep learning models (the multilayer perceptron (MLP), the Residual Network (ResNet) [306], the Attentive Interpretable Tabular Learning (TabNet) [149], the Tabular Transformer (TabTransformer) [150] and the Tabular Prior-Data Fitted Network (TabPFN) [311]). The different models, as well as their respective performance metrics, are listed in Table 8.2. The FT-Transformer model consistently outperformed the other tabular learning models in terms of accuracy and AUC. Moreover, we compared the performance of different machine learning models trained on radiomics and foundation model-derived features extracted from CT images. The experimental results, presented in Figure 8.7, show that the random forest outperformed the other machine learning models in both cases.

8.5.6 Comparison of different fusion approaches

Furthermore, we compared the performance of the intermediate fusion approach adopted in the proposed pipeline to two other approaches : early and late fusion approaches. In the former approach, the RNA sequencing gene expression features, clinical features and ima-

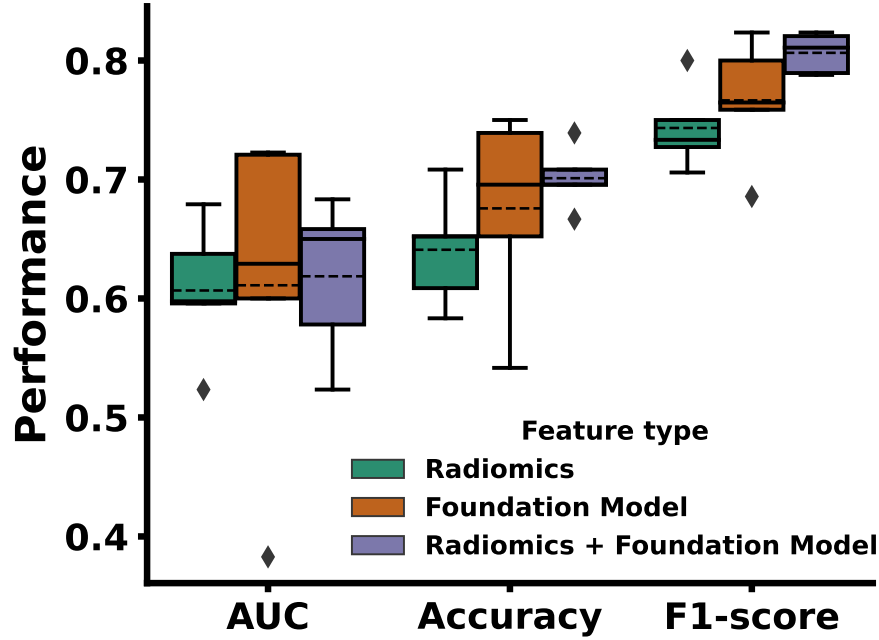


FIGURE 8.6 Ablation study results of the imaging module. The boxes cover the interquartile range, the whiskers the 1.5 interquartile range and the dots indicate the outliers. The dashed and straight lines represent the mean and median performance metrics, respectively.

ging features (both radiomics and foundation model-derived features) were concatenated to form a feature set of 33,900 features. The CTGAN model was applied to generate synthetic instances and the FT-Transformer was subsequently trained and evaluated using the same strategy adopted for the proposed pipeline. In the latter approach, the different modalities were separately processed using the following steps : (1) RNA sequencing was imputed using GAIN, augmented using the CTGAN for synthetic instance generation and the FT-Transformer was trained. (2) Radiomics and foundation-model features were used to train random forest models and the radiomics and FM scores were combined using an MLP. (3) The clinical features (Fong CRS, age and gender) were used to train a fully connected network. (4) Soft ensembling of the three models was performed by combining their respective scores and using the concept of majority voting. The results presented in Table 8.3 show that the proposed approach leads to the best performance.

8.5.7 Interpretability analysis

The top differentially expressed genes between the high and low risk groups, along with the volcano plot, are shown in Figure 8.10. Interestingly, amongst the top genes associated with low risk of recurrence, were the two pseudogenes RPL10P6 and RPL10P9 of the riboso-

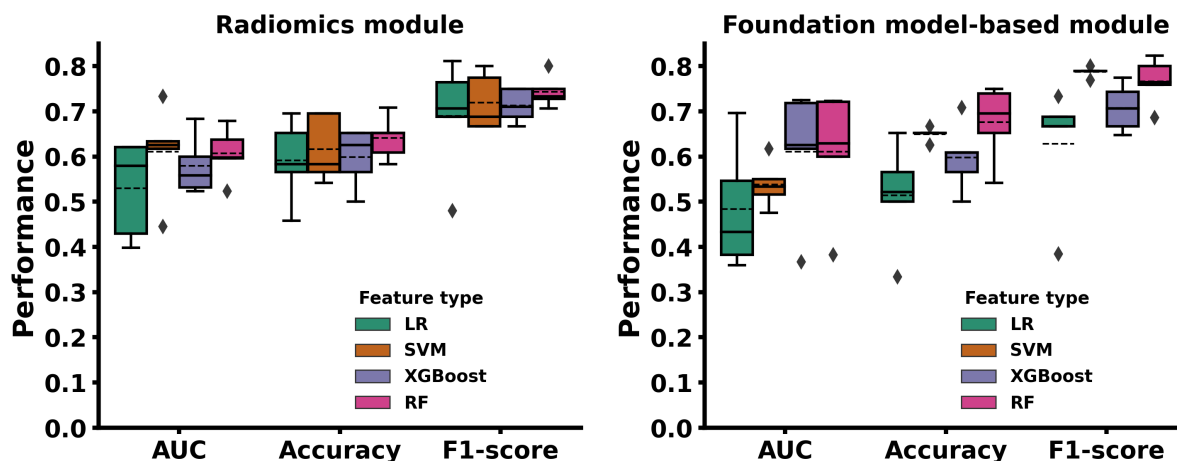


FIGURE 8.7 Comparison of the performance of different machine learning models for recurrence prediction using radiomics features (left) and features derived from the foundation model (right). The boxes indicate the interquartile range, the whiskers the 1.5 interquartile range and the dots represent the outliers.

mal protein RPL10 (ribosomal protein L10). While pseudogenes are traditionally considered non-functional "junk" DNA, emerging research is increasingly focusing on how pseudogenes may play regulatory roles in gene expression or other cellular processes in cancer [312]. The MAGE family (MAGEA2, MAGE2B, MAGA12, MAGE3) are genes normally expressed during embryogenesis, re-expressed by cancer cells during de-differentiation and are associated with poor prognosis in a number of common cancers [313]. Moreover, it currently represents a prime target for cancer immunotherapy, as MAGE proteins generate antigens recognized by immune T cells.

Moreover, Figure 8.12 shows the top 20 radiomics features listed in a decreasing order of their mean absolute SHAP value. Here, 17 out of the 20 top features were texture-related, highlighting the importance of the textural patterns in CRLM prognostication. On the other hand, the Fong CRS exhibited the largest SHAP value among clinical features (Figure 8.11), underlying its significant prognostic value.

To evaluate the correlation between the top differentially expressed genes, radiomics features and the Fong CRS, we depicted the correlation matrix between the different features. As shown in Figure 8.13, while gene expressions exhibited positive inter-correlations, different modalities were less correlated between each other. The absence of a significant correlation between the clinical score, the radiologic and the gene expression features, supports their complementary nature to improve prognostication in an integrative model as we proposed.

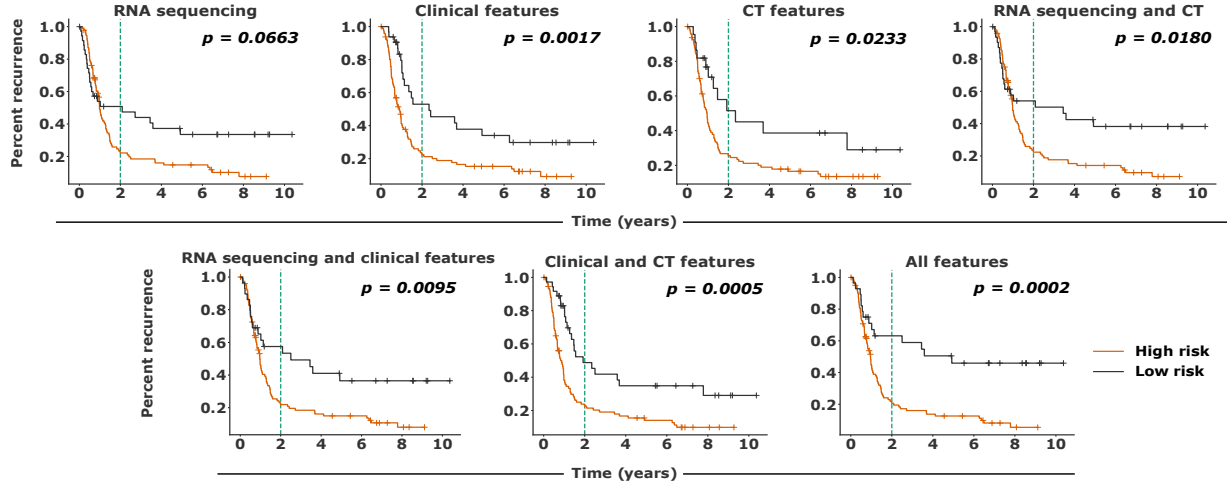


FIGURE 8.8 Kaplan-Meier curves obtained with the models trained using different combinations of modalities : RNA sequencing; CT features; Clinical features; RNA sequencing + CT features; RNA sequencing + clinical features; Clinical + CT features; RNA sequencing + clinical + CT features (proposed pipeline). The vertical dashed line highlights the stratification at 2 years post-hepatectomy.

8.6 Discussion

In this study, we proposed a prognostic pipeline integrating CT images, transcriptomics and clinical data. To our knowledge, this is the first work introducing a multi-omics pipeline for the prediction of the clinical outcomes of CRLM patients. Our model exhibited promising results, forecasting the odds of metastatic recurrence within two years after resective surgery with an AUC of 0.75 $\pm$ 0.05. Our pipeline learns complex biological and imaging features, while also integrating traditional scores widely used by oncologists, such as the clinical Fong CRS. The proposed multi-omics pipeline outperformed other approaches using single or dual data sources. Its performance was the only one that was significantly better than the Fong CRS, while its counterparts trained on one or two feature sources failed to reach statistical significance. Our approach also outperformed the multimodal pipelines conceived using early or late feature fusion methods. Processing each modality separately allowed to alleviate the burden of high-dimensionality compared to the early fusion approach, while also providing an integrative analysis of the different score types. In early fusion, a simple combination of features at the beginning of the pipeline leads to the introduction of unwanted noise due to the high dimensionality of the feature set, with no guidance of the model towards the most salient features. In the late fusion approach, features from the three modalities are processed separately and there is no interaction between them. Finally, the adopted intermediate fusion approach allows for a balanced model guidance and inter-modality interaction, leading to the

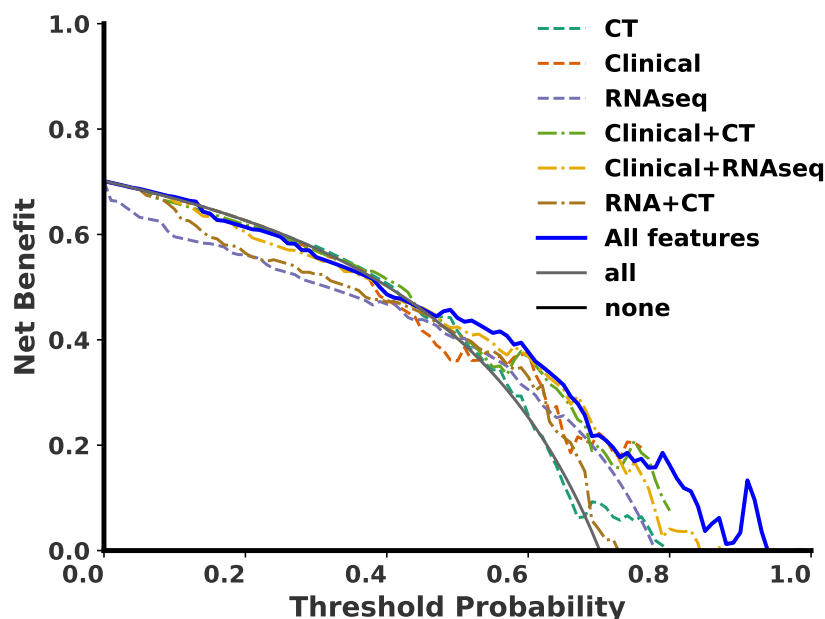


FIGURE 8.9 Decision curve analysis by modality. The gray curve represents the approach where all patients are considered to have a high recurrence risk and the black x-axis represents the approach where no patient is considered to have a high recurrence risk. The blue curve represents the net benefit of the proposed multi-omics pipeline.

best performance.

One of the major contributions of this work is the RNA sequencing component which tackles several challenges in handling transcriptomics data. Undetected transcripts, the high dimensionality of gene expression datasets and the relatively small sample size linked to the high cost of next-generation sequencing, are among the contributing factors that have hindered the deployment of deep learning models to the human transcriptome for CRLM prognostication. In this work, we first addressed the issue of undetected transcripts in some RNA sequencing samples by proposing a GAIN-based module which leverages an adapted GAN to learn the distribution of the gene expression dataset and generate plausible imputations of missing expression counts. Additionally, we mitigated the effects of the high dimensionality in RNA sequencing datasets by leveraging the CTGAN model which creates fictitious gene expression samples based on the original training instances and their corresponding recurrence outcomes. The subsequent differential gene expression analysis revealed several tumorigenesis-related genes. MAGE genes, including MAGEA2 and MAGEA2B, are expressed in different cancer types including CRC [314]. AMATS18 is a tumor suppressor and its deficiency has been associated with colon cancer formation [315]. CHGA has long been linked to neuroendo-

TABLEAU 8.2 Comparison of the performance of the different state-of-the-art tabular models for recurrence prediction using RNA sequencing data. Models with statistically significant differences with the FT-Transformer are marked with an asterisk ($p < 0.05$)

Category	Model	AUC	Accuracy	Precision	Recall	F1-score
Machine learning algorithms	Decision Tree	*0.45 (0.10)	0.48 (0.10)	0.61 (0.09)	0.54 (0.12)	0.57 (0.11)
	Random Forest	0.58 (0.06)	0.64 (0.03)	0.65 (0.02)	0.94 (0.04)	0.77 (0.02)
	SVM	0.62 (0.14)	0.63 (0.04)	0.64 (0.02)	0.96 (0.05)	0.77 (0.03)
Gradient boosting algorithms	CatBoost	0.56 (0.13)	0.64 (0.07)	0.65 (0.04)	0.95 (0.07)	0.77 (0.05)
	AdaBoost	*0.60 (0.09)	0.64 (0.04)	0.68 (0.02)	0.82 (0.11)	0.74 (0.04)
	XGBoost	0.61 (0.09)	0.62 (0.06)	0.69 (0.02)	0.75 (0.12)	0.72 (0.06)
Deep learning algorithms	MLP	0.60 (0.10)	0.58 (0.10)	0.77 (0.14)	0.67 (0.39)	0.60 (0.26)
	ResNet	0.61 (0.14)	0.64 (0.08)	0.69 (0.04)	0.80 (0.12)	0.74 (0.08)
	TabNet	0.59 (0.14)	0.65 (0.01)	0.65 (0.01)	1.00 (0.00)	0.79 (0.00)
	TabPFN	0.63 (0.09)	0.67 (0.03)	0.67 (0.03)	0.99 (0.02)	0.79 (0.01)
	TabTransformer	0.65 (0.07)	0.63 (0.07)	0.72 (0.07)	0.71 (0.11)	0.71 (0.06)
	FT-Transformer	0.68 (0.06)	0.67 (0.06)	0.73 (0.07)	0.81 (0.07)	0.76 (0.03)

TABLEAU 8.3 Comparison of different feature fusion approaches. Statistically significant differences ($p < 0.05$) with the proposed pipeline (using intermediate fusion) are marked by an asterisk.

Fusion	AUC	Acc	Prec	Rec	F1
Early	*0.62 (0.06)	0.67 (0.07)	0.69 (0.06)	0.94 (0.05)	0.79 (0.04)
Late	0.73 (0.08)	0.69 (0.07)	0.70 (0.05)	0.92 (0.07)	0.79 (0.05)
<i>Intermediate</i>	0.75 (0.05)	0.70 (0.07)	0.72 (0.06)	0.88 (0.08)	0.79 (0.05)

crine carcinoma and was recently proposed as a potential biomarker for CRC [316]. CNM1 has been associated with tumorigenesis in hepatocellular carcinoma [317]. When it comes to LCN15, although its exact function is unknown, it belongs to the family of lipocalins of which members have been associated with cell proliferation in CRC [318]. Moreover, two ribosomal protein pseudogenes were among the top differentially expressed. Ribosomal proteins are key constituents of ribosomes, which are the cellular structures responsible for protein synthesis. Pseudogenes are sequences resembling protein coding genes that can control gene expression and may interfere with the messenger RNA of the parent gene. Overall, this unbiased transcriptomic analysis may point to novel or neglected therapeutic targets that could be further explored to improve patient prognostication and clinical outcome.

The imaging component of the multi-omics pipeline leverages two complementary modules, working in tandem to boost its predictive capability : a radiomics module and a foundation model-based deep learning module. While the former extracts interpretable, well-established features, the latter derives abstract features, representative of the involved biological mechanisms. Incorporating these two feature types in a model cascading fashion allowed to circumvent the curse of dimensionality issue by training each predictor on a manageable fea-

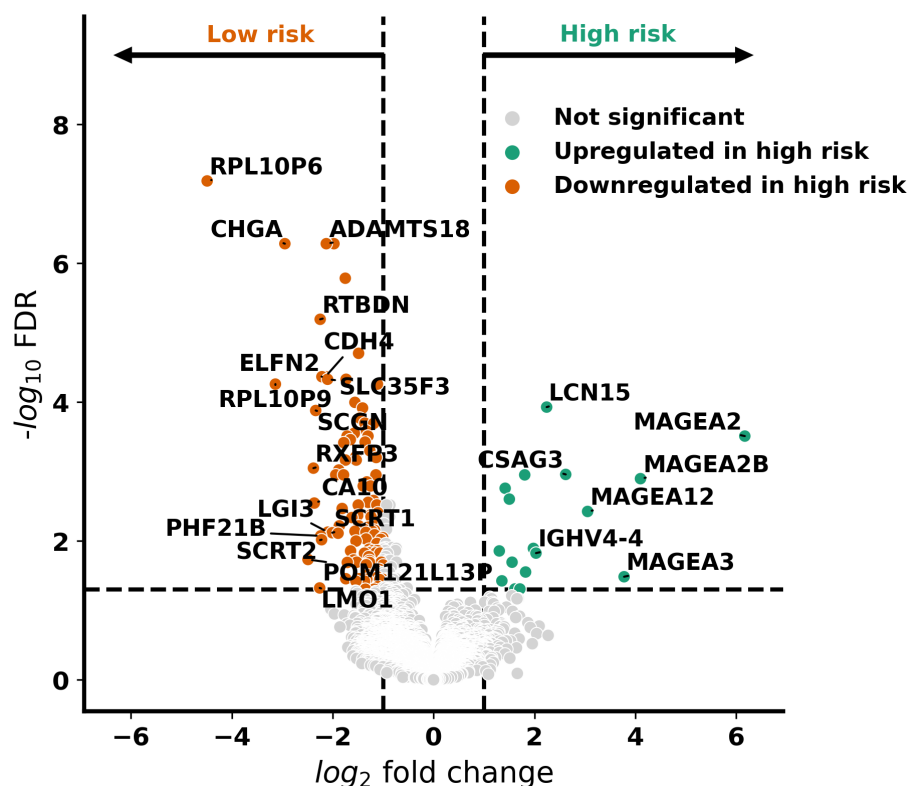


FIGURE 8.10 Volcano plot of the differential gene expression results. The gray dots represent the genes that are not differentially expressed, the green dots denote the differentially expressed genes, and the brown dots represent six chosen genes among the top differentially expressed. The fold change of a given gene is computed as the ratio of its expression in the high versus the low risk groups. FDR : False Discovery Rate. The dashed horizontal and vertical lines mark the thresholds for considering a gene as differentially expressed.

ture volume. The ensuing imaging component exhibited an improved performance with an accuracy of 0.70 and an AUC of 0.62. Foundation models have recently revolutionized the field of natural language processing, paving the way for the success of large language models. Foundation models, often characterized by a large number of parameters, are trained in a self-supervised manner on substantial, diverse datasets and serve as a foundation for a range of different tasks. In computer vision, several foundation models have been proposed for different purposes, including object detection (e.g. DETR [319]), segmentation (e.g. SAM [320]) and embedding extraction (e.g. CLIP [321], ImageBind [322]). Within the biomedical field, where expert annotations are often scarce, the application of foundation models constitutes a fertile ground for discovery. For instance, MedSAM [171], leveraged in this work, is an adaptation of SAM to segment medical images. Another example is XrayGPT [323], a language-vision model conceived to assess and respond to inquiries regarding chest radiographs. In this work,

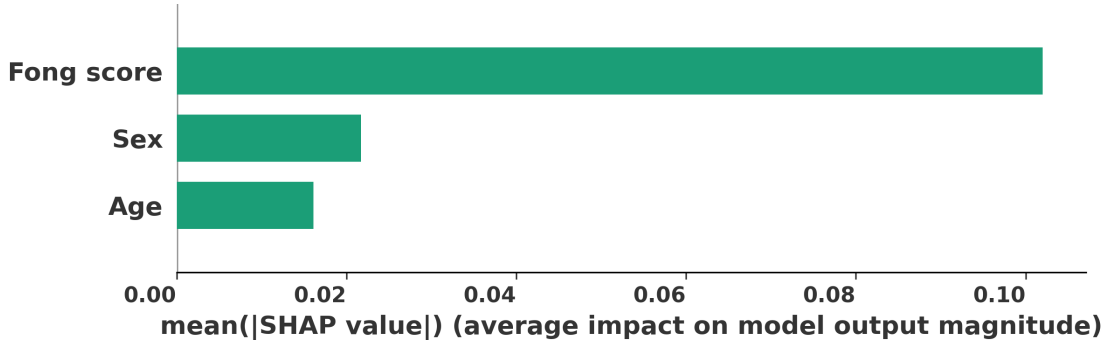


FIGURE 8.11 SHAP analysis of the clinical features.

we extended the application of the foundation model proposed in [7], initially validated for the development of diagnostic and prognostic biomarkers of lung lesions, to CRLM prognostication.

While previous studies used medical images to predict prognostic proxies of clinical outcomes of CRLM such as clinically established immune markers [172] or genetic mutations [324], the proposed pipeline successfully predicted the likelihood of recurrence and exhibited a superior performance in stratifying recurrence risk compared to its single and dual-modal counterparts. Its clinical utility was also validation based on the decision curve analysis method.

Our work has some limitations that should be addressed in future research. First, our study relies on data collected in a single center with limited heterogeneity in terms of patients characteristics and relatively limited sample size ($n = 129$). Therefore, it is essential to validate the prognostic capabilities of our model on external data in a multi-centric study including different acquisition parameters and patients with various ethnicity. Second, despite the uniqueness of our proposed pipeline in incorporating multiple modalities to provide a holistic evaluation of CRLM, additional modalities could be included to further boost its performance. In fact, several studies have reported the prognostic value of magnetic resonance and histopathological images [289]. Particularly, El Foraici et al. [154] has shown promising results leveraging histopathological images for the prediction of survival, recurrence and the tumor regression grade of CRLM. In our work, we chose to use CT images for their wide availability, especially in the pre-operative setting unlike histopathological images which are collected post-operatively. Nevertheless, it is important to concede that incorporating these imaging modalities could complement the extracted information provided by CT scans. Third, our pipeline leverages computationally expensive deep learning architectures such as Transformers and generative models. This might thwart its deployment in resource-constrained settings. The assessment of the benefits of our pipeline, along with its memory consumption, should

be conducted by the end user based on their specific needs.

8.7 Conclusion

In summary, this study proposes a multi-omics pipeline for the prediction of the risk of early recurrence in CRLM after curative resection. Our approach leverages RNA sequencing, CT images and prognostic clinical features simultaneously and tackles several challenges in handling each data type. In particular, generative imputation techniques and synthetic instance generation were proposed to handle missing expression counts and dimensionality issues in RNA sequencing, respectively. On the other hand, we harnessed advances in the conception of versatile foundation models to enrich the representations of the imaging features and enhance their predictive capabilities. The clinical score, imaging, and genomic features were not correlated and thus complementary in predicting recurrence outcomes. The pipeline demonstrated a high prognostic accuracy, as well as a promising clinical utility and risk stratification efficacy. It also outperformed several comparative approaches. Our future work should focus on addressing the challenges that might limit the clinical deployment of our pipeline. Specifically, building a larger dataset combining different medical centers would increase the trustworthiness of our pipeline. In addition, including supplementary modalities such as histopathological images could potentially boost its predictive performance and generalization on unseen data.

Acknowledgments

SK and ST are researchers at the Research Center of the University of Montreal Hospital (CRCHUM), supported by the Fonds de Recherche du Québec - Santé (FRQS). RS is supported by the Fonds de recherche du Québec - Nature et Technologie (FRQNT). ST is supported by the FRQS Clinician Scientist Senior Salary Award (No. 349577). ST obtained support for RNA sequencing by a Terry Fox Marathon of Hope Cancer Centres Network grant and pre-clinical research funding from Bristol Myers Squibb and Turnstone Biologics. The CRCHUM Hepatopancreatobiliary cancer biobank and prospective database is supported by the Université de Montréal Roger des Groseillers Research Chair in Hepatopancreatobiliary Surgical Oncology.

The authors wish to express their gratitude to L. Rousseau, S. Langevin, and J. Bilodeau for their contributions to patient recruitment, biospecimen acquisition and processing, and to the CRCHUM histopathology core facility and R. Kosovskaia for RNA extraction and sequencing logistics. Moreover, M. Cerny, V. Hamilton, T. Derennes, and A. Ilinca are recognized for

their work in editing segmentations and differentiating liver metastases from benign coexisting liver lesions.

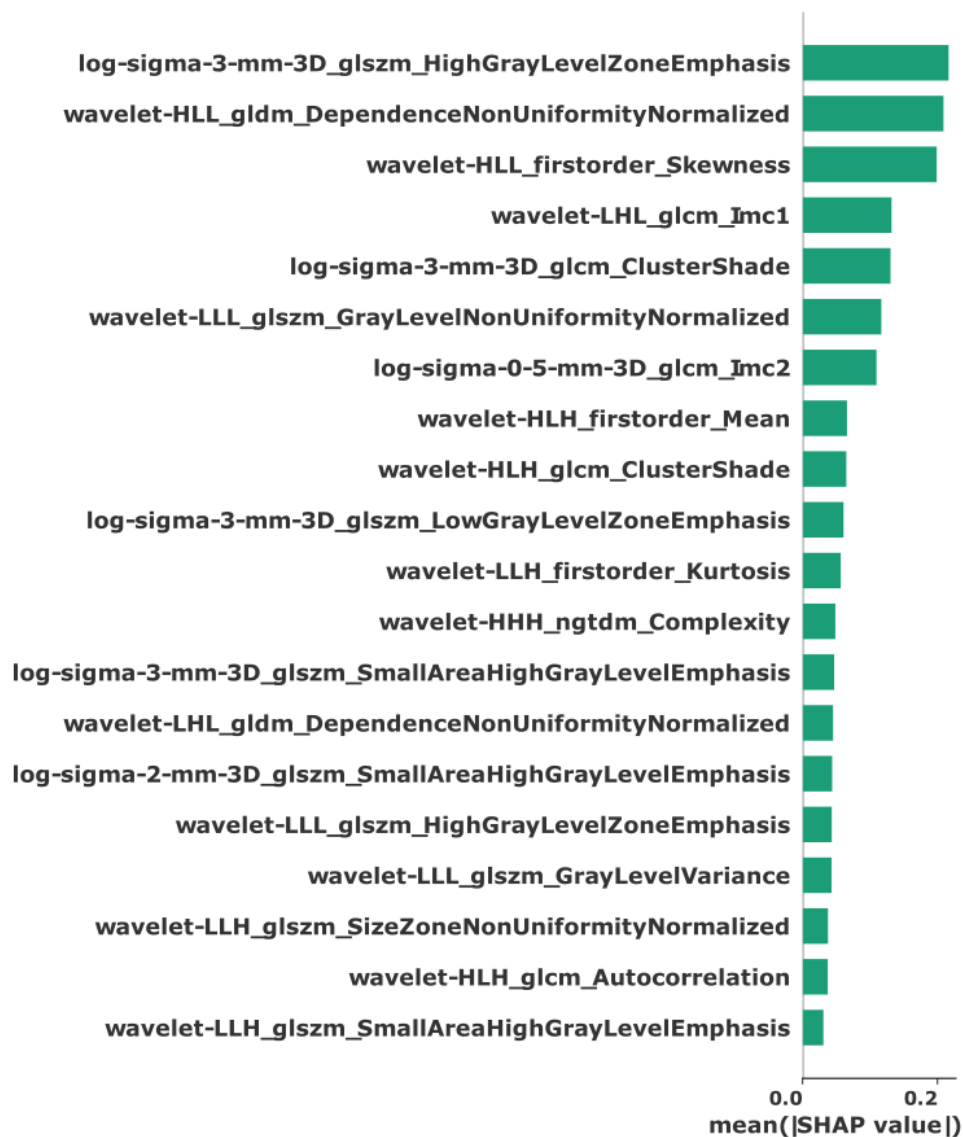


FIGURE 8.12 SHAP analysis of the radiomics features.

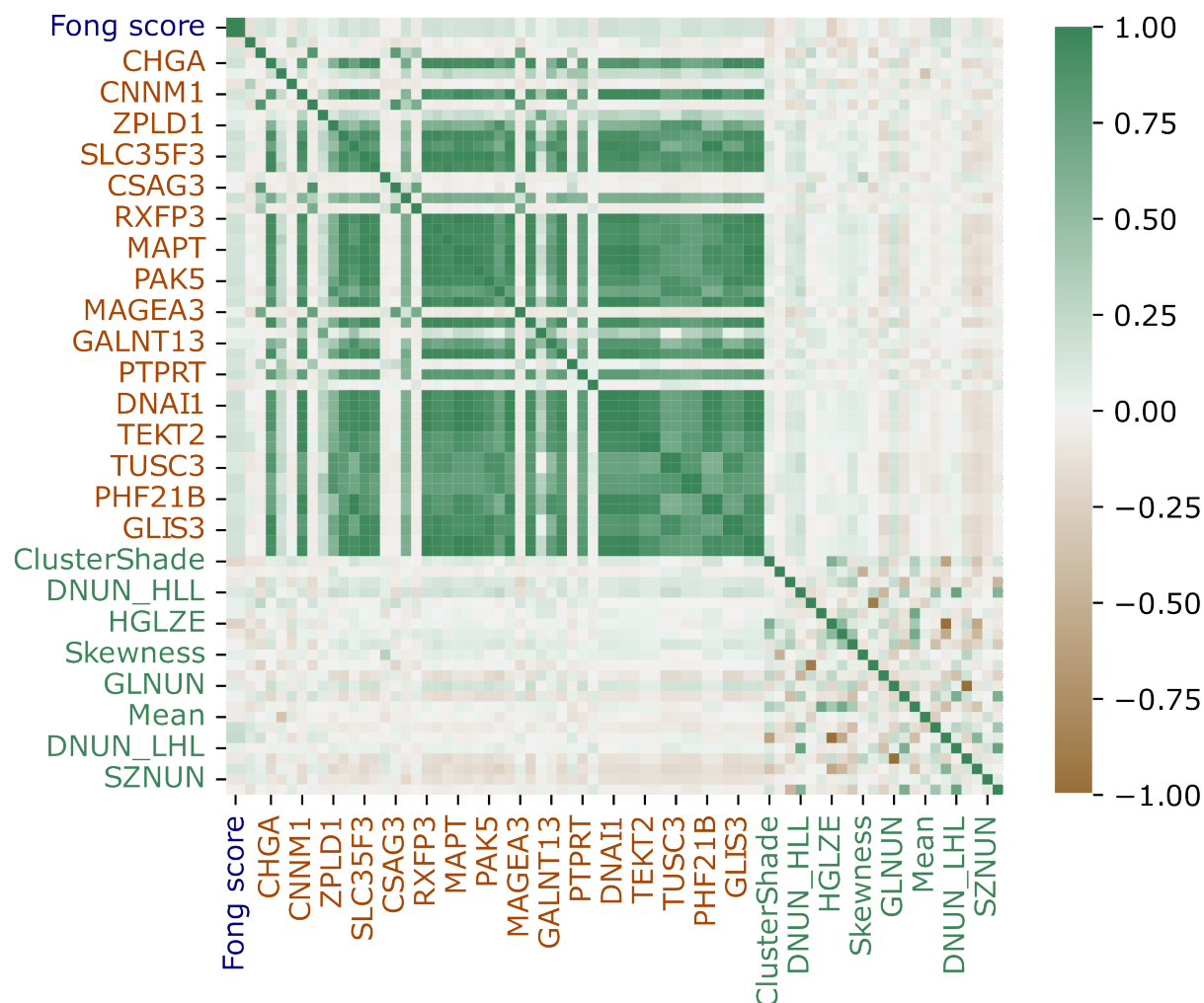


FIGURE 8.13 Correlation matrix between top differentially expressed genes, CT radiomics features and the Fong CRS. While some RNA sequencing features are inter-correlated, different modalities did not exhibit a high correlation between each other. DNUN-HLL, DependenceNonUniformityNormalized-HighLowLow; DNUN-LHL, DependenceNonUniformityNormalized-LowHighLow; HGLZE, HighGrayLevelZoneEmphasis; GLNUN, GrayLevelNonUniformityNormalized; SZNUN, SizeZoneNonUniformityNormalized. The clinical score is presented in blue, genes in brown and imaging features in green.

CHAPITRE 9 DISCUSSION

La méthodologie générale de cette thèse a été guidée par quatre objectifs de recherche interconnectés, motivant différentes solutions novatrices pour le pronostic des MHCR et du CPNPC. La thèse présente une exploration complète du pronostic multimodal qui comble les lacunes critiques de l'oncologie personnalisée. Ce travail fait progresser le domaine en intégrant des techniques d'apprentissage automatique de pointe avec diverses modalités de données pour améliorer la stratification des risques et l'orientation des traitements. Tout d'abord, un pipeline prédictif a été développé pour la prédiction de l'expression de CD73, un marqueur immunitaire hautement pronostique dans les MHCR. Deuxièmement, nous avons prédit un profil immunologique à partir des TDM basé sur la densité d'infiltration de CD3+ et l'expression du CMH-I dans les MHCR. Troisièmement, nous avons proposé un pipeline pour prédire l'expression de PD-L1 dans le CPNPC et validé sa généralisabilité à tous les stades de la maladie et à tous les schémas thérapeutiques. Enfin, nous avons proposé un pipeline multi-omique pour la pronostication des MHCR réséqués en combinant les modalités susmentionnées, tout en relevant les défis techniques liés à chaque modalité.

Cette thèse s'appuie sur des études antérieures qui ont démontré des associations entre les caractéristiques radiomiques extraites des images TDM et les profils d'infiltration immunitaire dans le microenvironnement tumoral. Notamment, une étude publiée dans *The Lancet Oncology* [104] a rapporté des corrélations significatives entre les caractéristiques radiomiques dérivées de la TDM et l'expression des CD8 dans plusieurs types de cancer. Cependant, des études similaires portant spécifiquement sur le MHCR restent rares ou pratiquement inexistantes.

Dans ce travail, nous avons d'abord étudié l'association entre les caractéristiques radiomiques et l'expression de CD73 dans le MHCR. À cette fin, nous avons utilisé TabNet, un modèle d'apprentissage profond de pointe au moment de l'étude conçu pour les données tabulaires. TabNet a été choisi en raison de ses performances compétitives par rapport aux algorithmes de boosting de gradient précédemment considérés comme de pointe, de sa robustesse lorsqu'il est appliqué à de petits ensembles de données et de son interprétabilité inhérente facilitée par des masques d'attention à l'échelle de l'instance. Nos résultats expérimentaux ont confirmé les performances supérieures du modèle dans ce contexte.

Pour la prédiction du profil immunitaire CD3-CMHI, TabNet seul n'a pas atteint le même niveau de performance que pour CD73. Par conséquent, nous avons proposé l'architecture Multi-TabNet, qui intègre plusieurs modèles TabNet afin d'améliorer la précision de la pré-

diction.

Lors de l'élaboration d'un cadre de prédiction multi-omique, nous avons entraîné un FT-Transformer sur des données de séquençage d'ARN. Ce modèle a été choisi pour sa capacité à transformer les caractéristiques d'entrée à haute dimension en représentations compactes et spécifiques à tâche en main. Pour remédier aux problèmes posés par la haute dimensionnalité et le nombre limité d'échantillons de séquençage d'ARN, nous avons synthétisé des instances d'ARN supplémentaires en utilisant le CTGAN. Le CTGAN a été choisi pour sa capacité à modéliser des distributions de données tabulaires grâce à des mécanismes de génération conditionnelle. Pour la composante imagerie de l'approche multi-omique, nous avons proposé de combiner des caractéristiques radiomiques avec des caractéristiques extraites d'un modèle de fondation. Nous avons ensuite appliqué des classificateurs de forêt aléatoire, car ils effectuent intrinsèquement une sélection des caractéristiques et sont bien adaptés à la réduction de la dimensionnalité tout en identifiant les caractéristiques les plus informatives.

Enfin, pour prédire l'expression de PD-L1 dans le CPNPC, nous avons proposé une approche d'auto-apprentissage. Étant donné le nombre limité d'échantillons annotés pour un ensemble de données multi-institutionnel, nous avons exploité des données publiques non annotées dans un schéma d'auto-apprentissage, ce qui a permis d'améliorer les performances prédictives.

9.1 Prédiction des marqueurs immunitaires non invasifs dans les MHCR

Dans le chapitre 5, le développement d'un pipeline radiomique basé sur la TDM pour la classification de l'expression de CD73 dans les MHCR représente un progrès significatif vers l'évaluation non invasive des biomarqueurs immunitaires. Avant notre étude, aucun marqueur immunitaire non invasif n'avait été étudié pour le pronostic des MHCR. Le CD73 joue un rôle important dans la suppression immunitaire dans le CCR primaire et les MHCR en ciblant la voie de l'adénosine dans le microenvironnement tumoral. En évaluant de manière non invasive l'expression du CD73 dans les MHCR à l'aide d'images de TDM préopératoires, nous fournissons un outil d'assistance qui peut être utilisé de manière longitudinale par les oncologues à tout moment pendant le traitement. En effet, la nature non invasive du biomarqueur permet des évaluations plus fréquentes de la biologie tumorale, ce qui est crucial pour suivre la réponse au traitement et la progression de la maladie.

Nos travaux ont révélé une corrélation significative entre le score rad-CD73 et les niveaux réels d'expression du CD73. Les patients classés comme rad-CD73^{High} présentaient une survie médiane sans récurrence et une SSC plus courtes, ce qui souligne la valeur pronostique de ce score radiomique. Les résultats confirment qu'une forte expression du CD73 est associée à de

moins bons résultats, ce qui renforce la nécessité de stratégies de traitement personnalisées dans le cas des MHCR.

Pour évaluer l'expression du marqueur immunitaire, nous avons utilisé TabNet, un modèle d'apprentissage profond conçu pour les données tabulaires. L'un des principaux avantages de TabNet est son interprétabilité intégrée, qui répond aux défis de transparence qui ont jusqu'à présent limité l'adoption clinique des modèles d'apprentissage profond. De plus, TabNet est entraîné sur des caractéristiques radiomiques interprétables, ce qui permet d'associer des modèles d'imagerie, tels que l'hétérogénéité texturale, à l'expression de CD73 et, en fin de compte, aux mécanismes immunosuppresseurs contribuant à la croissance tumorale, aux métastases et à la progression de la maladie. En particulier, les caractéristiques liées à la texture ont été identifiées comme des indicateurs clés de l'expression de CD73, les textures homogènes étant corrélées à des niveaux plus élevés de CD73. Cette découverte est vitale pour les cliniciens, car elle renforce la confiance dans son application en milieu clinique.

Malgré ses contributions significatives à la mise au point d'un marqueur immunitaire non invasif pour le cancer du foie métastatique, l'étude a été limitée à un seul centre avec un échantillon de taille relativement petite. Ces limites, ainsi que la nature rétrospective de l'analyse, peuvent affecter la généralisabilité des résultats. Il est donc important de valider le biomarqueur proposé de manière prospective avec de grands ensembles de données intégrant des patients présentant diverses caractéristiques démographiques. De telles études prospectives à grande échelle pourraient confirmer la valeur prédictive du biomarqueur dans l'identification des patients les plus susceptibles de répondre aux agents thérapeutiques ciblant le CD73 ou les récepteurs de l'adénosine. L'évaluation prospective de ce biomarqueur pourrait également élucider plusieurs hypothèses cliniquement pertinentes, notamment si une surveillance postopératoire intensifiée après une résection des MHCR facilite la détection et le traitement précoces de la récurrence et si un traitement systémique adjuvant après une résection des MHCR réduit le risque de récurrence et de mortalité chez les patients positifs au biomarqueur. Ces recherches pourraient permettre d'élaborer des stratégies de traitement personnalisées et d'optimiser la prise en charge des patients atteints d'une métastatique hépatique résécable d'origine colorectale .

Dans une étude de suivi présentée au chapitre 6, nous avons cherché à prédire de manière non invasive un autre marqueur pronostique du CCR métastatique : le profil immunitaire $CD3^{Low}CMH^{High}$. Alors qu'une forte expression de CD73 est associée à un pronostic plus défavorable en cas de MHCR, les patients qui présentent une faible densité d'infiltration de cellules T CD3+, ainsi qu'une forte expression de CMH-I ($CD3^{Low}CMH^{High}$), ont une récurrence plus précoce, une SSC plus courte et une réponse plus faible à la chimiothérapie. Alors que

le CD73 est une enzyme associée à l'évasion immunitaire liée à l'adénosine, les cellules T CD3+ sont les cellules tueuses de tumeurs et le CMH-I permet la présentation de l'antigène aux cellules T CD8+. La combinaison de la densité des cellules T CD3+ et de l'expression du CMH-I reflète l'immunité antitumorale et offre une valeur pronostique plus prononcée.

L'étude présentée au chapitre 6 montre que l'approche d'ensemble Multi-TabNet proposée a donné de bons résultats dans la prédiction du profil immunitaire $CD3^{Low}CMH^{High}$ à l'aide d'images de TDM préopératoires. L'approche proposée a surpassé d'autres modèles d'apprentissage automatique, y compris TabNet. L'étude repose sur l'hypothèse qu'en utilisant les caractéristiques radiomiques extraites des images de TDM et en formant des modèles robustes, il serait possible de comprendre en profondeur le microenvironnement tumoral et ses caractéristiques immunologiques. Le potentiel de cette approche en tant que biomarqueur non invasif met en évidence sa pertinence dans les efforts en cours pour personnaliser le traitement du cancer.

En ce qui concerne la pertinence clinique des études susmentionnées, un rad-CD73 validé de manière prospective pourrait servir de biomarqueur non invasif des agents thérapeutiques ciblant les récepteurs CD73 ou l'adénosine, tandis que le profil d'imagerie $CD3^{Low}CMH^{High}$ pourrait permettre une évaluation immunitaire non invasive des MHCR. La pertinence pronostique des deux scores pronostiques a également été démontrée dans nos travaux.

9.2 Prédiction de l'expression de PD-L1 dans le CPNPC

Le troisième objectif de la thèse décrit au chapitre 7 est de proposer un autre outil non invasif qui pourrait guider la sélection de candidats appropriés pour les ICI, en prédisant l'expression de PD-L1 dans le CPNPC à l'aide d'images de TDM de prétraitement. Comme dans le cas du CD73, le PD-L1 présente des effets immunosuppresseurs bien qu'il agisse par un mécanisme différent. Le CD73 dans les MHCR favorise l'immunosuppression en catalysant la dégradation de l'AMP extracellulaire en adénosine, ce qui crée un microenvironnement tumoral immunosuppresseur. En revanche, les ligands PD-L1 dans le CPNPC inhibent directement le rôle des cellules immunitaires en se liant au PD-1 exprimé sur les lymphocytes T, ce qui entraîne un épuisement immunitaire et une réduction des réponses immunitaires antitumorales. Pour les patients atteints d'un CPNPC avancé sans mutations conductrices, le PD-L1 est le biomarqueur ICI le plus fiable et dicte la stratégie de traitement. Pour les patients atteints d'un CPNPC à un stade limité qui peuvent être opérés, l'association néoadjuvante d'ICI et de chimiothérapie est devenue la norme de soins et les patients présentant un taux de PD-L1 > 50% bénéficient davantage de la chimiothérapie néoadjuvante associée à l'ICI.

Le développement d'un biomarqueur non invasif basé sur la radiomique pour l'expression de PD-L1 dans le CPNPC, appelé rad-PDL1, constitue une avancée significative dans la médecine personnalisée. En exploitant les techniques d'apprentissage automatique et l'imagerie TDM, l'étude surmonte les limites associées à l'évaluation traditionnelle de la PD-L1 basée sur la biopsie, telles que l'hétérogénéité intratumorale et intertumorale. Cette méthode non invasive pourrait potentiellement améliorer la prise de décision thérapeutique grâce à des tests répétés, réduire le besoin de procédures invasives et améliorer les résultats pour les patients en guidant mieux l'utilisation de l'immunothérapie dans le CPNPC.

L'étude du biomarqueur rad-PDL1 améliore considérablement les approches radiomiques existantes grâce à quatre innovations clés. Premièrement, sa nature multi-institutionnelle, qui intègre des patients atteints de CPNPC avancé dans trois centres médicaux, résout directement la question de la reproductibilité qui affectait les études radiomiques monocentriques antérieures. Deuxièmement, la validation du modèle à la fois sur le CPNPC avancé et le CPNPC à un stade limité représente une avancée cruciale par rapport aux études antérieures, notamment grâce à la démonstration de la généralisabilité du modèle dans des contextes néoadjuvants combinant chimiothérapie et ICI. Cela souligne la robustesse et la généralisabilité du biomarqueur rad-PDL1 et constitue l'une des principales contributions de ce travail, car les études précédentes n'étudiaient pas la généralisabilité des modèles entre les stades. Troisièmement, le pipeline d'auto-apprentissage est une solution novatrice au problème de pénurie des données annotées qui limitait les modèles antérieurs. En combinant un FT-Transformer avec un pseudo-étiquetage, les performances prédictives du modèle se sont améliorées. Quatrièmement, notre pipeline développé démontre son interprétabilité grâce à des analyses révélant que les caractéristiques hétérogènes d'imagerie des lésions du CPNPC étaient significativement associées à des niveaux élevés d'expression du biomarqueur rad-PDL1. Cette corrélation entre les caractéristiques radiologiques et les profils moléculaires fournit des informations sur les mécanismes biologiques sous-jacents et soutient l'utilité clinique potentielle du modèle.

Lors de l'interprétation des biomarqueurs non invasifs dérivés de l'imagerie, il est essentiel de reconnaître que, bien que ces biomarqueurs soient conceptuellement liés aux niveaux d'expression de protéines importantes pour le pronostic telles que PD-L1 et CD73, ils représentent des entités fondamentalement différentes. En conséquence, ils doivent être considérés comme des compléments plutôt que comme des remplacements des marqueurs moléculaires. En effet, les caractéristiques radiologiques identifiées par les modèles basés sur les TDM peuvent ne pas refléter directement l'expression des protéines, mais plutôt capturer des effets biologiques associés indirects qui sont plus facilement détectables par l'imagerie. Par exemple, la présence d'une nécrose centrale peut provoquer des microenvironnements hypoxiques qui, à leur tour,

induisent une surexpression de CD73 [44]. Les régions nécrotiques sont généralement visibles au TDM sous la forme de régions d’atténuation du liquide. De même, des niveaux élevés d’expression de PD-L1 ont été associés à une quantité accrue d’acides gras polyinsaturés [325], et l’accumulation de graisse est connue pour être détectable par TDM.

L’intégration clinique des biomarqueurs proposés dans cette thèse nécessite une compréhension approfondie du comportement du modèle, une généralisation démontrée sur des ensembles de données externes et une validation par le biais d’études à grande échelle impliquant des patients présentant des caractéristiques démographiques distinctes et des protocoles d’acquisition d’imagerie variés. En outre, le déploiement réussi dans les établissements de soins de santé exige une approbation réglementaire (par exemple, par Santé Canada, FDA), une intégration facile dans les flux de travail cliniques existants (par exemple, PACS), et des preuves provenant d’essais cliniques prospectifs qui démontrent l’utilité clinique et la rentabilité.

9.3 Stratification du risque de récurrence multi-omique dans le cas des MHCR

L’intégration de la radiomique préopératoire par TDM, du séquençage de l’ARN et des variables cliniques dans un pipeline pronostique unifié représente une contribution significative à la pronostication des MHCR. Les études antérieures sur le pronostic des MHCR se sont principalement concentrées sur des modalités uniques, ce qui limite leur capacité à fournir une évaluation complète de la biologie tumorale. En surpassant de manière significative le SRC de Fong dans la prédiction des récurrences, la thèse démontre le potentiel synergique de l’intégration des données multi-omiques.

La fusion des données radiomiques avec les caractéristiques transcriptomiques pourrait permettre de créer des modèles prédictifs encore plus puissants, en combinant la nature non invasive des biomarqueurs d’imagerie avec la profondeur moléculaire fournie par le séquençage de l’ARN. Néanmoins, cela nécessite de surmonter au préalable les défis techniques liés au traitement des données transcriptomiques. L’analyse des données transcriptomiques, en particulier le séquençage de l’ARN, présente des défis importants, notamment le problème des transcrits géniques non détectés et la dimensionnalité élevée des ensembles de données d’expression génique. Dans cette thèse, nous avons développé une approche innovante pour relever ces défis grâce à l’utilisation de modèles génératifs.

Les transcrits géniques non détectés, peuvent être attribués à différents facteurs comme les pertes lors de la préparation des bibliothèques ou à une profondeur de séquençage insuffisante [35]. Les données manquantes ont plusieurs implications qui peuvent affecter l’analyse et l’interprétation des profils transcriptomiques, conduisant à une perte d’informations biologiques

précieuses. Le fait de rejeter les valeurs manquantes est souvent préjudiciable, car cela compromet l'intégrité des analyses en aval et peut fausser les résultats, en particulier dans les ensembles de données complexes tels que ceux utilisés dans la recherche sur le cancer.

L'application réussie de GAIN pour traiter les données manquantes en transcriptomique a des implications plus larges pour le domaine de la recherche biomédicale. En améliorant le traitement des valeurs manquantes, nous pouvons renforcer la robustesse des analyses dans divers contextes, notamment la classification des cancers et d'autres études omiques. L'imputation générative des valeurs d'expression manquantes du séquençage de l'ARN a permis d'améliorer la prédiction des performances des sous-types de CPNPC (adénocarcinome vs carcinome épidermoïde) [177]. L'application de GAIN pourrait donc être explorée dans différents ensembles de données intégrant d'autres techniques d'apprentissage automatique afin d'affiner davantage les pipelines prédictifs.

Pour remédier à la malédiction de la dimensionnalité inhérente à l'analyse des données de séquençage de l'ARN, qui est exacerbée par la taille relativement petite des échantillons, nous avons utilisé des techniques d'apprentissage génératif conditionnel. Cette approche a permis la synthèse d'instances d'expression génique artificielles qui suivent la même distribution que les données originales. En augmentant l'ensemble de données de séquençage de l'ARN avec ces instances synthétiques, nous avons préservé les informations biologiques complexes intégrées dans les données tout en atténuant les défis associés aux ensembles de données à haute dimension et à échantillons limités.

De plus, l'analyse TDM améliorée par le modèle de fondation augmente la valeur prédictive de la radiomique interprétable, qui a dominé la recherche sur les MHCR. Le modèle de fondation utilisé dans nos travaux s'appuie sur un pré-entraînement auto-supervisé effectué sur de vastes ensembles de données TDM, ce qui lui permet de déchiffrer les propriétés texturales complexes des tumeurs associés à la progression métastatique et à la récurrence de la maladie.

9.4 Résumé

Les recherches présentées dans ces quatre chapitres offrent un aperçu significatif des différents aspects du développement des biomarqueurs et de la stratification des risques. Ensemble, ils représentent une approche hybride pour faire progresser la médecine personnalisée basée sur l'intelligence artificielle, en particulier dans le cas des MHCR et du CPNPC.

La synergie entre les trois premiers chapitres réside dans leur potentiel à améliorer la sélection des patients pour l'immunothérapie. Alors que le rad-CD73 pourrait être validé dans des

études prospectives en tant que biomarqueur de réponse aux inhibiteurs de CD73 ou de la voie de l'adénosine, le rad-PDL1 pourrait être utilisé pour identifier les patients atteints de CPNPC les plus susceptibles de bénéficier d'ICI seuls ou en combinaison avec la chimiothérapie à différents stades de la maladie. De plus, le profilage immunitaire non invasif CD3/CMH-I inspiré de l'IHC pourrait fournir un score pronostique reflétant la réponse immunitaire contre les cellules tumorales. Ces approches complémentaires soulignent l'importance de développer de nouveaux biomarqueurs prédictifs pour optimiser les résultats de l'immunothérapie pour différents types de cancer.

L'approche multi-omique présentée dans la quatrième étude représente l'aboutissement des efforts visant à comprendre de manière exhaustive la progression des MHCR. En intégrant des caractéristiques de différents niveaux, notamment cliniques, d'imagerie et de transcriptomique, cette approche fournit une évaluation holistique de la biologie des MHCR. L'approche basée sur la radiomique pour le développement de biomarqueurs complète les connaissances acquises grâce au séquençage de l'ARN et les caractéristiques cliniques, ce qui permet d'améliorer considérablement les performances par rapport au SRC établi cliniquement. En conclusion, les résultats collectifs de ces études suggèrent un potentiel important pour les approches d'apprentissage automatique dans l'avancement des stratégies de médecine personnalisée.

Afin de garantir la reproductibilité des résultats et l'applicabilité clinique des approches développées, il est essentiel de prendre en compte des considérations techniques importantes telles que l'optimisation des modèles d'apprentissage profond et leur calibration. En effet, les modèles qui ne sont pas correctement optimisés peuvent présenter un surajustement ou un sous-ajustement. Les modèles surajustés présentent souvent des performances d'apprentissage élevées mais ne parviennent pas à se généraliser sur des données non vues. Les modèles sous-ajustés présentent des performances sous-optimales sur les données vues et non vues. Il est donc primordial de s'assurer que les hyperparamètres sont bien sélectionnés. A cette fin, nous avons ajouté deux annexes à la fin de cette thèse pour détailler le choix des hyperparamètres dans chaque objectif et assurer la reproductibilité des modèles.

Le calibrage d'un modèle reflète le degré d'alignement entre les probabilités de classe prédites par le modèle et les probabilités réelles observées des résultats. Un modèle mal calibré peut présenter un excès de confiance (prédire des probabilités plus élevées que les probabilités réelles) ou un manque de confiance (prédire des probabilités plus faibles que les probabilités réelles). En outre, un modèle peut être bien calibré sur les classes majoritaires mais mal calibré sur les classes sous-représentées, ce qui peut entraîner des biais potentiels. Il est donc utile d'évaluer le calibrage des modèles développés dans les travaux futurs en utilisant des techniques qualitatives et quantitatives telles que les courbes de calibrage et l'erreur

de calibrage attendue. Il pourrait également être utile d'appliquer certaines techniques de calibrage à posteriori telles que l'approche paramétrique de l'échelle de Platt ou la technique de régression isotonique non paramétrique [326].

CHAPITRE 10 CONCLUSION ET RECOMMANDATIONS

10.1 Contribution à l'avancement des connaissances

Cette thèse aborde l'écart persistant entre la biologie du cancer et la prise de décision clinique, en se concentrant sur les MHCR et le CPNPC. Nous avons développé des pipelines computationnels innovants pour la découverte de biomarqueurs immunitaires non invasifs ainsi que des approches multi-omiques pour améliorer la personnalisation du traitement et le pronostic du cancer. Collectivement, les contributions de la thèse visent à transformer le pronostic du cancer, passant de paradigmes réactifs à des approches proactives et multimodales où les données multi-omiques guident collectivement les décisions thérapeutiques. En intégrant les modalités de données grâce à un apprentissage automatique avancé, ce travail vise à fournir aux cliniciens des outils interprétables pour naviguer dans les complexités de la gestion moderne du cancer, avec un accent particulier sur l'immuno-oncologie.

Nos recherches sur la prédiction de l'expression de CD73 dans les MHCR réséqués à l'aide d'images de TDM préopératoires sont pionnières dans le domaine du développement de biomarqueurs immunitaires non invasifs pour les MHCR. La pertinence pronostique indépendante démontrée du biomarqueur non invasif rad-CD73 proposé et sa corrélation avec l'expression de CD73 dérivée de l'IHC justifient des études prospectives à grande échelle pour établir son rôle en tant qu'outil non invasif permettant d'identifier les répondeurs potentiels aux inhibiteurs anti-CD73 ou aux inhibiteurs des récepteurs de l'adénosine et de guider de manière dynamique les thérapies adjuvantes et la fréquence des suivis après la chirurgie.

Dans le chapitre 6, le travail sur la classification des phénotypes immunologiques CD3/CMH-I dans les MHCR à l'aide d'images de TDM complète la première étude en mappant les caractéristiques d'imagerie avec l'infiltration immunitaire. Ils contribuent à faire progresser notre compréhension du microenvironnement tumoral en proposant un score immunitaire non invasif des MHCR. Une fois validé de manière prospective, ce score pourrait également contribuer à guider le suivi postopératoire des patients et à adapter les stratégies d'immunothérapie.

Nos travaux sur l'exploitation des images de TDM de pré-traitement pour prédire l'expression de PD-L1 dans le CPNPC ont démontré une performance robuste à différents stades de la maladie et avec différents régimes thérapeutiques. Notre pipeline, entraîné sur un ensemble de données multi-institutionnelles de patients atteints d'un CPNPC avancé pouvant être traité par immunothérapie, a montré une grande généralisation sur la cohorte de patients atteints d'une maladie à un stade limité traités par le nouveau régime thérapeutique combinant im-

munothérapie et chimiothérapie, actuellement considéré comme le traitement standard en situation néoadjuvante pour le CPNPC résécable. De plus, notre pipeline d'auto-apprentissage proposé, intégrant le modèle FT-Transformer, a surpassé les méthodes comparatives entraînées pour effectuer la même tâche. Cette approche généralisable répond au besoin critique de biomarqueurs fiables et non invasifs pour guider les décisions d'ICI chez les patients atteints d'un cancer du poumon.

Enfin, nous avons proposé un pipeline multi-omique pour prédire le risque de récurrence dans les deux ans suivant la résection chirurgicale d'une MHCR. En combinant des images de TDM préopératoires, des données de séquençage d'ARN et des variables cliniques, nous avons développé un modèle intégratif qui démontre une performance pronostique supérieure à celle du SRC conventionnel de Fong. Notre étude représente une approche pionnière dans le pronostic des MHCR, car les études précédentes se concentrent uniquement sur des approches à modalité unique. L'évaluation holistique de la biologie tumorale permet de mieux comprendre le risque de récurrence et peut orienter les décisions de suivi et de traitement adjuvant plus personnalisées.

Collectivement, ces études mettent en évidence les avantages de la combinaison de données multi-omiques, notamment des données d'imagerie, cliniques et génomiques. Notre thèse fait progresser le domaine vers une stratification plus précise des patients à l'aide de techniques avancées d'apprentissage automatique. La nature non invasive des biomarqueurs d'imagerie établit une base pour le suivi longitudinal des patients et l'évaluation du microenvironnement tumoral, palliant ainsi les limites et les risques procéduraux des biopsies répétées. De plus, les analyses d'interprétabilité associées à nos modèles et la reproductibilité des résultats de prédiction du PD-L1 à travers différentes caractéristiques des patients et différents contextes de traitement augmentent leur potentiel d'intégration clinique à grande échelle.

En conclusion, cette thèse a fait progresser le domaine de pronostication du cancer en proposant des approches innovantes non invasives pour la prédiction des biomarqueurs et un pipeline multimodal pour la stratification du risque de récurrence du MHCR. Lorsqu'ils seront validés de manière prospective, ces outils ont le potentiel de révolutionner la planification et le suivi des traitements dans les MHCR et le CPNPC, ouvrant la voie à des soins plus personnalisés et plus efficaces contre le cancer.

10.2 Limites

Malgré les contributions de cette thèse, certaines limites doivent être abordées dans les travaux futurs :

1. Nature rétrospective : toutes les études de cette thèse sont menées rétrospectivement. La biobanque des MHCR est maintenue de manière prospective en suivant les patients au-delà de la récurrence et jusqu'à leur décès pour les données clinico-pathologiques et d'imagerie dans le registre. Néanmoins, les données ont été analysées de manière rétrospective dans notre travail, ce qui limite les inférences causales sur les relations observées.
2. Ensembles de données monocentriques ou multicentriques limités : les études réalisées sur les MHCR étaient basées sur un ensemble de données monocentriques de patients recrutés au CHUM. En ce qui concerne l'étude menée sur le CPNPC, elle a intégré des patients de plusieurs centres, mais ils sont tous confinés au Québec, Canada. La limitation géographique et démographique peut avoir un impact sur la généralisabilité des modèles à d'autres régions et caractéristiques des patients.
3. Taille des échantillons : le nombre de patients inclus dans chacune des études était relativement limité. Néanmoins, les résultats prometteurs des études soulignent la nécessité de mener des études à grande échelle combinées à des approches avancées afin de respecter les réglementations en matière de confidentialité des données des patients.
4. Fluctuations temporelles des biomarqueurs : l'expression des biomarqueurs du cancer tels que PD-L1 et CD73 est susceptible de changer de manière dynamique au cours du traitement. Nous avons mitigé ce problème en appliquant des contraintes temporelles dans nos analyses. Pour la prédiction de l'expression de PD-L1, nous avons limité notre analyse aux TDM acquis dans un délai de trois mois avant le début du traitement. Pour la prédiction des marqueurs immunitaires du MHCR, nous avons utilisé le TDM le plus récent avant la chirurgie.
5. Complexité informatique : l'approche multi-omique, bien que puissante, introduit des complexités informatiques et analytiques afin d'intégrer efficacement diverses modalités et d'en tirer des informations biologiques pertinentes.

10.3 Améliorations futures

Les résultats de cette thèse suggèrent plusieurs recommandations pour les orientations de la recherche en matière de pronostic du cancer.

Recommandation 1. Des études multi-institutionnelles à grande échelle de nature prospective sont nécessaires pour valider les résultats prometteurs obtenus dans la thèse. Pour garantir le respect des réglementations en matière de confidentialité des données des patients tout en permettant des résultats solides et généralisables, des méthodologies avancées

pourraient être intégrées, telles que l'apprentissage fédéré.

Recommandation 2. Le pipeline multi-omique pourrait être amélioré en exploitant des modalités supplémentaires pour accroître sa précision et son utilité clinique. Par exemple, le statut mutationnel de gènes pronostiques connus tels que KRAS pourrait être inclus. L'expression d'autres biomarqueurs tels que TIGIT pourrait également être intégrée dans le pipeline. En outre, l'intégration d'autres modalités d'imagerie telles que les images histopathologiques ou la TEP pourrait également améliorer les performances de l'approche.

Recommandation 3. Des études longitudinales pourraient être réalisées en exploitant l'imagerie post-traitement pour suivre les changements temporels dans l'expression des biomarqueurs. En recueillant des données d'imagerie et potentiellement génomiques à plusieurs moments au cours du traitement, des modèles peuvent être formés pour prédire comment les patients répondront aux lignes de thérapie ultérieures ou aux changements de régimes de traitement.

Recommandation 4. Les pipelines proposés pourraient être adaptés à d'autres types de cancer, ce qui contribuerait à élargir leur impact. Les approches pan-cancer pourraient être exploitées pour révéler des profils pronostiques communs à toutes les tumeurs malignes. Par exemple, le travail de prédiction du PD-L1 pourrait être étendu à d'autres types de cancer pour lesquels une autorisation de la FDA a été obtenue pour guider l'ICI.

RÉFÉRENCES

- [1] I. Perrot, H.-A. Michaud, M. Giraudon-Paoli, S. Augier, A. Docquier, L. Gros, R. Courtois, C. Déjou, D. Jecko, O. Becquart, H. Rispaud-Blanc, L. Gauthier, B. Rossi, S. Chanteux, N. Gourdin, B. Amigues, A. Roussel, A. Bensussan, J.-F. Eliaou, J. Bastid, F. Romagné, Y. Morel, E. Narni-Mancinelli, E. Vivier, C. Paturel et N. Bonnefoy, “Blocking antibodies targeting the CD39/CD73 immunosuppressive pathway unleash immune responses in combination cancer therapies,” *Cell Reports*, vol. 27, n^o. 8, p. 2411–2425.e9, mai 2019.
- [2] H. Petrowsky, R. Fritsch, M. Guckenberger, M. L. De Oliveira, P. Dutkowski et P.-A. Clavien, “Modern therapeutic approaches for the treatment of malignant liver tumours,” *Nature Reviews Gastroenterology & Hepatology*, vol. 17, n^o. 12, p. 755–772, déc. 2020.
- [3] American Cancer Society, “Treatment choices for non-small cell lung cancer, by stage,” accessed : March 6, 2025. [En ligne]. Disponible : <https://www.cancer.org/cancer/types/lung-cancer/treating-non-small-cell/by-stage.html>
- [4] F. Alshohoumi, A. Al-Hamdani, R. Hedjam, A. AlAbdulsalam et A. Al Zaabi, “A review of radiomics in predicting therapeutic response in colorectal liver metastases : From traditional to artificial intelligence techniques,” *Healthcare*, vol. 10, n^o. 10, 2022. [En ligne]. Disponible : <https://www.mdpi.com/2227-9032/10/10/2075>
- [5] A. Kline, H. Wang, Y. Li, S. Dennis, M. Hutch, Z. Xu, F. Wang, F. Cheng et Y. Luo, “Multimodal machine learning in precision health : A scoping review,” *npj Digital Medicine*, vol. 5, n^o. 1, p. 171, nov. 2022.
- [6] Y. Pan, L. Shi, Y. Liu, J. cheng Chen et J. Qiu, “Multi-omics models for predicting prognosis in non-small cell lung cancer patients following chemotherapy and radiotherapy : A multi-center study,” *Radiotherapy and Oncology*, vol. 204, p. 110715, 2025. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0167814025000106>
- [7] S. Pai, D. Bontempi, I. Hadzic, V. Prudente, M. Sokač, T. L. Chaunzwa, S. Bernatz, A. Hosny, R. H. Mak, N. J. Birkbak et H. J. W. L. Aerts, “Foundation model for cancer imaging biomarkers,” *Nature Machine Intelligence*, vol. 6, n^o. 3, p. 354–367, mars 2024.
- [8] J. Yoon *et al.*, “Gain : Missing data imputation using generative adversarial nets,” 2018.
- [9] C. C. Society, “Colorectal cancer statistics,” 2024, accessed : March 23, 2025. [En

- ligne]. Disponible : <https://cancer.ca/en/cancer-information/cancer-types/colorectal/statistics>
- [10] Canadian Cancer Society, “Lung and bronchus cancer statistics,” accessed : March 6, 2025. [En ligne]. Disponible : <https://cancer.ca/en/cancer-information/cancer-types/lung/statistics>
 - [11] World Health Organization (WHO), “Colorectal cancer,” 2024, accessed : March 23, 2025; Last updated : March 2024. [En ligne]. Disponible : <https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer>
 - [12] M. Van den Eynde, B. Mlecnik, G. Bindea *et al.*, “The link between the multiverse of immune microenvironments in metastases and the survival of colorectal cancer patients,” *Cancer Cell*, vol. 34, n^o. 6, p. 1012–1026.e3, déc. 2018.
 - [13] X. Yu, L. Zhu, J. Liu, M. Xie, J. Chen et J. Li, “Emerging role of immunotherapy for colorectal cancer with liver metastasis,” *Onco Targets Ther*, vol. 13, p. 11 645–11 658, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/33223838>
 - [14] Y. Wang, K. Kondrat, J. Adhikari, Q. Nguyen, Q. Yu et D. Uprety, “Survival trends among patients with metastatic non-small cell lung cancer before and after the approval of immunotherapy in the united states : A surveillance, epidemiology, and end results database-based study,” *Cancer*, vol. 131, n^o. 1, p. e35476, juill. 2024.
 - [15] A. Jones, A. Findlay, S. R. Knight, J. Rees, D. O’Reilly, R. P. Jones et S. Pathak, “Follow up after surgery for colorectal liver metastases : A systematic review,” *Eur J Surg Oncol*, vol. 49, n^o. 12, p. 107103, oct. 2023.
 - [16] H. Chen, Y. Xu, J. Liu, S. Yang, H. Jiang et Z. Chen, “Analysis of the association between immune-related adverse events and the effectiveness in patients with advanced non-small-cell-lung cancer,” *Discov Oncol*, vol. 15, n^o. 1, p. 534, oct. 2024.
 - [17] M. Donadon, N. Cortese, F. Marchesi, M. Cimino, A. Mantovani et G. Torzilli, “Hepatobiliary surgeons meet immunologists : the case of colorectal liver metastases patients,” *Hepatobiliary Surgery and Nutrition*, vol. 8, n^o. 4, 2019. [En ligne]. Disponible : <https://hbsn.amegroups.org/article/view/24636>
 - [18] L. Rubbia-Brandt, E. Giostra, C. Brezault, A. D. Roth, A. Andres, V. Audard, P. Sartoretti, B. Dousset, P. E. Majno, O. Soubrane, S. Chaussade, G. Mentha et B. Terris, “Importance of histological tumor response assessment in predicting the outcome in patients with colorectal liver metastases treated with neo-adjuvant chemotherapy followed by liver surgery,” *Ann Oncol*, vol. 18, n^o. 2, p. 299–304, 2007.
 - [19] D. G. Blazer, 3rd, Y. Kishi, D. M. Maru, S. Kopetz, Y. S. Chun, M. J. Overman, D. Fogelman, C. Eng, D. Z. Chang, H. Wang, D. Zorzi, D. Ribero, L. M. Ellis, K. Y.

- Glover, R. A. Wolff, S. A. Curley, E. K. Abdalla et J.-N. Vauthey, “Pathologic response to preoperative chemotherapy : a new outcome end point after resection of hepatic colorectal metastases,” *J Clin Oncol*, vol. 26, n°. 33, p. 5344–5351, oct. 2008.
- [20] Y. Fong, J. Fortner, R. L. Sun, M. F. Brennan et L. H. Blumgart, “Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer : analysis of 1001 consecutive cases,” *Ann Surg*, vol. 230, n°. 3, p. 309–18 ; discussion 318–21, 1999.
- [21] D. de Jong, J. P. Das, H. Ma, J. Pailey Valiplackal, C. Prendergast, T. Roa, B. Braumuller, A. Deng, L. Dercle, R. Yeh, M. M. Salvatore et K. M. Capaccione, “Novel targets, novel treatments : The changing landscape of Non-Small cell lung cancer,” *Cancers (Basel)*, vol. 15, n°. 10, mai 2023.
- [22] D. J. Myers et J. M. Wallen, *Lung Adenocarcinoma*. StatPearls Publishing, juin 2023. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/books/NBK519578/>
- [23] National Cancer Institute, “Non-small cell lung cancer treatment (pdq®)–health professional version,” accessed : March 6, 2025. [En ligne]. Disponible : <https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq>
- [24] L.-H. Xu, S.-J. Cai, G.-X. Cai et W.-J. Peng, “Imaging diagnosis of colorectal liver metastases,” *World J Gastroenterol*, vol. 17, n°. 42, p. 4654–4659, nov. 2011.
- [25] M. Renzulli, A. Clemente, A. M. Ierardi, I. Pettinari, F. Tovoli, S. Brocchi, G. Peta, S. Cappabianca, G. Carrafiello et R. Golfieri, “Imaging of colorectal liver metastases : New developments and pending issues,” *Cancers (Basel)*, vol. 12, n°. 1, janv. 2020.
- [26] C. d. Margerie-Mellon, *Imagerie médicale pratique*, 3^e éd. Paris : Vuibert, 2019.
- [27] P. Saukko et S. Pollak, *Histopathology*. Waltham : Academic Press, 2013, p. 210–216. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/B9780123821652001860>
- [28] N. Knijn, J. A. M. de Ridder, C. J. A. Punt, J. H. W. de Wilt et I. D. Nagtegaal, “Histopathological evaluation of resected colorectal cancer liver metastases : what should be done ?” *Histopathology*, vol. 63, n°. 2, p. 149–156, juin 2013.
- [29] P. Baldin, M. Van den Eynde, B. Mlecnik, G. Bindea, G. Beniuga, J. Carrasco, N. Haicheur, F. Marliot, L. Lafontaine, T. Fredriksen, N. Lanthier, C. Hubert, B. Navez, N. Huyghe, F. Pages, A. Jouret-Mourin, J. Galon et M. Komuta, “Prognostic assessment of resected colorectal liver metastases integrating pathological features, RAS mutation and immunoscore,” *J Pathol Clin Res*, vol. 7, n°. 1, p. 27–41, 2021. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32902189>

- [30] N. M. T. Jawhar, “Tissue microarray : A rapidly evolving diagnostic and research tool,” *Ann Saudi Med*, vol. 29, n^o. 2, p. 123–127, mars 2009.
- [31] S. Ergin, N. Kherad et M. Alagoz, “RNA sequencing and its applications in cancer and rare diseases,” *Mol Biol Rep*, vol. 49, n^o. 3, p. 2325–2333, janv. 2022.
- [32] Y. Zhang, G. Parmigiani et W. E. Johnson, “ComBat-seq : batch effect adjustment for RNA-seq count data,” *NAR Genom. Bioinform.*, vol. 2, n^o. 3, p. lqaa078, 09 2020.
- [33] D. Deshpande, K. Chhugani, Y. Chang, A. Karlsberg, C. Loeffler, J. Zhang, A. Muszyńska, V. Munteanu, H. Yang, J. Rotman, L. Tao, B. Balliu, E. Tseng, E. Eskin, F. Zhao, P. Mohammadi, P. P. Łabaj et S. Mangul, “Rna-seq data science : From raw data to effective interpretation,” *Frontiers in Genetics*, vol. 14, 2023. [En ligne]. Disponible : <https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2023.997383>
- [34] S. A. Byron, K. R. Van Keuren-Jensen, D. M. Engelthaler, J. D. Carpten et D. W. Craig, “Translating RNA sequencing into clinical diagnostics : opportunities and challenges,” *Nature Reviews Genetics*, vol. 17, n^o. 5, p. 257–271, mai 2016.
- [35] A. Adil, V. Kumar, A. T. Jan et M. Asger, “Single-cell transcriptomics : Current methods and challenges in data acquisition and analysis,” *Frontiers in Neuroscience*, vol. 15, 2021. [En ligne]. Disponible : <https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2021.591122>
- [36] G. Rossi, A. Russo, M. Tagliamento, A. Tuzi, O. Nigro, G. Vallome, C. Sini, M. Grassi, M. G. Dal Bello, S. Coco, L. Longo, L. Zullo, E. T. Tanda, C. Dellepiane, P. Pronzato et C. Genova, “Precision medicine for NSCLC in the era of immunotherapy : New biomarkers to select the most suitable treatment or the most suitable patient,” *Cancers*, vol. 12, n^o. 5, 2020. [En ligne]. Disponible : <https://www.mdpi.com/2072-6694/12/5/1125>
- [37] W. You, B. Shang, J. Sun, X. Liu, L. Su et S. Jiang, “Mechanistic insight of predictive biomarkers for antitumor PD-1/PD-L1 blockade : A paradigm shift towards immunome evaluation (review),” *Oncol Rep*, vol. 44, n^o. 2, p. 424–437, août 2020.
- [38] T. Donnem, T. Kilvaer, S. Andersen, E. Richardsen, E. Paulsen, S. Hald, S. Al-Saad, O. Brustugun, A. Helland, M. Lund-Iversen, S. Solberg, B. Gronberg, S. Wahl, L. Helgeland, O. Fløtten, M. Pohl, K. Al-Shibli, T. Sandanger, F. Pezzella, L. Busund et R. Bremnes, “Strategies for clinical implementation of tnm-immunoscore in resected nonsmall-cell lung cancer,” *Annals of Oncology*, vol. 27, n^o. 2, p. 225–232, 2016. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0923753419355589>

- [39] Y. Geng, Y. Shao, W. He, W. Hu, Y. Xu, J. Chen, C. Wu et J. Jiang, “Prognostic role of tumor-infiltrating lymphocytes in lung cancer : a meta-analysis,” *Cellular Physiology and Biochemistry*, vol. 37, n^o. 4, p. 1560–1571, 10 2015. [En ligne]. Disponible : <https://doi.org/10.1159/000438523>
- [40] F. Pagès, B. Mlecnik, F. Marliot, G. Bindea, F.-S. Ou, C. Bifulco, A. Lugli, I. Zlobec, T. T. Rau, M. D. Berger, I. D. Nagtegaal, E. Vink-Börger, A. Hartmann, C. Gelpert, J. Kolwelter, S. Merkel, R. Grützmann, M. Van den Eynde, A. Jouret-Mourin, A. Kartheuser, D. Léonard, C. Remue, J. Y. Wang, P. Bavi, M. H. A. Roehrl, P. S. Ohashi, L. T. Nguyen, S. Han, H. L. MacGregor, S. Hafezi-Bakhtiari, B. G. Wouters, G. V. Masucci, E. K. Andersson, E. Zavadova, M. Vocka, J. Spacek, L. Petruzella, B. Konopasek, P. Dundr, H. Skalova, K. Nemejcova, G. Botti, F. Tatangelo, P. Delrio, G. Ciliberto, M. Maio, L. Laghi, F. Grizzi, T. Fredriksen, B. Buttard, M. Angelova, A. Vasaturo, P. Maby, S. E. Church, H. K. Angell, L. Lafontaine, D. Bruni, C. El Sissy, N. Haicheur, A. Kirilovsky, A. Berger, C. Lagorce, J. P. Meyers, C. Paustian, Z. Feng, C. Ballesteros-Merino, J. Dijkstra, C. van de Water, S. van Lent-van Vliet, N. Knijn, A.-M. Muşină, D.-V. Scripcariu, B. Popivanova, M. Xu, T. Fujita, S. Hazama, N. Suzuki, H. Nagano, K. Okuno, T. Torigoe, N. Sato, T. Furuhashi, I. Takemasa, K. Itoh, P. S. Patel, H. H. Vora, B. Shah, J. B. Patel, K. N. Rajvik, S. J. Pandya, S. N. Shukla, Y. Wang, G. Zhang, Y. Kawakami, F. M. Marincola, P. A. Ascierto, D. J. Sargent, B. A. Fox et J. Galon, “International validation of the consensus immunescore for the classification of colon cancer : a prognostic and accuracy study,” *Lancet*, vol. 391, n^o. 10135, p. 2128–2139, mai 2018.
- [41] B. Mlecnik, M. Van den Eynde, G. Bindea, S. E. Church, A. Vasaturo, T. Fredriksen, L. Lafontaine, N. Haicheur, F. Marliot, D. Debetancourt, G. Pairet, A. Jouret-Mourin, J.-F. Gigot, C. Hubert, E. Danse, C. Dragean, J. Carrasco, Y. Humblet, V. Valge-Archer, A. Berger, F. Pagès, J.-P. Machiels et J. Galon, “Comprehensive intrametastatic immune quantification and major impact of immunescore on survival,” *JNCI : Journal of the National Cancer Institute*, vol. 110, n^o. 1, p. 97–108, 08 2017. [En ligne]. Disponible : <https://doi.org/10.1093/jnci/djx123>
- [42] S. Turcotte, S. C. Katz, J. Shia *et al.*, “Tumor MHC class I expression improves the prognostic value of t-cell density in resected colorectal liver metastases,” *Cancer Immunol Res*, vol. 2, n^o. 6, p. 530–537, févr. 2014.
- [43] M. V. d. E. Bernhard Mlecnik, Gabriela Bindea et J. Galon, “Metastasis immune-based scores predict patient survival,” *OncoImmunology*, vol. 9, n^o. 1, p. 1806000, 2020.
- [44] N. Messaoudi, I. Cousineau, E. Arslanian, D. Henault, D. Stephen, F. Vandembroucke-Menu, M. Dagenais, R. Letourneau, M. Plasse, A. Roy, R. Lapointe, D. Ysebaert,

- D. Trudel, G. Soucy, J. Stagg et S. Turcotte, “Prognostic value of cd73 expression in resected colorectal cancer liver metastasis,” *Oncoimmunology*, vol. 9, n° 1, p. 1746138, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32363113>
- [45] S. Chen, D. A. Wainwright, J. D. Wu, Y. Wan, D. E. Matei, Y. Zhang et B. Zhang, “CD73 : an emerging checkpoint for cancer immunotherapy,” *Immunotherapy*, vol. 11, n° 11, p. 983–997, juin 2019.
- [46] J. Linden et C. Cekic, “Regulation of lymphocyte function by adenosine,” *Arterioscler. Thromb. Vasc. Biol.*, vol. 32, n° 9, p. 2097–2103, sept. 2012.
- [47] X. R. Wu, X. S. He, Y. F. Chen, R. X. Yuan, Y. Zeng, L. Lian, Y. F. Zou, N. Lan, X. J. Wu et P. Lan, “High expression of cd73 as a poor prognostic biomarker in human colorectal cancer,” *J Surg Oncol*, vol. 106, n° 2, p. 130–7, 2012.
- [48] Y. Han, D. Liu et L. Li, “PD-1/PD-L1 pathway : current researches in cancer,” *Am J Cancer Res*, vol. 10, n° 3, p. 727–742, mars 2020.
- [49] A. P. George, T. M. Kuzel, Y. Zhang et B. Zhang, “The discovery of biomarkers in cancer immunotherapy,” *Comput Struct Biotechnol J*, vol. 17, p. 484–497, avr. 2019.
- [50] L. Wang, Y. Hu, S. Wang, J. Shen et X. Wang, “Biomarkers of immunotherapy in non-small cell lung cancer,” *Oncol Lett*, vol. 20, n° 5, p. 139, août 2020.
- [51] E. Van Cutsem, A. Cervantes, R. Adam, A. Sobrero, J. H. Van Krieken, D. Aderka, E. Aranda Aguilar, A. Bardelli, A. Benson, G. Bodoky, F. Ciardiello, A. D’Hoore, E. Diaz-Rubio, J.-Y. Douillard, M. Ducreux, A. Falcone, A. Grothey, T. Gruenberger, K. Haustermans, V. Heinemann, P. Hoff, C.-H. Köhne, R. Labianca, P. Laurent-Puig, B. Ma, T. Maughan, K. Muro, N. Normanno, P. Österlund, W. J. G. Oyen, D. Papamichael, G. Pentheroudakis, P. Pfeiffer, T. J. Price, C. Punt, J. Ricke, A. Roth, R. Salazar, W. Scheithauer, H. J. Schmoll, J. Tabernero, J. Taïeb, S. Tejpar, H. Wasan, T. Yoshino, A. Zaanani et D. Arnold, “ESMO consensus guidelines for the management of patients with metastatic colorectal cancer,” *Ann Oncol*, vol. 27, n° 8, p. 1386–1422, juill. 2016.
- [52] Y. Dai, W. Zhao, L. Yue, X. Dai, D. Rong, F. Wu, J. Gu et X. Qian, “Perspectives on immunotherapy of metastatic colorectal cancer,” *Front Oncol*, vol. 11, p. 659964, 2021.
- [53] Canadian Cancer Society, “Treatments for non-small cell lung cancer,” accessed : March 6, 2025. [En ligne]. Disponible : <https://cancer.ca/en/cancer-information/cancer-types/lung/treatment>
- [54] I. Houda, C. Dickhoff, C. A. Uyl-de Groot, N. Reguart, M. Provencio, A. Levy, R. Dziadziuszko, C. Pompili, M. Di Maio, M. Thomas, A. Brunelli, S. Popat, S. Senan et I. Bahce, “New systemic treatment paradigms in resectable non-small cell lung cancer and variations in patient access across europe,” *The Lancet Regional Health – Europe*, vol. 38, mars 2024.

- [55] N. H. Hanna, B. J. Schneider, S. Temin, S. Baker, J. Brahmer, P. M. Ellis, L. E. Gaspar, R. Y. Haddad, P. J. Hesketh, D. Jain, I. Jaiyesimi, D. H. Johnson, N. B. Leighl, T. Phillips, G. J. Riely, A. G. Robinson, R. Rosell, J. H. Schiller, N. Singh, D. R. Spigel, J. O. Stabler, J. Tashbar et G. Masters, “Therapy for stage iv non-small-cell lung cancer without driver alterations : Asco and oh (cco) joint guideline update,” *Journal of Clinical Oncology*, vol. 38, n°. 14, p. 1608–1632, 2020, pMID : 31990617. [En ligne]. Disponible : <https://ascopubs.org/doi/abs/10.1200/JCO.19.03022>
- [56] I. Houda, C. Dickhoff, C. A. Uyl-de Groot, R. A. M. Damhuis, N. Reguart, M. Provencio, A. Levy, R. Dziadziuszko, C. Pompili, M. Di Maio, M. Thomas, A. Brunelli, S. Popat, S. Senan et I. Bahce, “Challenges and controversies in resectable non-small cell lung cancer : a clinician’s perspective,” *The Lancet Regional Health – Europe*, vol. 38, mars 2024.
- [57] A. Sridhar, H. Khan, B. Yohannan, K. H. Chan, N. Kataria et S. H. Jafri, “A review of the current approach and treatment landscape for stage iii non-small cell lung cancer,” *Journal of Clinical Medicine*, vol. 13, n°. 9, 2024. [En ligne]. Disponible : <https://www.mdpi.com/2077-0383/13/9/2633>
- [58] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. van Stiphout, P. Granton, C. M. Zegers, R. Gillies, R. Boellard, A. Dekker et H. J. Aerts, “Radiomics : extracting more information from medical images using advanced feature analysis,” *Eur J Cancer*, vol. 48, n°. 4, p. 441–6, 2012. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/22257792>
- [59] E. J. Limkin, R. Sun, L. Derle, E. I. Zacharaki, C. Robert, S. Reuze, A. Schernberg, N. Paragios, E. Deutsch et C. Ferte, “Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology,” *Ann Oncol*, vol. 28, n°. 6, p. 1191–1206, 2017. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/28168275>
- [60] M. Porcu, C. Solinas, L. Mannelli, G. Micheletti, M. Lambertini, K. Willard-Gallo, E. Neri, A. E. Flanders et L. Saba, “Radiomics and “radi-... omics” in cancer immunotherapy : a guide for clinicians,” *Critical Reviews in Oncology/Hematology*, vol. 154, p. 103068, 2020. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1040842820302043>
- [61] S. S. F. Yip et H. J. W. L. Aerts, “Applications and limitations of radiomics,” *Phys Med Biol*, vol. 61, n°. 13, p. R150–66, juin 2016.
- [62] A. K. Jha, S. Mithun, V. Jaiswar, U. B. Sherkhane, N. C. Purandare, K. Prabhash, V. Rangarajan, A. Dekker, L. Wee et A. Traverso, “Repeatability and reproducibility

- study of radiomic features on a phantom and human cohort,” *Scientific Reports*, vol. 11, n° 1, p. 2055, janv. 2021.
- [63] S. Badillo, B. Banfai, F. Birzele, I. I. Davydov, L. Hutchinson, T. Kam-Thong, J. Siebourg-Polster, B. Steiert et J. D. Zhang, “An introduction to machine learning,” *Clin Pharmacol Ther*, vol. 107, n° 4, p. 871–885, mars 2020.
- [64] F. P. Romero, A. Diler, G. Bisson-Gregoire, S. Turcotte, R. Lapointe, F. Vandenbroucke-Menu, A. Tang et S. Kadoury, “End-to-end discriminative deep network for liver lesion classification,” dans *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019, p. 1243–1246.
- [65] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens et Z. Wojna, “Rethinking the inception architecture for computer vision,” 2015. [En ligne]. Disponible : <https://arxiv.org/abs/1512.00567>
- [66] Y. Li, J. Gong, X. Shen, M. Li, H. Zhang, F. Feng et T. Tong, “Assessment of primary colorectal cancer CT radiomics to predict metachronous liver metastasis,” *Frontiers in Oncology*, vol. 12, 2022. [En ligne]. Disponible : <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2022.861892>
- [67] S. Grosu, P. Wesp, A. Graser, S. Maurus, C. Schulz, T. Knösel, C. C. Cyran, J. Ricke, M. Ingrisch et P. M. Kazmierczak, “Machine learning-based differentiation of benign and premalignant colorectal polyps detected with CT colonography in an asymptomatic screening population : A proof-of-concept study,” *Radiology*, vol. 299, n° 2, p. 326–335, 2021, pMID : 33620287. [En ligne]. Disponible : <https://doi.org/10.1148/radiol.2021202363>
- [68] C. Jacob, C. Gopakumar et F. Nazarudeen, “Optimized radiomics-based machine learning approach for lung cancer subtype classification,” *Biomedical Engineering : Applications, Basis and Communications*, vol. 35, n° 05, p. 2350023, 2023. [En ligne]. Disponible : <https://doi.org/10.4015/S1016237223500230>
- [69] A. Kapil, A. Meier, A. Zuraw, K. E. Steele, M. C. Rebelatto, G. Schmidt et N. Brieu, “Deep semi supervised generative learning for automated tumor proportion scoring on NSCLC tissue needle biopsies,” *Scientific Reports*, vol. 8, n° 1, p. 17343, nov. 2018.
- [70] D. Krijgsman, M. B. van Leeuwen, J. van der Ven, V. Almeida, R. Vlutters, D. Halter, P. J. K. Kuppen, C. J. H. van de Velde et R. Wimberger-Friedl, “Quantitative whole slide assessment of tumor-infiltrating cd8-positive lymphocytes in er-positive breast cancer in relation to clinical outcome,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, n° 2, p. 381–392, 2021.
- [71] E. Vorontsov, G. Chartrand, A. Tang, C. Pal et S. Kadoury, “Liver lesion segmentation informed by joint liver segmentation,” *2018 IEEE 15th International Symposium on*

- Biomedical Imaging (ISBI 2018)*, p. 1332–1335, 2018.
- [72] M. A. de Boisredon d’Assier, E. Vorontsov et S. Kadoury, “M-genseg : Domain adaptation for target modality tumor segmentation with annotation-efficient supervision,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2212.07276>
 - [73] S. Yoon, Y. J. Kim, J. S. Jeon, S. J. Ahn et S. J. Choi, “Radiomics and machine learning analysis of liver magnetic resonance imaging for prediction and early detection of tumor response in colorectal liver metastases,” *Korean J Clin Oncol*, vol. 20, n^o. 1, p. 27–35, juin 2024.
 - [74] Y. Chen, T. Lu, Y. Zhang, H. Li, J. Xu et M. Li, “Baseline hepatobiliary MRI for predicting chemotherapeutic response and prognosis in initially unresectable colorectal cancer liver metastases,” *Abdom Radiol (NY)*, vol. 49, n^o. 8, p. 2585–2594, juill. 2024.
 - [75] M. Taghavi, M. Maas, F. C. R. Staal, R. G. H. Beets-Tan et S. Benson, “CNN-based tumor progression prediction after thermal ablation with CT imaging,” dans *Medical Imaging 2022 : Computer-Aided Diagnosis*, K. Drukker et K. M. Iftikharuddin, édit., vol. 12033, International Society for Optics and Photonics. SPIE, 2022, p. 1203335. [En ligne]. Disponible : <https://doi.org/10.1117/12.2611516>
 - [76] C.-H. Li, D. Cai, M.-E. Zhong, M.-Y. Lv, Z.-P. Huang, Q. Zhu, C. Hu, H. Qi, X. Wu et F. Gao, “Multi-size deep learning based preoperative computed tomography signature for prognosis prediction of colorectal cancer,” *Frontiers in Genetics*, vol. 13, 2022. [En ligne]. Disponible : <https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2022.880093>
 - [77] M. Li, J. Gong, Y. Bao, D. Huang, J. Peng et T. Tong, “Special issue “the advance of solid tumor research in china” : Prognosis prediction for stage II colorectal cancer by fusing computed tomography radiomics and deep-learning features of primary lesions and peripheral lymph nodes,” *Int J Cancer*, vol. 152, n^o. 1, p. 31–41, mai 2022.
 - [78] A. Maaref, F. P. Romero, E. Montagnon, M. Cerny, B. Nguyen, F. Vandenbroucke, G. Soucy, S. Turcotte, A. Tang et S. Kadoury, “Predicting the response to FOLFOX-based chemotherapy regimen from untreated liver metastases on baseline CT : A deep neural network approach,” *J. Digit. Imaging*, vol. 33, n^o. 4, p. 937–945, août 2020.
 - [79] V. Giannini, A. Defeudis, S. Rosati, G. Cappello, S. Mazzetti, J. Panic, D. Regge et G. Balestra, “An innovative radiomics approach to predict response to chemotherapy of liver metastases based on CT images,” dans *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, p. 1339–1342.
 - [80] Y. Chen, T. Lu, Y. Zhang, H. Li, J. Xu et M. Li, “Baseline hepatobiliary MRI for predicting chemotherapeutic response and prognosis in initially unresectable colorectal cancer liver metastases,” *Abdominal Radiology*, vol. 49, n^o. 8, p. 2585–2594, août 2024.

- [81] W. Chen, X. Hou, Y. Hu, G. Huang, X. Ye et S. Nie, “A deep learning- and CT image-based prognostic model for the prediction of survival in non-small cell lung cancer,” *Med Phys*, vol. 48, n^o. 12, p. 7946–7958, oct. 2021.
- [82] C. Li, Z. Zhou, L. Hou, K. Hu, Z. Wu, Y. Xie, J. Ouyang et X. Cai, “A novel machine learning model for efficacy prediction of immunotherapy-chemotherapy in NSCLC based on CT radiomics,” *Computers in Biology and Medicine*, vol. 178, p. 108638, 2024. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0010482524007236>
- [83] J. Yu, Y. Velichko, H. Kim, N. Gennaro, M. Soliman, L. Kim, Y. Oh, T. A. Djunadi, J. Lee, L. I.-Y. Chung, S. M. Yoon, Z. Shah, W. J. Yang, H. S. Kim, Y. Lee, S. Lee, D. Kang, R. Agrawal, P. Aouad et Y. K. Chae, “Machine learning models based on radiomics features to predict treatment response, biomarker status, and bone metastasis in patients with non-small cell lung cancer (nsccl) treated with immune checkpoint inhibitors (icis).” *Journal of Clinical Oncology*, vol. 41, n^o. 16_suppl, p. e21217–e21217, 2023, pMID : [En ligne]. Disponible : https://ascopubs.org/doi/abs/10.1200/JCO.2023.41.16_suppl.e21217
- [84] K. H. Lee, J. H. Lee, S. Park, Y. K. Jeon, D. H. Chung, Y. T. Kim, J. M. Goo et H. Kim, “Computed tomography–based prognostication in lung adenocarcinomas through histopathological feature learning : A retrospective multicenter study,” *Annals of the American Thoracic Society*, vol. 20, n^o. 7, p. 1020–1028, 2023, pMID : 37075305. [En ligne]. Disponible : <https://doi.org/10.1513/AnnalsATS.202210-895OC>
- [85] P. Tian, B. He, W. Mu, K. Liu, L. Liu, H. Zeng, Y. Liu, L. Jiang, P. Zhou, Z. Huang, D. Dong et W. Li, “Assessing pd-l1 expression in non-small cell lung cancer and predicting responses to immune checkpoint inhibitors using deep learning on computed tomography images,” *Theranostics*, vol. 11, p. 2098–2107, 2021. [En ligne]. Disponible : <https://www.thno.org/v11p2098.htm>
- [86] Q. Wen, Z. Yang, H. Dai, A. Feng et Q. Li, “Radiomics study for predicting the expression of PD-L1 and tumor mutation burden in Non-Small cell lung cancer based on CT images and clinicopathological features,” *Front Oncol*, vol. 11, p. 620246, août 2021.
- [87] J. Yoon, Y. J. Suh, K. Han, H. Cho, H.-J. Lee, J. Hur et B. W. Choi, “Utility of CT radiomics for prediction of PD-L1 expression in advanced lung adenocarcinomas,” *Thoracic Cancer*, vol. 11, n^o. 4, p. 993–1004, 2020. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1111/1759-7714.13352>
- [88] S. Yolchuyeva, E. Giacomazzi, M. Tonneau, F. Lamaze, M. Orain, F. Coulombe, J. Malo, W. Belkaid, B. Routy, P. Joubert et V. S. K. Manem, “Radiomics approaches to predict PD-L1 and PFS in advanced non-small cell lung patients treated with im-

- munotherapy : a multi-institutional study,” *Scientific Reports*, vol. 13, n^o. 1, p. 11065, juill. 2023.
- [89] Y. Iwatate, I. Hoshino, H. Yokota, F. Ishige, M. Itami, Y. Mori, S. Chiba, H. Arimitsu, H. Yanagibashi, H. Nagase et W. Takayama, “Radiogenomics for predicting p53 status, PD-L1 expression, and prognosis with machine learning in pancreatic cancer,” *Br J Cancer*, vol. 123, n^o. 8, p. 1253–1261, juill. 2020.
- [90] M. A. Hussain, G. Hamarneh et R. Garbi, “Noninvasive determination of gene mutations in clear cell renal cell carcinoma using multiple instance decisions aggregated cnn,” dans *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López et G. Fichtinger, édit. Cham : Springer International Publishing, 2018, p. 657–665.
- [91] G. Liu, Z. Xu, Y. Ge, B. Jiang, H. Groen, R. Vliegenthart et X. Xie, “3d radiomics predicts egfr mutation, exon-19 deletion and exon-21 l858r mutation in lung adenocarcinoma,” *Transl Lung Cancer Res*, vol. 9, n^o. 4, p. 1212–1224, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32953499>
- [92] L. Lu, S. H. Sun, H. Yang, L. E, P. Guo, L. H. Schwartz et B. Zhao, “Radiomics prediction of egfr status in lung cancer-our experience in using multiple feature extractors and the cancer imaging archive data,” *Tomography*, vol. 6, n^o. 2, p. 223–230, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32548300>
- [93] D. Mei, Y. Luo, Y. Wang et J. Gong, “CT texture analysis of lung adenocarcinoma : can radiomic features be surrogate biomarkers for EGFR mutation statuses,” *Cancer Imaging*, vol. 18, n^o. 1, p. 52, 2018. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/30547844>
- [94] K. Ninomiya, H. Arimura, W. Y. Chan, K. Tanaka, S. Mizuno, N. F. Muhammad Gowdh, N. A. Yaakup, C. K. Liam, C. S. Chai et K. H. Ng, “Robust radiogenomics approach to the identification of egfr mutations among patients with NSCLC from three different countries using topologically invariant betti numbers,” *PLoS One*, vol. 16, n^o. 1, p. e0244354, 2021. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/33428651>
- [95] G. Rossi, E. Barabino, A. Fedeli, G. Ficarra, S. Coco, A. Russo, V. Adamo, F. Buemi, L. Zullo, M. Dono, G. De Luca, L. Longo, M. G. Dal Bello, M. Tagliamento, A. Alama, G. Cittadini, P. Pronzato et C. Genova, “Radiomic detection of egfr mutations in nsclc,” *Cancer Res*, vol. 81, n^o. 3, p. 724–731, 2021. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/33148663>
- [96] I. Shiri, H. Maleki, G. Hajianfar, H. Abdollahi, S. Ashrafinia, M. Hatt, H. Zaidi, M. Oveisi et A. Rahmim, “Next-generation radiogenomics sequencing for prediction

- of egfr and kras mutation status in NSCLC patients using multimodal imaging and machine learning algorithms,” *Mol Imaging Biol*, vol. 22, n°. 4, p. 1132–1148, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32185618>
- [97] V. González-Castro, E. Cernadas, E. Huelga, M. Fernández-Delgado, J. Porto, J. R. Antunez et M. Souto-Bayarri, “Ct radiomics in colorectal cancer : Detection of kras mutation using texture analysis and machine learning,” *Applied Sciences*, vol. 10, n°. 18, 2020.
- [98] K. He, X. Liu, M. Li, X. Li, H. Yang et H. Zhang, “Noninvasive kras mutation estimation in colorectal cancer using a deep learning method based on CT imaging,” *BMC Med Imaging*, vol. 20, n°. 1, p. 59, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32487083>
- [99] X. Wu, Y. Li, X. Chen, Y. Huang, L. He, K. Zhao, X. Huang, W. Zhang, Y. Huang, Y. Li, M. Dong, J. Huang, T. Xia, C. Liang et Z. Liu, “Deep learning features improve the performance of a radiomics signature for predicting kras status in patients with colorectal cancer,” *Acad Radiol*, vol. 27, n°. 11, p. e254–e262, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/31982342>
- [100] R. Shi, W. Chen, B. Yang, J. Qu, Y. Cheng, Z. Zhu, Y. Gao, Q. Wang, Y. Liu, Z. Li et X. Qu, “Prediction of kras, nras and braf status in colorectal cancer patients with liver metastasis using a deep artificial neural network based on radiomics and semantic features,” *American journal of cancer research*, vol. 10, n°. 12, p. 4513–4526, 2020. [En ligne]. Disponible : <https://pubmed.ncbi.nlm.nih.gov/33415015https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7783758/>
- [101] W. Cao, H. Hu, J. Guo, Q. Qin, Y. Lian, J. Li, Q. Wu, J. Chen, X. Wang et Y. Deng, “CT-based deep learning model for the prediction of DNA mismatch repair deficient colorectal cancer : a diagnostic study,” *Journal of Translational Medicine*, vol. 21, n°. 1, p. 214, mars 2023.
- [102] Q. Li, X. Long, Y. Lin, R. Liang, Y. Li et L. Ge, “Computed tomography radiomics signature via machine learning predicts rrm2 and overall survival in hepatocellular carcinoma,” *Journal of Gastrointestinal Oncology*, vol. 14, n°. 3, 2023. [En ligne]. Disponible : <https://jgo.amegroups.org/article/view/76278>
- [103] H. Liao, Z. Zhang, J. Chen, M. Liao, L. Xu, Z. Wu, K. Yuan, B. Song et Y. Zeng, “Preoperative radiomic approach to evaluate tumor-infiltrating cd8+ t cells in hepatocellular carcinoma patients using contrast-enhanced computed tomography,” *Annals of Surgical Oncology*, vol. 26, n°. 13, p. 4537–4547, 2019. [En ligne]. Disponible : <https://doi.org/10.1245/s10434-019-07815-9>

- [104] R. Sun, E. J. Limkin, M. Vakalopoulou, L. Dercle, S. Champiat, S. R. Han, L. Verlingue, D. Brandao, A. Lancia, S. Ammari, A. Hollebecque, J. Y. Scoazec, A. Marabelle, C. Massard, J. C. Soria, C. Robert, N. Paragios, E. Deutsch et C. F  rt  , “A radiomics approach to assess tumour-infiltrating cd8 cells and response to anti-pd-1 or anti-pd-l1 immunotherapy : an imaging biomarker, retrospective multicohort study,” *Lancet Oncol*, vol. 19, n  . 9, p. 1180–1191, 2018.
- [105] Y. Liu, D. Francis, V. Groza, C. Voyton, A. Baili-Laya, B. Huet et N. Boujemaa, “Machine learning based biomarkers to predict CD8 infiltration in non-small cell lung cancers using CT imaging.” *Journal of Clinical Oncology*, vol. 39, n  . 15_suppl, p. e21145–e21145, 2021, pMID :. [En ligne]. Disponible : https://ascopubs.org/doi/abs/10.1200/JCO.2021.39.15_suppl.e21145
- [106] F. Zerka, M. Felfli, C. M. Voyton, A. Thinn  s, S. Jacques, Y. Liu et A. Iannessi, “A ct-based radiomics approach for cd8+ lymphocytes infiltration stratification in patients with non-small cell lung cancer,” *medRxiv*, 2024. [En ligne]. Disponible : <https://www.medrxiv.org/content/early/2024/05/24/2024.05.23.24307791>
- [107] H. J. Yoon, J. Kang, H. Park, I. Sohn, S. H. Lee et H. Y. Lee, “Deciphering the tumor microenvironment through radiomics in non-small cell lung cancer : Correlation with immune profiles,” *PLoS One*, vol. 15, n  . 4, p. e0231227, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32251447>
- [108] J. Li, Z. Shi, F. Liu, X. Fang, K. Cao, Y. Meng, H. Zhang, J. Yu, X. Feng, Q. Li, Y. Liu, L. Wang, H. Jiang, J. Lu, C. Shao et Y. Bian, “Xgboost classifier based on computed tomography radiomics for prediction of tumor-infiltrating cd8(+) t-cells in patients with pancreatic ductal adenocarcinoma,” *Front Oncol*, vol. 11, p. 671333, 2021. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/34094971>
- [109] Y. Jiang, H. Wang, J. Wu, C. Chen, Q. Yuan, W. Huang, T. Li, S. Xi, Y. Hu, Z. Zhou, Y. Xu, G. Li et R. Li, “Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer,” *Ann Oncol*, vol. 31, n  . 6, p. 760–768, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32240794>
- [110] X. Gao, T. Ma, S. Bai, Y. Liu, Y. Zhang, Y. Wu, H. Li et Z. Ye, “A ct-based radiomics signature for evaluating tumor infiltrating treg cells and outcome prediction of gastric cancer,” *Ann Transl Med*, vol. 8, n  . 7, p. 469, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32395513>
- [111] K. Huang, X. Huang, C. Zeng, S. Wang, Y. Zhan, Q. Cai, G. Peng, Z. Yang, L. Zhou, J. Chen et C. Chen, “Radiomics signature for dynamic changes of tumor-infiltrating

- CD8+ T cells and macrophages in cervical cancer during chemoradiotherapy,” *Cancer Imaging*, vol. 24, n° 1, p. 54, avr. 2024.
- [112] Y. Bian, C. Liu, Q. Li, Y. Meng, F. Liu, H. Zhang, X. Fang, J. Li, J. Yu, X. Feng, C. Ma, Z. Zhao, L. Wang, J. Xu, C. Shao et J. Lu, “Preoperative radiomics approach to evaluating Tumor-Infiltrating CD8(+) T cells in patients with pancreatic ductal adenocarcinoma using noncontrast magnetic resonance imaging,” *J Magn Reson Imaging*, vol. 55, n° 3, p. 803–814, août 2021.
- [113] L. Sun, L. Mu, J. Zhou, W. Tang, L. Zhang, S. Xie, J. Chen et J. Wang, “Imaging features of gadoteric acid-enhanced mr imaging for evaluation of tumor-infiltrating cd8 cells and pd-l1 expression in hepatocellular carcinoma,” *Cancer Immunol Immunother*, 2021. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/33993366>
- [114] Y. J. Ku, H. H. Kim, J. H. Cha, H. J. Shin, S. H. Baek, H. J. Lee et G. Gong, “Correlation between mri and the level of tumor-infiltrating lymphocytes in patients with triple-negative breast cancer,” *AJR Am J Roentgenol*, vol. 207, n° 5, p. 1146–1151, 2016. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/27556407>
- [115] S. Chen, S. Feng, J. Wei, F. Liu, B. Li, X. Li, Y. Hou, D. Gu, M. Tang, H. Xiao, Y. Jia, S. Peng, J. Tian et M. Kuang, “Pretreatment prediction of immunoscore in hepatocellular cancer : a radiomics-based clinical model based on gd-eob-dtpa-enhanced mri imaging,” *Eur Radiol*, vol. 29, n° 8, p. 4177–4187, 2019. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/30666445>
- [116] J. Wu et C. Hicks, “Breast cancer type classification using machine learning,” *Journal of Personalized Medicine*, vol. 11, n° 2, 2021. [En ligne]. Disponible : <https://www.mdpi.com/2075-4426/11/2/61>
- [117] V. Bobée, F. Drieux, V. Marchand, V. Sater, L. Veresezan, J.-M. Picquenot, P.-J. Vially, M.-D. Lanic, M. Viennot, E. Bohers, L. Oberic, C. Copie-Bergman, T. J. Molina, P. Gaulard, C. Haioun, G. Salles, H. Tilly, F. Jardin et P. Ruminy, “Combining gene expression profiling and machine learning to diagnose b-cell non-hodgkin lymphoma,” *Blood Cancer Journal*, vol. 10, n° 5, p. 59, mai 2020.
- [118] Z. Yu, Z. Wang, X. Yu et Z. Zhang, “Rna-seq-based breast cancer subtypes classification using machine learning approaches,” *Computational Intelligence and Neuroscience*, vol. 2020, n° 1, p. 4737969, 2020. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1155/2020/4737969>
- [119] X. Zhou, Y. Qian, C. Ling, Z. He, P. Shi, Y. Gao et X. Sui, “An integrated framework for prognosis prediction and drug response modeling in colorectal liver metastasis drug discovery,” *Journal of Translational Medicine*, vol. 22, n° 1, p. 321, mars 2024.

- [120] S. Xie, S. Hou, J. Chen et X. Qi, “Integrative single-cell and bulk rna sequencing identifies a macrophage-related prognostic signature for predicting prognosis and therapy responses in colorectal cancer,” *International Journal of Molecular Sciences*, vol. 26, n°. 2, 2025. [En ligne]. Disponible : <https://www.mdpi.com/1422-0067/26/2/811>
- [121] Z. Huang, T. S. Johnson, Z. Han, B. Helm, S. Cao, C. Zhang, P. Salama, M. Rizkalla, C. Y. Yu, J. Cheng, S. Xiang, X. Zhan, J. Zhang et K. Huang, “Deep learning-based cancer survival prognosis from RNA-seq data : approaches and evaluations,” *BMC Medical Genomics*, vol. 13, n°. 5, p. 41, avr. 2020.
- [122] R. J. Pais, F. Lopes, I. Parreira, M. Silva, M. Silva et M. G. Moutinho, “Predicting cancer prognostics from tumour transcriptomics using an auto machine learning approach,” *Medical Sciences Forum*, vol. 22, n°. 1, 2023. [En ligne]. Disponible : <https://www.mdpi.com/2673-9992/22/1/6>
- [123] S. Bhattacharjee, B. Saha et S. Saha, “Prediction of recurrence in non small cell lung cancer patients with gene expression data using machine learning techniques,” dans *2023 International Conference on Computer, Electrical Communication Engineering (ICCECE)*, 2023, p. 1–8.
- [124] K. Cai, W. Fu, H. Liu, X. Yang, Z. Wang et X. Zhao, “Leveraging bioinformatics and machine learning for identifying prognostic biomarkers and predicting clinical outcomes in lung adenocarcinoma,” *Genes*, vol. 15, n°. 12, 2024. [En ligne]. Disponible : <https://www.mdpi.com/2073-4425/15/12/1497>
- [125] M. Wang, J. Chen et Y. Liu, “Optimized deep learning model for predicting liver metastasis in colorectal cancer patients,” *Symmetry*, vol. 17, n°. 1, 2025. [En ligne]. Disponible : <https://www.mdpi.com/2073-8994/17/1/103>
- [126] Z. Jiang, L. Yan, S. Deng, J. Gu, L. Qin, F. Mao, Y. Xue, W. Cai, X. Nie, H. Liu, F. Shang, K. Tao, J. Wang, K. Wu, Y. Cao et K. Cai, “Development and interpretation of a Clinicopathological-Based model for the identification of microsatellite instability in colorectal cancer,” *Dis Markers*, vol. 2023, p. 5178750, févr. 2023.
- [127] A. Janik, M. Torrente, L. Costabello, V. Calvo, B. Walsh, C. Camps, S. K. Mohamed, A. L. Ortega, V. Nováček, B. Massutí, P. Minervini, M. R. G. Campelo, E. del Barco, J. Bosch-Barrera, E. Menasalvas, M. Timilsina et M. Provencio, “Machine learning–assisted recurrence prediction for patients with early-stage non–small-cell lung cancer,” *JCO Clinical Cancer Informatics*, n°. 7, p. e2200062, 2023, pMID : 37428988. [En ligne]. Disponible : <https://ascopubs.org/doi/abs/10.1200/CCI.22.00062>
- [128] F. Carrillo-Perez, J. C. Morales, D. Castillo-Secilla, Y. Molina-Castro, A. Guillén, I. Rojas et L. J. Herrera, “Non-small-cell lung cancer classification via RNA-Seq and his-

- tology imaging probability fusion,” *BMC Bioinformatics*, vol. 22, n° 1, p. 454, sept. 2021.
- [129] Y. Cao, J. Zhang, L. Huang, Z. Zhao, G. Zhang, J. Ren, H. Li, H. Zhang, B. Guo, Z. Wang, Y. Xing et J. Zhou, “Construction of prediction model for KRAS mutation status of colorectal cancer based on CT radiomics,” *Jpn J Radiol*, vol. 41, n° 11, p. 1236–1246, juin 2023.
- [130] M. Taghavi, F. Staal, F. Gomez Munoz, F. Imani, D. B. Meek, R. Simões, L. G. Klompenhouwer, U. A. van der Heide, R. G. H. Beets-Tan et M. Maas, “CT-Based radiomics analysis before thermal ablation to predict local tumor progression for colorectal liver metastases,” *CardioVascular and Interventional Radiology*, vol. 44, n° 6, p. 913–920, juin 2021.
- [131] D. J. van der Reijnd, C. Guerendel, F. C. R. Staal, M. P. Busard, M. De Oliveira Taveira, E. G. Klompenhouwer, K. F. D. Kuhlmann, A. Moelker, C. Verhoef, M. P. A. Star-mans, D. M. J. Lambregts, R. G. H. Beets-Tan, S. Benson et M. Maas, “Independent validation of CT radiomics models in colorectal liver metastases : predicting local tumour progression after ablation,” *European Radiology*, vol. 34, n° 6, p. 3635–3643, juin 2024.
- [132] H. Zhang, X. Meng, Z. Wang, X. Zhou, Y. Liu et N. Li, “Predicting pd-l1 in lung adenocarcinoma using 18f-fdg pet/ct radiomic features,” *Diagnostics*, vol. 15, n° 5, 2025. [En ligne]. Disponible : <https://www.mdpi.com/2075-4418/15/5/543>
- [133] R. Da-ano, G. Andrade-Miranda, O. Tankyevych, D. Visvikis, P.-H. Conze et C. C. L. Rest, “Automated PD-L1 status prediction in lung cancer with multi-modal PET/CT fusion,” *Scientific Reports*, vol. 14, n° 1, p. 16720, juill. 2024.
- [134] C. Liang, M. Zheng, H. Zou, Y. Han, Y. Zhan, Y. Xing, C. Liu, C. Zuo et J. Zou, “Deep learning-based image analysis predicts pd-l1 status from 18f-fdg pet/ct images in non-small-cell lung cancer,” *Frontiers in Oncology*, vol. 14, 2024. [En ligne]. Disponible : <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2024.1402994>
- [135] G. Mehri-Kakavand, S. Mdletshe et A. Wang, “A comprehensive review on the application of artificial intelligence for predicting postsurgical recurrence risk in Early-Stage Non-Small cell lung cancer using computed tomography, positron emission tomography, and clinical data,” *J Med Radiat Sci*, janv. 2025.
- [136] A. Valter, T. Kordemets, A. Gasimova, N. Waterfield Price, L. Freitag, T. Vanakesa, I. Almre et K. Oselin, “1269P machine learning model for predicting lung cancer recurrence after surgical treatment : A retrospective study using NLST and european hospital data,” *Annals of Oncology*, vol. 34, p. S735, oct. 2023.

- [137] J. R. Christie, O. Daher, M. Abdelrazek, P. E. Romine, R. A. Malthaner, M. Qiabi, R. Nayak, S. Napel, V. S. Nair et S. A. Mattonen, “Predicting recurrence risks in lung cancer patients using multimodal radiomics and random survival forests,” *J Med Imaging (Bellingham)*, vol. 9, n^o. 6, p. 066001, nov. 2022.
- [138] Y. She, B. He, F. Wang, Y. Zhong, T. Wang, Z. Liu, M. Yang, B. Yu, J. Deng, X. Sun, C. Wu, L. Hou, Y. Zhu, Y. Yang, H. Hu, D. Dong, C. Chen et J. Tian, “Deep learning for predicting major pathological response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer : A multicentre study,” *eBioMedicine*, vol. 86, p. 104364, 2022, chengqi Chen. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S2352396422005461>
- [139] P. Grossmann, O. Stringfield, N. El-Hachem, M. M. Bui, E. Rios Velazquez, C. Parmar, R. T. Leijenaar, B. Haibe-Kains, P. Lambin, R. J. Gillies et H. J. Aerts, “Defining the biological basis of radiomic phenotypes in lung cancer,” *eLife*, vol. 6, p. e23421, jul 2017. [En ligne]. Disponible : <https://doi.org/10.7554/eLife.23421>
- [140] S. Starke, A. Zwanenburg, K. Leger, K. Zöphel, J. Kotzerke, M. Krause, M. Baumann, E. G. C. Troost et S. Löck, “Longitudinal and multimodal radiomics models for head and neck cancer outcome prediction,” *Cancers*, vol. 15, n^o. 3, 2023. [En ligne]. Disponible : <https://www.mdpi.com/2072-6694/15/3/673>
- [141] R. S. Vanguri, J. Luo, A. T. Aukerman, J. V. Egger, C. J. Fong, N. Horvat, A. Pagano, J. d. A. B. Araujo-Filho, L. Geneslaw, H. Rizvi, R. Sosa, K. M. Boehm, S.-R. Yang, F. M. Bodd, K. Ventura, T. J. Hollmann, M. S. Ginsberg, J. Gao, R. Vanguri, M. D. Hellmann, J. L. Sauter, S. P. Shah et MSK MIND Consortium, “Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer,” *Nature Cancer*, vol. 3, n^o. 10, p. 1151–1164, oct. 2022.
- [142] N. Captier, M. Lerousseau, F. Orlhac, N. Hovhannisyan-Baghdasarian, M. Luporsi, E. Woff, S. Lagha, P. Salamoun Feghali, C. Lonjou, C. Beaulaton, A. Zinovyev, H. Salmon, T. Walter, I. Buvat, N. Girard et E. Barillot, “Integration of clinical, pathological, radiological, and transcriptomic data improves prediction for first-line immunotherapy outcome in metastatic non-small cell lung cancer,” *Nature Communications*, vol. 16, n^o. 1, p. 614, janv. 2025.
- [143] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit et N. Houlsby, “An image is worth 16x16 words : Transformers for image recognition at scale,” 2021.
- [144] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser et I. Polosukhin, “Attention is all you need,” 2023. [En ligne]. Disponible :

<https://arxiv.org/abs/1706.03762>

- [145] Y. Dai, Y. Gao et F. Liu, “Transmed : Transformers advance multi-modal medical image classification,” *Diagnostics*, vol. 11, n°. 8, 2021. [En ligne]. Disponible : <https://www.mdpi.com/2075-4418/11/8/1384>
- [146] J. Chen, J. Mei, X. Li, Y. Lu, Q. Yu, Q. Wei, X. Luo, Y. Xie, E. Adeli, Y. Wang, M. P. Lungren, S. Zhang, L. Xing, L. Lu, A. Yuille et Y. Zhou, “Transunet : Rethinking the u-net architecture design for medical image segmentation through the lens of transformers,” *Medical Image Analysis*, vol. 97, p. 103280, 2024. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1361841524002056>
- [147] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin et B. Guo, “Swin transformer : Hierarchical vision transformer using shifted windows,” *CoRR*, vol. abs/2103.14030, 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2103.14030>
- [148] S. Somvanshi, S. Das, S. A. Javed, G. Antariksa et A. Hossain, “A survey on deep tabular learning,” 2024. [En ligne]. Disponible : <https://arxiv.org/abs/2410.12034>
- [149] S. Arik et T. Pfister, “Tabnet : Attentive interpretable tabular learning,” *ArXiv*, vol. abs/1908.07442, 2021.
- [150] X. Huang, A. Khetan, M. Cvitkovic et Z. Karnin, “Tabtransformer : Tabular data modeling using contextual embeddings,” 2020. [En ligne]. Disponible : <https://arxiv.org/abs/2012.06678>
- [151] Y. Gorishniy, I. Rubachev, V. Khrulkov et A. Babenko, “Revisiting deep learning models for tabular data,” 2023.
- [152] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville et Y. Bengio, “Generative adversarial networks,” 2014. [En ligne]. Disponible : <https://arxiv.org/abs/1406.2661>
- [153] Y. Wang, L. Liu et C. Wang, “Trends in using deep learning algorithms in biomedical prediction systems,” *Front Neurosci*, vol. 17, p. 1256351, nov. 2023.
- [154] M. E. A. Elforaici, E. Montagnon, F. P. Romero, W. T. Le, F. Azzi, D. Trudel, B. Nguyen, S. Turcotte, A. Tang et S. Kadoury, “Semi-supervised vit knowledge distillation network with style transfer normalization for colorectal liver metastases survival prediction,” *Med. Image Anal.*, vol. 99, p. 103346, 2025.
- [155] Annapoorani Gopal, Lathaselvi Gandhimaruthian et Javid Ali, “Role of general adversarial networks in mammogram analysis : A review,” p. 863–877, 2020.
- [156] G. O. Ghosheh, J. Li et T. Zhu, “A survey of generative adversarial networks for synthesizing structured electronic health records,” *ACM Comput. Surv.*, vol. 56, n°. 6, janv. 2024. [En ligne]. Disponible : <https://doi.org/10.1145/3636424>

- [157] N. Park, M. Mohammadi, K. Gorde, S. Jajodia, H. Park et Y. Kim, “Data synthesis based on generative adversarial networks,” *Proceedings of the VLDB Endowment*, vol. 11, n^o. 10, p. 1071–1083, juin 2018. [En ligne]. Disponible : <http://dx.doi.org/10.14778/3231751.3231757>
- [158] E. Choi, S. Biswal, B. Malin, J. Duke, W. F. Stewart et J. Sun, “Generating multi-label discrete patient records using generative adversarial networks,” 2018. [En ligne]. Disponible : <https://arxiv.org/abs/1703.06490>
- [159] L. Xu, M. Skoularidou, A. Cuesta-Infante et K. Veeramachaneni, “Modeling tabular data using conditional GAN,” 2019. [En ligne]. Disponible : <https://arxiv.org/abs/1907.00503>
- [160] V. Agiwal et S. Chaudhuri, “Methods and implications of addressing missing data in health-care research,” *Current Medical Issues*, vol. 22, n^o. 1, 2024.
- [161] W. Dong, D. Y. T. Fong, J.-S. Yoon, E. Y. F. Wan, L. E. Bedford, E. H. M. Tang et C. L. K. Lam, “Generative adversarial networks for imputing missing data for big data clinical research,” *BMC Med Res Methodol*, vol. 21, n^o. 1, p. 78, avr. 2021.
- [162] S. Azizi, L. Culp, J. Freyberg, B. Mustafa, S. Baur, S. Kornblith, T. Chen, P. MacWilliams, S. S. Mahdavi, E. Wulczyn, B. Babenko, M. Wilson, A. Loh, P.-H. C. Chen, Y. Liu, P. Bavishi, S. M. McKinney, J. Winkens, A. G. Roy, Z. Beaver, F. Ryan, J. Krogue, M. Etemadi, U. Telang, Y. Liu, L. Peng, G. S. Corrado, D. R. Webster, D. Fleet, G. Hinton, N. Houlsby, A. Karthikesalingam, M. Norouzi et V. Natarajan, “Robust and efficient medical imaging with self-supervision,” 2022. [En ligne]. Disponible : <https://arxiv.org/abs/2205.09723>
- [163] T. Chen, S. Kornblith, M. Norouzi et G. E. Hinton, “A simple framework for contrastive learning of visual representations,” *CoRR*, vol. abs/2002.05709, 2020. [En ligne]. Disponible : <https://arxiv.org/abs/2002.05709>
- [164] L. Zhu, Z. Ke et R. Lau, “Towards self-adaptive pseudo-label filtering for semi-supervised learning,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2309.09774>
- [165] A. Tarvainen et H. Valpola, “Mean teachers are better role models : Weight-averaged consistency targets improve semi-supervised deep learning results,” 2018. [En ligne]. Disponible : <https://arxiv.org/abs/1703.01780>
- [166] M.-R. Amini, V. Feofanov, L. Pauletto, L. Hadjadj, Devijver et Y. Maximov, “Self-training : A survey,” *Neurocomputing*, vol. 616, p. 128904, 2025. [En ligne]. Disponible : <http://dx.doi.org/10.1016/j.neucom.2024.128904>
- [167] R. C. Godwin, A. Tung, D. E. Berkowitz et R. L. Melvin, “Transforming physiology and healthcare through foundation models,” *Physiology*, vol. 40, n^o. 3, p. 000–000, 2025, pMID : 39832521. [En ligne]. Disponible : <https://doi.org/10.1152/physiol.00048.2024>

- [168] Y. Han, X. Liu, X. Zhang et C. Ding, “Foundation models in electrocardiogram : A review,” 2024. [En ligne]. Disponible : <https://arxiv.org/abs/2410.19877>
- [169] M. Sevgi, F. Antaki et P. A. Keane, “Medical education with large language models in ophthalmology : custom instructions and enhanced retrieval capabilities,” *Br J Ophthalmol*, vol. 108, n^o. 10, p. 1354–1361, sept. 2024.
- [170] C. Shi, R. Rezai, J. Yang, Q. Dou et X. Li, “A survey on trustworthiness in foundation models for medical image analysis,” 2024. [En ligne]. Disponible : <https://arxiv.org/abs/2407.15851>
- [171] J. Ma, Y. He, F. Li, L. Han, C. You et B. Wang, “Segment anything in medical images,” *Nat. Commun.*, vol. 15, n^o. 1, janv. 2024.
- [172] R. Saber *et al.*, “Radiomics using computed tomography to predict CD73 expression and prognosis of colorectal cancer liver metastases,” *J Transl Med*, vol. 21, n^o. 1, p. 507, juill. 2023.
- [173] R. Saber, D. Henault, R. Rebolledo, S. Turcotte et S. Kadoury, “Ensemble tabnet predicting a t-cell/mhc-i-based immune profile biomarker for colorectal liver metastases from CT images,” dans *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, 2023, p. 1–5.
- [174] D. Henault, D. Stephen, P.-A. St-Hilaire *et al.*, “Prognostic immune scoring of colorectal cancer liver metastasis with mhc class-i expression combined to t cell quantification.” *Journal of Clinical Oncology*, vol. 36, n^o. 15_suppl, p. 3586–3586, 2018. [En ligne]. Disponible : https://doi.org/10.1200/JCO.2018.36.15_suppl.3586
- [175] R. Saber, D. Henault, E. Vorontsov, E. Montagnon, A. Tang, S. Turcotte et S. Kadoury, “Prediction of CD3 T-cell infiltration status in colorectal liver metastases : a radiomics-based imaging biomarker,” dans *Medical Imaging 2022 : Computer-Aided Diagnosis*, K. Drukker et K. M. Iftexharuddin, édit., vol. 12033, International Society for Optics and Photonics. SPIE, 2022, p. 120331L. [En ligne]. Disponible : <https://doi.org/10.1117/12.2612696>
- [176] R. Saber, M. Tonneau, S. Bahig, J. Malo, W. Belkaid, M. Messaoudene, N. Bouchard, F. Coulombe, P. Joubert, H. Bahig, S. Turcotte, B. Routy et S. Kadoury, “Feature tokenizer-transformers with self-training for the prediction of pd-l1 expression of non-small cell lung cancer from ct,” dans *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024, p. 1–5.
- [177] R. Saber, B. Routy, S. Turcotte et S. Kadoury, “RNA sequencing-based histological subtyping of non-small cell lung cancer with generative adversarial data imputation,” dans *2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2023, p. 1–4.

- [178] R. L. Siegel, K. D. Miller, A. Goding Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith et A. Jemal, “Colorectal cancer statistics, 2020,” *CA Cancer J Clin*, vol. 70, n°. 3, p. 145–164, 2020.
- [179] J. S. Tomlinson, W. R. Jarnagin, R. P. DeMatteo, Y. Fong, P. Kornprat, M. Gonen, N. Kemeny, M. F. Brennan, L. H. Blumgart et M. D’Angelica, “Actual 10-year survival after resection of colorectal liver metastases defines cure,” *J Clin Oncol*, vol. 25, n°. 29, p. 4575–80, 2007.
- [180] Y. Kanemitsu, Y. Shimizu, J. Mizusawa, Y. Inaba, T. Hamaguchi, D. Shida, M. Ohue, K. Komori, A. Shiomi, M. Shiozawa, J. Watanabe, T. Suto, Y. Kinugasa, Y. Takii, H. Bando, T. Kobatake, M. Inomata, Y. Shimada, H. Katayama et H. Fukuda, “Hepatectomy followed by mfolfox6 versus hepatectomy alone for liver-only metastatic colorectal cancer (jcog0603) : A phase ii or iii randomized controlled trial,” *J Clin Oncol*, vol. 39, n°. 34, p. 3789–3799, 2021.
- [181] B. Zhang, B. Song, X. Wang, X. S. Chang, T. Pang, X. Zhang, K. Yin et G. E. Fang, “The expression and clinical significance of cd73 molecule in human rectal adenocarcinoma,” *Tumour Biol*, vol. 36, n°. 7, p. 5459–66, 2015.
- [182] R. D. Leone et L. A. Emens, “Targeting adenosine for cancer immunotherapy,” *J Immunother Cancer*, vol. 6, n°. 1, p. 57, 2018.
- [183] P. W. Eide, S. H. Moosavi, I. A. Eilertsen, T. H. Brunsell, J. Langerud, K. C. G. Berg, B. I. Røsok, B. A. Bjørnbeth, A. Nesbakken, R. A. Lothe et A. Sveen, “Metastatic heterogeneity of the consensus molecular subtypes of colorectal cancer,” *NPJ Genom Med*, vol. 6, n°. 1, p. 59, 2021.
- [184] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. van der Laak, B. van Ginneken et C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Med Image Anal*, vol. 42, p. 60–88, 2017.
- [185] W. Hu, H. Yang, H. Xu et Y. Mao, “Radiomics based on artificial intelligence in liver diseases : where we are?” *Gastroenterol Rep (Oxf)*, vol. 8, n°. 2, p. 90–97, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/32280468>
- [186] D. Nam, J. Chapiro, V. Paradis, T. P. Seraphin et J. N. Kather, “Artificial intelligence in liver diseases : Improving diagnostics, prognostics and response prediction,” *JHEP Rep*, vol. 4, n°. 4, p. 100443, 2022. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/35243281>
- [187] S. Naganawa, K. Enooku, R. Tateishi, H. Akai, K. Yasaka, J. Shibahara, T. Ushiku, O. Abe, K. Ohtomo et S. Kiryu, “Imaging prediction of nonalcoholic steatohepatitis using computed tomography texture analysis,” *Eur Radiol*, vol. 28, n°. 7, p. 3050–3058, 2018.

- [188] J. Cheng, J. Wei, T. Tong, W. Sheng, Y. Zhang, Y. Han, D. Gu, N. Hong, Y. Ye, J. Tian et Y. Wang, “Prediction of histopathologic growth patterns of colorectal liver metastases with a noninvasive imaging method,” *Ann Surg Oncol*, vol. 26, n^o. 13, p. 4587–4598, 2019.
- [189] F. Fiz, L. Vigano, N. Gennaro, G. Costa, L. La Bella, A. Boichuk, L. Cavinato, M. Sollini, L. S. Politi, A. Chiti et G. Torzilli, “Radiomics of liver metastases : A systematic review,” *Cancers (Basel)*, vol. 12, n^o. 10, 2020. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/33036490>
- [190] J. J. M. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J. C. Fillion-Robin, S. Pieper et H. Aerts, “Computational radiomics system to decode the radiographic phenotype,” *Cancer Res*, vol. 77, n^o. 21, p. e104–e107, 2017.
- [191] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society : Series B (Methodological)*, vol. 58, n^o. 1, p. 267–288, 1996. [En ligne]. Disponible : <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [192] D. P. Kingma et J. Ba, “Adam : A method for stochastic optimization,” *arXiv e-prints*, p. arXiv :1412.6980, 2014. [En ligne]. Disponible : <https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6980K>
- [193] S. M. Lundberg et S.-I. Lee, “A unified approach to interpreting model predictions,” *ArXiv*, vol. abs/1705.07874, 2017.
- [194] A. J. Vickers et E. B. Elkin, “Decision curve analysis : a novel method for evaluating prediction models,” *Med Decis Making*, vol. 26, n^o. 6, p. 565–74, 2006.
- [195] L. Buisseret, S. Pommey, B. Allard, S. Garaud, M. Bergeron, I. Cousineau, L. Ameye, Y. Bareche, M. Paesmans, J. P. A. Crown, A. Di Leo, S. Loi, M. Piccart-Gebhart, K. Willard-Gallo, C. Sotiriou et J. Stagg, “Clinical significance of cd73 in triple-negative breast cancer : multiplex analysis of a phase iii clinical trial,” *Ann Oncol*, vol. 29, n^o. 4, p. 1056–1062, 2018.
- [196] B. G. Leclerc, R. Charlebois, G. Chouinard, B. Allard, S. Pommey, F. Saad et J. Stagg, “Cd73 expression is an independent prognostic factor in prostate cancer,” *Clin Cancer Res*, vol. 22, n^o. 1, p. 158–66, 2016.
- [197] C. Jacobberger-Foissac, I. Cousineau, Y. Bareche, D. Allard, P. Chrobak, B. Allard, S. Pommey, N. Messaoudi, Y. McNicoll, G. Soucy, S. Koseoglu, R. Masia, A. C. Lake, H. Seo, C. B. Eeles, N. Rohatgi, S. C. Robson, S. Turcotte, B. Haibe-Kains et J. Stagg, “Cd73 inhibits cgas-sting and cooperates with cd39 to promote pancreatic cancer,” *Cancer Immunol Res*, vol. 11, n^o. 1, p. 56–71, 2023.

- [198] M. Turcotte, K. Spring, S. Pommey, G. Chouinard, I. Cousineau, J. George, G. M. Chen, D. M. Gendoo, B. Haibe-Kains, T. Karn, K. Rahimi, C. Le Page, D. Provencher, A. M. Mes-Masson et J. Stagg, “Cd73 is associated with poor prognosis in high-grade serous ovarian cancer,” *Cancer Res*, vol. 75, n°. 21, p. 4494–503, 2015.
- [199] V. S. Parekh et M. A. Jacobs, “Deep learning and radiomics in precision medicine,” *Expert Rev Precis Med Drug Dev*, vol. 4, n°. 2, p. 59–72, 2019.
- [200] H. J. Aerts, “The potential of radiomic-based phenotyping in precision medicine : A review,” *JAMA Oncol*, vol. 2, n°. 12, p. 1636–1642, 2016.
- [201] R. J. Gillies et M. B. Schabath, “Radiomics improves cancer screening and early detection,” *Cancer Epidemiol Biomarkers Prev*, vol. 29, n°. 12, p. 2556–2567, 2020.
- [202] S. Li et B. Zhou, “A review of radiomics and genomics applications in cancers : the way towards precision medicine,” *Radiat Oncol*, vol. 17, n°. 1, p. 217, 2022.
- [203] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert et L. Cilar, “Interpretability of machine learning-based prediction models in healthcare,” *WIREs Data Mining and Knowledge Discovery*, vol. 10, n°. 5, p. e1379, 2020. [En ligne]. Disponible : <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1379>
- [204] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser et I. Polosukhin, “Attention is all you need,” *ArXiv*, vol. abs/1706.03762, 2017.
- [205] A. K. Mohankumar, P. Nema, S. Narasimhan, M. M. Khapra, B. V. Srinivasan et B. Ravindran, “Towards transparent and explainable attention models,” dans *ACL, Conference Proceedings*.
- [206] J. H. Wang, K. A. Wahid, L. V. van Dijk, K. Farahani, R. F. Thompson et C. D. Fuller, “Radiomic biomarkers of tumor immune biology and immunotherapy response,” *Clin Transl Radiat Oncol*, vol. 28, p. 97–115, 2021.
- [207] C. Tang, B. Hobbs, A. Amer, X. Li, C. Behrens, J. R. Canales, E. P. Cuentas, P. Villalobos, D. Fried, J. Y. Chang, D. S. Hong, J. W. Welsh, B. Sepesi, L. Court, I. Wistuba et E. J. Koay, “Development of an immune-pathology informed radiomics model for non-small cell lung cancer,” *Sci Rep*, vol. 8, n°. 1, p. 1922, 2018. [En ligne]. Disponible : <https://www.ncbi.nlm.nih.gov/pubmed/29386574>
- [208] A. Eftekhari, C. Krysch, D. Pamies, S. Gulec, E. Ahmadian, D. Janas, S. Davaran et R. Khalilov, “Natural and synthetic nanovectors for cancer therapy,” *Nanotheranostics*, vol. 7, n°. 3, p. 236–257, 2023. [En ligne]. Disponible : <https://www.ntno.org/v07p0236.htm>
- [209] V. N. Ramazanli et I. S. Ahmadov, “Synthesis of silver nanoparticles by using extract of olive leaves,” *Advances in Biology Earth Sciences*, vol. 7, n°. 3, p. 238–244, 2022.

- [210] L. Devkota, Z. Starosolski, C. H. Rivas, I. Stupin, A. Annapragada, K. B. Ghaghada et R. Parihar, “Detection of response to tumor microenvironment-targeted cellular immunotherapy using nano-radiomics,” *Sci Adv*, vol. 6, n^o. 28, p. eaba6156, 2020.
- [211] C. Y. Kang, S. E. Duarte, H. S. Kim, E. Kim, J. Park, A. D. Lee, Y. Kim, L. Kim, S. Cho, Y. Oh, G. Gim, I. Park, D. Lee, M. Abazeed, Y. S. Velichko et Y. K. Chae, “Artificial intelligence-based radiomics in the era of immuno-oncology,” *Oncologist*, vol. 27, n^o. 6, p. e471–e483, 2022.
- [212] Z. Bodalal, S. Trebeschi, T. D. L. Nguyen-Kim, W. Schats et R. Beets-Tan, “Radiogenomics : bridging imaging and genomics,” *Abdom Radiol (NY)*, vol. 44, n^o. 6, p. 1960–1984, 2019.
- [213] S.-X. Rao, D. M. Lambregts, R. S. Schnerr, R. C. Beckers, M. Maas, F. Albarello, R. G. Riedl, C. H. Dejong, M. H. Martens, L. A. Heijnen, W. H. Backes, G. L. Beets, M.-S. Zeng et R. G. Beets-Tan, “Ct texture analysis in colorectal liver metastases : A better way than size and volume measurements to assess response to chemotherapy ?” *United European gastroenterology journal*, vol. 4, n^o. 2, p. 257–263, 2016. [En ligne]. Disponible : <http://europepmc.org/abstract/MED/27087955><https://doi.org/10.1177/2050640615601603><https://europepmc.org/articles/PMC4804371><https://europepmc.org/articles/PMC4804371?pdf=render>
- [214] S. J. Ahn, J. H. Kim, S. M. Lee, S. J. Park et J. K. Han, “Ct reconstruction algorithms affect histogram and texture analysis : evidence for liver parenchyma, focal solid liver lesions, and renal cysts,” *Eur Radiol*, vol. 29, n^o. 8, p. 4008–4015, 2019.
- [215] I. R. Andersen, K. Thorup, M. B. Andersen, R. Olesen, F. V. Mortensen, D. T. Nielsen et F. Rasmussen, “Texture in the monitoring of regorafenib therapy in patients with colorectal liver metastases,” *Acta Radiol*, vol. 60, n^o. 9, p. 1084–1093, 2019.
- [216] D. B. Johnson, R. J. Sullivan et A. M. Menzies, “Immune checkpoint inhibitors in challenging populations,” *Cancer*, vol. 123, n^o. 11, p. 1904–1911, 2017.
- [217] D. Vijayan, A. Young, M. W. L. Teng et M. J. Smyth, “Targeting immunosuppressive adenosine in cancer,” *Nat Rev Cancer*, vol. 17, n^o. 12, p. 709–724, 2017.
- [218] B. Zhang, “Cd73 : a novel target for cancer immunotherapy,” *Cancer Res*, vol. 70, n^o. 16, p. 6407–11, 2010.
- [219] M. Roh, D. A. Wainwright, J. D. Wu, Y. Wan et B. Zhang, “Targeting cd73 to augment cancer immunotherapy,” *Curr Opin Pharmacol*, vol. 53, p. 66–76, 2020.
- [220] M. Rebersek, “Gut microbiome and its role in colorectal cancer,” *BMC Cancer*, vol. 21, n^o. 1, p. 1325, déc. 2021.

- [221] A. I. Valderrama-Treviño, B. Barrera-Mera, J. C. Ceballos-Villalva *et al.*, “Hepatic metastasis from colorectal cancer,” *Euroasian J Hepatogastroenterol*, vol. 7, n°. 2, p. 166–175, sept. 2017.
- [222] P. Kanikarla Marie, C. Haymaker, E. R. Parra *et al.*, “Pilot Clinical Trial of Perioperative Durvalumab and Tremelimumab in the Treatment of Resectable Colorectal Cancer Liver Metastases,” *Clinical Cancer Research*, vol. 27, n°. 11, p. 3039–3049, 06 2021. [En ligne]. Disponible : <https://doi.org/10.1158/1078-0432.CCR-21-0163>
- [223] F. E. Buisman, B. Galjart, E. P. van der Stok *et al.*, “Recurrence patterns after resection of colorectal liver metastasis are modified by perioperative systemic chemotherapy,” *World J Surg*, vol. 44, n°. 3, p. 876–886, mars 2020.
- [224] M. J. M. Gooden, G. H. de Bock, N. Leffers *et al.*, “The prognostic influence of tumour-infiltrating lymphocytes in cancer : a systematic review with meta-analysis,” *Br J Cancer*, vol. 105, n°. 1, p. 93–103, mai 2011.
- [225] J. Galon, A. Costes, F. Sanchez-Cabo *et al.*, “Type, density, and location of immune cells within human colorectal tumors predict clinical outcome,” *Science*, vol. 313, n°. 5795, p. 1960–1964, 2006. [En ligne]. Disponible : <https://www.science.org/doi/abs/10.1126/science.1129139>
- [226] H. Yu, L. T. Yang, Q. Zhang *et al.*, “Convolutional neural networks for medical image analysis : State-of-the-art, comparisons, improvement and perspectives,” *Neurocomputing*, vol. 444, p. 92–110, juill. 2021.
- [227] K. H. Cha, L. Hadjiiski, H.-P. Chan *et al.*, “Bladder cancer treatment response assessment in CT using radiomics with Deep-Learning,” *Scientific Reports*, vol. 7, n°. 1, p. 8738, août 2017.
- [228] O. Sagi et L. Rokach, “Ensemble learning : A survey,” *WIREs Data Mining and Knowledge Discovery*, vol. 8, n°. 4, p. e1249, 2018. [En ligne]. Disponible : <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1249>
- [229] N. Messaoudi, D. Henault, D. Stephen *et al.*, “Prognostic implications of adaptive immune features in MMR-proficient colorectal liver metastases classified by histopathological growth patterns,” *British Journal of Cancer*, vol. 126, n°. 9, p. 1329–1338, mai 2022.
- [230] R. L. Camp, M. Dolled-Fillhart et D. L. Rimm, “X-tile : a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization,” *Clin Cancer Res*, vol. 10, n°. 21, p. 7252–7259, nov. 2004.
- [231] S. Kumari, D. Kumar et M. Mittal, “An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier,” *International Journal of Cognitive Computing in Engineering*, vol. 2, p. 40–46, juin 2021.

- [232] N. V. Chawla *et al.*, “SMOTE : Synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, p. 321–357, jun 2002.
- [233] L. Gandhi, D. Rodríguez-Abreu, S. Gadgeel, E. Esteban, E. Felip, F. D. Angelis, M. Domine, P. Clingan, M. J. Hochmair, S. F. Powell, S. Y.-S. Cheng, H. G. Bischoff, N. Peled, F. Grossi, R. R. Jennens, M. Reck, R. Hui, E. B. Garon, M. Boyer, B. Rubio-Viqueira, S. Novello, T. Kurata, J. E. Gray, J. Vida, Z. Wei, J. Yang, H. Raftopoulos, M. C. Pietanza et M. C. Garassino, “Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer,” *New England Journal of Medicine*, vol. 378, n°. 22, p. 2078–2092, 2018. [En ligne]. Disponible : <https://www.nejm.org/doi/full/10.1056/NEJMoa1801005>
- [234] M. A. Socinski, R. M. Jotte, F. Cappuzzo, F. Orlandi, D. Stroyakovskiy, N. Nogami, D. Rodríguez-Abreu, D. Moro-Sibilot, C. A. Thomas, F. Barlesi, G. Finley, C. Kelsch, A. Lee, S. Coleman, Y. Deng, Y. Shen, M. Kowanetz, A. Lopez-Chavez, A. Sandler et M. Reck, “Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC,” *New England Journal of Medicine*, vol. 378, n°. 24, p. 2288–2301, 2018. [En ligne]. Disponible : <https://www.nejm.org/doi/full/10.1056/NEJMoa1716948>
- [235] W. Kian, L. C. Roisman, D. Levitas, N. Wallach, A. Abo-Sharb, A. Y. Cohen, A. Yakobson, N. Peled et Y. Dudnik, “Non-small cell lung cancer pdl1 >50
- [236] Q. Sun, X. Wei, Z. Wang, Y. Zhu, W. Zhao et Y. Dong, “Primary and acquired resistance against immune check inhibitors in non-small cell lung cancer,” *Cancers (Basel)*, vol. 14, n°. 14, 2022.
- [237] F. Ghiringhelli, F. Bibeau, L. Greillier, J.-D. Fumet, A. Ilie, F. Monville, C. Laugé, A. Catteau, I. Boquet, A. Majdi, E. Morgand, Y. Oulkhoudir, N. Brandone, J. Adam, T. Sbarrato, A. Kassambara, J. Fieschi, S. Garcia, A. L. Lepage, P. Tomasini et J. Galon, “Immunoscore immune checkpoint using spatial quantitative analysis of cd8 and pd-l1 markers is predictive of the efficacy of anti- pd1/pd-l1 immunotherapy in non-small cell lung cancer,” *eBioMedicine*, vol. 92, 2023. [En ligne]. Disponible : <https://doi.org/10.1016/j.ebiom.2023.104633>
- [238] K. Tostes, A. P. Siqueira, R. M. Reis, L. F. Leal et L. Arantes, “Biomarkers for immune checkpoint inhibitor response in NSCLC : Current developments and applicability,” *Int J Mol Sci*, vol. 24, n°. 15, 2023.
- [239] M. Mathew, R. A. Safyan et C. A. Shu, “PD-L1 as a biomarker in NSCLC : challenges and future directions,” *Ann Transl Med*, vol. 5, n°. 18, p. 375, 2017.
- [240] M. Reck, D. Rodríguez-Abreu, A. G. Robinson, R. Hui, T. Csőszi, A. Fülöp, M. Gottfried, N. Peled, A. Tafreshi, S. Cuffe, M. O’Brien, S. Rao, K. Hotta, T. A. Leal, J. W. Riess, E. Jensen, B. Zhao, M. C. Pietanza et J. R. Brahmer, “Five-year outcomes with

- pembrolizumab versus chemotherapy for metastatic non-small-cell lung cancer with pd-l1 tumor proportion score ≥ 50 ,” *J Clin Oncol*, vol. 39, n $^{\circ}$. 21, p. 2339–2349, 2021.
- [241] L. Gandhi, D. Rodríguez-Abreu, S. Gadgeel, E. Esteban, E. Felip, F. De Angelis, M. Domine, P. Clingan, M. J. Hochmair, S. F. Powell, S. Y. Cheng, H. G. Bischoff, N. Peled, F. Grossi, R. R. Jennens, M. Reck, R. Hui, E. B. Garon, M. Boyer, B. Rubio-Viqueira, S. Novello, T. Kurata, J. E. Gray, J. Vida, Z. Wei, J. Yang, H. Raftopoulos, M. C. Pietanza et M. C. Garassino, “Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer,” *N Engl J Med*, vol. 378, n $^{\circ}$. 22, p. 2078–2092, 2018.
- [242] M. Reck, T. E. Ciuleanu, M. Cobo, M. Schenker, B. Zurawski, J. Menezes, E. Richardet, J. Bennouna, E. Felip, O. Juan-Vidal, A. Alexandru, H. Sakai, A. Lingua, F. Reyes, P. J. Souquet, P. De Marchi, C. Martin, M. Pérol, A. Scherpereel, S. Lu, L. Paz-Ares, D. P. Carbone, A. Memaj, S. Marimuthu, X. Zhang, P. Tran et T. John, “First-line nivolumab plus ipilimumab with two cycles of chemotherapy versus chemotherapy alone (four cycles) in advanced non-small-cell lung cancer : Checkmate 91a 2-year update,” *ESMO Open*, vol. 6, n $^{\circ}$. 5, p. 100273, 2021.
- [243] L. Paz-Ares, A. Luft, D. Vicente, A. Tafreshi, M. Gümüş, J. Mazières, B. Hermes, F. Çay Şenler, T. Csősz, A. Fülöp, J. Rodríguez-Cid, J. Wilson, S. Sugawara, T. Kato, K. H. Lee, Y. Cheng, S. Novello, B. Halmos, X. Li, G. M. Lubiniecki, B. Piperdi et D. M. Kowalski, “Pembrolizumab plus chemotherapy for squamous non-small-cell lung cancer,” *N Engl J Med*, vol. 379, n $^{\circ}$. 21, p. 2040–2051, 2018.
- [244] P. M. Forde, J. Spicer, S. Lu, M. Provencio, T. Mitsudomi, M. M. Awad, E. Felip, S. R. Broderick, J. R. Brahmer, S. J. Swanson, K. Kerr, C. Wang, T. E. Ciuleanu, G. B. Saylor, F. Tanaka, H. Ito, K. N. Chen, M. Liberman, E. E. Vokes, J. M. Taube, C. Dorange, J. Cai, J. Fiore, A. Jarkowski, D. Balli, M. Sausen, D. Pandya, C. Y. Calvet et N. Girard, “Neoadjuvant nivolumab plus chemotherapy in resectable lung cancer,” *N Engl J Med*, vol. 386, n $^{\circ}$. 21, p. 1973–1985, 2022.
- [245] J. V. Heymach, D. Harpole, T. Mitsudomi, J. M. Taube, G. Galffy, M. Hochmair, T. Winder, R. Zukov, G. Garbaos, S. Gao, H. Kuroda, G. Ostoros, T. V. Tran, J. You, K. Y. Lee, L. Antonuzzo, Z. Papai-Szekely, H. Akamatsu, B. Biswas, A. Spira, J. Crawford, H. T. Le, M. Aperghis, G. J. Doherty, H. Mann, T. M. Fouad et M. Reck, “Perioperative durvalumab for resectable non-small-cell lung cancer,” *N Engl J Med*, vol. 389, n $^{\circ}$. 18, p. 1672–1684, 2023.
- [246] J. D. Spicer, M. C. Garassino, H. Wakelee, M. Liberman, T. Kato, M. Tsuboi, S.-H. Lee, K.-N. Chen, C. Doooms, M. Majem, E. Eigendorff, G. L. Martinengo, O. Bylicki, D. Rodríguez-Abreu, J. E. Chaft, S. Novello, J. Yang, A. Arunachalam, S. M. Keller, A. Samkari, S. Gao, S. Afanasyev, S. Ahmed, T. Alekshun, G. Alves, I. Anderson,

- L. H. Araujo, A. Arkhipov, A. Arora, J. Bai, P. Begin, A. Belonogov, H. Berard, R. Berceanu-Ion, R. Bernabe Caro, I. Bondarenko, R. Bonnet, J. Bosch Barrera, C. Brocca, M. Bryl, A. Bulotta, O. Bylicki, A. Calles Blanco, E. Carcereny, L. Carvalho, C. Cebotaru, J. Chaft, V. Charu, F. Chaves, J. Chen, K.-N. Chen, H. Chen, Q. Chen, K. Chen, C.-L. Chiang, C.-H. Chiu, S. Cicenias, E. Ciubotaru, T. Ciuleanu, I. Ciurescu, P. Cobb, C. Coetzee, D. Collins, D. Cortinovic, K. Costas, D. Costin, E. H. Cronemberger, R. Cuevo, S. Cuffe, P. R. M. De Marchi, T. de Paiva Junior, A. Delmonte, I. Demedts, K. Deschepper, J. Dias, C. Dooms, B. Duchemann, C. Dutra, H. Duvivier, E. Eigendorff, V. Ernani, M. Faehling, L. Faria, A. Fedenko, H. Fernando, R. Ferrara, V. Ferrari, G. Finley, P. Fix, M. Flores, S. Fourie, F. Franke, K.-P. Frohling, M. Furqan, C. Gal, R. Galamaga, D. Ganea, A. K. Ganti, S. Gao, M. Garassino *et al.*, “Neoadjuvant pembrolizumab plus chemotherapy followed by adjuvant pembrolizumab compared with neoadjuvant chemotherapy alone in patients with early-stage non-small-cell lung cancer (keynote-671) : a randomised, double-blind, placebo-controlled, phase 3 trial,” *The Lancet*, vol. 404, n°. 10459, p. 1240–1252, 2024. [En ligne]. Disponible : [https://doi.org/10.1016/S0140-6736\(24\)01756-2](https://doi.org/10.1016/S0140-6736(24)01756-2)
- [247] “A study of neoadjuvant chemotherapy plus nivolumab versus neoadjuvant chemotherapy plus placebo, followed by surgical removal and adjuvant treatment with nivolumab or placebo for participants with surgically removable early stage non-small cell lung cancer,” 2023. [En ligne]. Disponible : <https://tinyurl.com/38rnbx7a>
- [248] S. W. Kim, J. Roh et C. S. Park, “Immunohistochemistry for pathologists : Protocols, pitfalls, and tips,” *J Pathol Transl Med*, vol. 50, n°. 6, p. 411–418, 2016.
- [249] A. Moran et M. E. Daly, “Surveillance imaging for non-small cell lung cancer : mounting evidence that less is more,” *Transl Lung Cancer Res*, vol. 8, n°. Suppl 4, p. S343–s346, 2019.
- [250] J. Yang, W. Shi, Z. Yang, H. Yu, M. Wang, Y. Wei, J. Wen, W. Zheng, P. Zhang, W. Zhao et L. Chen, “Establishing a predictive model for tumor mutation burden status based on CT radiomics and clinical features of non-small cell lung cancer patients,” *Transl Lung Cancer Res*, vol. 12, n°. 4, p. 808–823, 2023.
- [251] M. Tonneau, K. Phan, V. S. K. Manem, C. Low-Kam, F. Dutil, S. Kazandjian, D. Vanderweyen, J. Panasci, J. Malo, F. Coulombe, A. Gagné, A. Elkrief, W. Belkaïd, L. Di Jorio, M. Orain, N. Bouchard, T. Muanza, F. J. Rybicki, K. Kafi, D. Huntsman, P. Joubert, F. Chandelier et B. Routy, “Generalization optimizing machine learning to improve CT scan radiomics and assess immune checkpoint inhibitors’ response in non-small cell lung cancer : a multicenter cohort study,” *Frontiers in Oncology*, vol. 13, 2023. [En ligne]. Disponible :

- <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2023.1196414>
- [252] M. D. Hellmann, L. Paz-Ares, R. B. Caro, B. Zurawski, S.-W. Kim, E. C. Costa, K. Park, A. Alexandru, L. Lupinacci, E. d. l. M. Jimenez, H. Sakai, I. Albert, A. Vergnenegre, S. Peters, K. Syrigos, F. Barlesi, M. Reck, H. Borghaei, J. R. Brahmer, K. J. O’Byrne, W. J. Geese, P. Bhagavatheeswaran, S. K. Rabindran, R. S. Kasinathan, F. E. Nathan et S. S. Ramalingam, “Nivolumab plus ipilimumab in advanced non–small-cell lung cancer,” *New England Journal of Medicine*, vol. 381, n°. 21, p. 2020–2031, 2019. [En ligne]. Disponible : <https://www.nejm.org/doi/full/10.1056/NEJMoa1910231>
 - [253] H. J. W. L. Aerts, L. Wee, E. Rios Velazquez, R. T. H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. M. Rietbergen, C. R. Leemans, A. Dekker, J. Quackenbush, R. J. Gillies et P. Lambin, “Data from NSCLC-Radiomics,” 2019.
 - [254] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox et F. Prior, “The cancer imaging archive (tcia) : maintaining and operating a public information repository,” *J Digit Imaging*, vol. 26, n°. 6, p. 1045–57, 2013.
 - [255] J. Wang et H. Guo, “Automatic approach for lung segmentation with juxta-pleural nodules from thoracic CT based on contour tracing and correction,” *Comput Math Methods Med*, vol. 2016, p. 2962047, 2016.
 - [256] S. P. Primakov, A. Ibrahim, J. E. van Timmeren, G. Wu, S. A. Keek, M. Beuque, R. W. Y. Granzier, E. Lavrova, M. Scrivener, S. Sanduleanu, E. Kayan, I. Halilaj, A. Lenaers, J. Wu, R. Monshouwer, X. Geets, H. A. Gietema, L. E. L. Hendriks, O. Morin, A. Jochems, H. C. Woodruff et P. Lambin, “Automated detection and segmentation of non-small cell lung cancer computed tomography images,” *Nature Communications*, vol. 13, n°. 1, p. 3423, 2022. [En ligne]. Disponible : <https://doi.org/10.1038/s41467-022-30841-3>
 - [257] B. Norman, V. Pedoia et S. Majumdar, “Use of 2d u-net convolutional neural networks for automated cartilage and meniscus segmentation of knee mr imaging data to determine relaxometry and morphometry,” *Radiology*, vol. 288, n°. 1, p. 177–185, 2018. [En ligne]. Disponible : <https://pubs.rsna.org/doi/abs/10.1148/radiol.2018172322>
 - [258] M.-R. Amini, V. Feofanov, L. Pauletto, L. Hadjadj, E. Devijver et Y. Maximov, “Self-training : A survey,” 2023.
 - [259] G. Somepalli, M. Goldblum, A. Schwarzschild, C. B. Bruss et T. Goldstein, “SAINT : Improved neural networks for tabular data via row attention and contrastive pre-training,” *arXiv [cs.LG]*, 2021. [En ligne]. Disponible : <http://arxiv.org/abs/2106.01342>

- [260] A. Hassani, S. Walton, N. Shah, A. Abuduweili, J. Li et H. Shi, “Escaping the big data paradigm with compact transformers,” *arXiv [cs.CV]*, 2022. [En ligne]. Disponible : <http://arxiv.org/abs/2104.05704>
- [261] S. M. Lundberg et S.-I. Lee, “A unified approach to interpreting model predictions,” dans *Adv. Neural Inf. Process. Syst.*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et R. Garnett, édit. Curran Associates, Inc., 2017, p. 4765–4774.
- [262] Y. Xu, B. Wan, X. Chen, P. Zhan, Y. Zhao, T. Zhang, H. Liu, M. Z. Afzal, S. Dermime, S. N. Hochwald, P. Hofman, H. Borghaei, D. Lin, T. Lv, Y. Song et w. o. b. o. AME Lung Cancer Collaborative Group, “The association of PD-L1 expression with the efficacy of anti- PD-1/PD-L1 immunotherapy and survival of non-small cell lung cancer patients : a meta-analysis of randomized controlled trials,” *Translational Lung Cancer Research*, vol. 8, n^o. 4, p. 413–428, 2019. [En ligne]. Disponible : <https://tlcr.amegroups.org/article/view/31044>
- [263] J. Vallejo, H. Singh, E. Larkins, N. Drezner, B. Ricciuti, P. Mishra-Kalyani, S. Tang, J. A. Beaver et M. M. Awad, “Impact of increasing PD-L1 levels on outcomes to PD-1/PD-L1 inhibition in patients with nslc : A pooled analysis of 11 prospective clinical trials,” *The Oncologist*, vol. 29, n^o. 5, p. 422–430, 2024. [En ligne]. Disponible : <https://doi.org/10.1093/oncolo/oyae006>
- [264] M. Majem, M. Cobo, D. Isla, D. Marquez-Medina, D. Rodriguez-Abreu, J. Casal-Rubio, T. M. Bueno, R. Bernabé-Caro, D. P. Parente, P. Ruiz-Gracia, M. M. Arroyo et L. Paz-Ares, “Pd-(l)1 inhibitors as monotherapy for the first-line treatment of non-small-cell lung cancer patients with high pd-l1 expression : A network meta-analysis,” *J Clin Med*, vol. 10, n^o. 7, 2021.
- [265] Y. Wang, Z. Tong, W. Zhang, W. Zhang, A. Buzdin, X. Mu, Q. Yan, X. Zhao, H.-H. Chang, M. Duhon, X. Zhou, G. Zhao, H. Chen et X. Li, “FDA-approved and emerging next generation predictive biomarkers for immune checkpoint inhibitors in cancer patients,” *Frontiers in Oncology*, vol. 11, 2021. [En ligne]. Disponible : <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.683419>
- [266] D. B. Doroshow, W. Wei, S. Gupta, J. Zugazagoitia, C. Robbins, B. Adamson et D. L. Rimm, “Programmed death-ligand 1 tumor proportion score and overall survival from first-line pembrolizumab in patients with nonsquamous versus squamous nslc,” *Journal of Thoracic Oncology*, vol. 16, n^o. 12, p. 2139–2143, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1556086421023923>
- [267] N. John, V. Schlintl, T. Sassmann, J. Lindenmann, M. Fediuk, R. Wurm, P. Douschan, M. Zacharias, L. Kalson, F. Posch, G. Absenger, L. Brcic, P. J. Jost et A. Terbuch,

- “Longitudinal analysis of pd-l1 expression in patients with relapsed nslc,” *Journal for ImmunoTherapy of Cancer*, vol. 12, n°. 4, p. e008592, 2024. [En ligne]. Disponible : <https://jitc.bmj.com/content/jitc/12/4/e008592.full.pdf>
- [268] M. K. Moutafi, W. Tao, R. Huang, J. Haberberger, B. Alexander, S. Ramkissoon, J. S. Ross, K. Syrigos, W. Wei, L. Pusztai, D. L. Rimm et I. A. Vathiotis, “Comparison of programmed death-ligand 1 protein expression between primary and metastatic lesions in patients with lung cancer,” *J Immunother Cancer*, vol. 9, n°. 4, 2021.
- [269] Y. Zhang, R. Cheng, T. Ding et J. Wu, “Discrepancies in PD-L1 expression, lymphocyte infiltration, and tumor mutational burden in non-small cell lung cancer and matched brain metastases,” *Translational Lung Cancer Research*, vol. 13, n°. 12, p. 3590–3602, 2024. [En ligne]. Disponible : <https://tldr.amegroups.org/article/view/94723>
- [270] M. Rozenblit, R. Huang, N. Danziger, P. Hegde, B. Alexander, S. Ramkissoon, K. Blenman, J. S. Ross, D. L. Rimm et L. Pusztai, “Comparison of PD-L1 protein expression between primary tumors and metastatic lesions in triple negative breast cancers,” *J Immunother Cancer*, vol. 8, n°. 2, 2020.
- [271] S. Y. Jeong, J. Y. Jeong, C. M. Hong, J. H. Jeong, S.-W. Lee, B.-C. Ahn et J. Lee, “Association of tumor pd-l1 expression and texture features of¹⁸f-fdg pet in lung cancer,” *Journal of Nuclear Medicine*, vol. 60, n°. supplement 1, p. 149–149, 2019.
- [272] X. Zhao, Y. Zhao, J. Zhang, Z. Zhang, L. Liu et X. Zhao, “Predicting PD-L1 expression status in patients with non-small cell lung cancer using [18F]FDG PET/CT radiomics,” *EJNMMI Research*, vol. 13, n°. 1, p. 4, 2023. [En ligne]. Disponible : <https://doi.org/10.1186/s13550-023-00956-9>
- [273] R. Da-ano, G. Andrade-Miranda, O. Tankyevych, D. Visvikis, P.-H. Conze et C. C. L. Rest, “Automated pd-l1 status prediction in lung cancer with multi-modal pet/ct fusion,” *Scientific Reports*, vol. 14, n°. 1, p. 16720, 2024. [En ligne]. Disponible : <https://doi.org/10.1038/s41598-024-66487-y>
- [274] U. Leung, M. Gönen, P. J. Allen, T. P. Kingham, R. P. DeMatteo, W. R. Jarnagin et M. I. D’Angelica, “Colorectal cancer liver metastases and concurrent extrahepatic disease treated with resection,” *Ann Surg*, vol. 265, n°. 1, p. 158–165, janv. 2017.
- [275] M. Amoretti, C. Amsler, G. Bonomi, A. Bouchta, P. Bowe, C. Carraro, C. L. Cesar, M. Charlton, M. J. T. Collier, M. Doser, V. Filippini, K. S. Fine, A. Fontana, M. C. Fujiwara, R. Funakoshi, P. Genova, J. S. Hangst, R. S. Hayano, M. H. Holzscheiter, L. V. Jørgensen, V. Lagomarsino, R. Landua, D. Lindelöf, E. Lodi Rizzini, M. Macrì, N. Madsen, G. Manuzio, M. Marchesotti, P. Montagna, H. Pruys, C. Regenfus, P. Riedler, J. Rochet, A. Rotondi, G. Rouleau, G. Testera, A. Variola, T. L. Wat-

- son, D. P. van der Werf et ATHENA Collaboration, “Production and detection of cold antihydrogen atoms,” *Nature*, vol. 419, n°. 6906, p. 456–459, oct. 2002.
- [276] G. Y. M. Wong, N. Bhimani, B. Mol, C. Diakos, P. de Reuver, M. P. Molloy et T. J. Hugh, “Performance of prognostic models incorporating KRAS mutation status to predict survival after resection of colorectal liver metastases,” *HPB*, vol. 24, n°. 8, p. 1316–1325, 2022.
- [277] S. Teng, Y. E. Li, M. Yang, R. Qi, Y. Huang, Q. Wang, Y. Zhang, S. Chen, S. Li, K. Lin, Y. Cao, Q. Ji, Q. Gu, Y. Cheng, Z. Chang, W. Guo, P. Wang, I. Garcia-Bassets, Z. J. Lu et D. Wang, “Tissue-specific transcription reprogramming promotes liver metastasis of colorectal cancer,” *Cell Res*, vol. 30, n°. 1, p. 34–49, déc. 2019.
- [278] L. Li, M. Sun, J. Wang et S. Wan, “Chapter nine - multi-omics based artificial intelligence for cancer research,” dans *Cutting Edge Artif. Intell., Spatial Transcriptomics and Proteomics Approaches Analyze Cancer*, ser. Advances in Cancer Research, E. Madan, P. B. Fisher et R. Gogna, édit. Academic Press, 2024, vol. 163, p. 303–356.
- [279] B. Wu, J. Yang, Z. Qin, H. Yang, J. Shao et Y. Shang, “Prognosis prediction of stage IV colorectal cancer patients by mRNA transcriptional profile,” *Cancer Med.*, vol. 11, n°. 24, p. 4900–4912, 2022.
- [280] F. C. Staal, D. J. van der Reijd, M. Taghavi, D. M. Lambregts, R. G. Beets-Tan et M. Maas, “Radiomics for the prediction of treatment outcome and survival in patients with colorectal cancer : A systematic review,” *Clin. Colorectal Cancer*, vol. 20, n°. 1, p. 52–71, 2021.
- [281] E. Montagnon, M. Cerny, V. Hamilton, T. Derennes, A. Ilinca, M. E. A. Elforaici, G. Jabbour, E. Rafie, A. Wu, F. Perdigon Romero, A. Cadrin-Chênevert, S. Kadoury, S. Turcotte et A. Tang, “Radiomics analysis of baseline computed tomography to predict oncological outcomes in patients treated for resectable colorectal cancer liver metastasis,” *PLoS One*, vol. 19, n°. 9, p. e0307815, sept. 2024.
- [282] M. G. Lubner, N. Stabo, S. J. Lubner, A. M. del Rio, C. Song, R. B. Halberg et P. J. Pickhardt, “CT textural analysis of hepatic metastatic colorectal cancer : pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes,” *Abdom Imaging*, vol. 40, n°. 7, p. 2331–2337, oct. 2015.
- [283] S. Ye, Y. Han, X. Pan, K. Niu, Y. Liao et X. Meng, “Association of ct-based delta radiomics biomarker with progression-free survival in patients with colorectal liver metastases undergo chemotherapy,” *Front. Oncol.*, vol. 12, 2022.
- [284] H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haihe-Kains, D. Rietveld, F. Hoebbers,

- M. M. Rietbergen, C. R. Leemans, A. Dekker, J. Quackenbush, R. J. Gillies et P. Lambin, “Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach,” *Nat. Commun.*, vol. 5, n^o. 1, p. 4006, juin 2014.
- [285] S. Zhang et D. Metaxas, “On the challenges and perspectives of foundation models for medical image analysis,” *Med. Image Anal.*, vol. 91, p. 102996, 2024.
- [286] K. Alleman, E. Knecht, J. Huang, L. Zhang, S. Lam et M. DeCuyper, “Multimodal deep learning-based prognostication in glioma patients : A systematic review,” *Cancers*, vol. 15, n^o. 2, 2023.
- [287] D. Ho, I. B. H. Tan et M. Motani, “Predictive models for colorectal cancer recurrence using multi-modal healthcare data,” dans *Proc. Conf. Health, Inference, Learn.*, ser. CHIL ’21. New York, NY, USA : Association for Computing Machinery, 2021, p. 204–213.
- [288] K. Chaudhary, O. B. Poirion, L. Lu et L. X. Garmire, “Deep Learning-Based Multi-Omics integration robustly predicts survival in liver cancer,” *Clin Cancer Res*, vol. 24, n^o. 6, p. 1248–1259, oct. 2017.
- [289] N. Wesdorp, V. van Goor, R. Kemna, E. Jansma, J. van Waesberghe, R. Swijnenburg, C. Punt, J. Huiskens et G. Kazemier, “Advanced image analytics predicting clinical outcomes in patients with colorectal liver metastases : A systematic review of the literature,” *Surg Oncol.*, vol. 38, p. 101578, 2021.
- [290] V. Granata, R. Fusco, F. De Muzio, M. C. Brunese, S. V. Setola, A. Ottaiano, C. Cardone, A. Avallone, R. Patrone, S. Pradella, V. Miele, F. Tatangelo, C. Cutolo, N. Maggialelli, D. Caruso, F. Izzo et A. Petrillo, “Radiomics and machine learning analysis by computed tomography and magnetic resonance imaging in colorectal liver metastases prognostic assessment,” *Radiol. Med.*, vol. 128, n^o. 11, p. 1310–1332, nov. 2023.
- [291] X. Luo, H. Deng, F. Xie, L. Wang, J. Liang, X. Zhu, T. Li, X. Tang, W. Liang, Z. Xiang et J. He, “Prognostication of colorectal cancer liver metastasis by CE-based radiomics and machine learning,” *Transl Oncol.*, vol. 47, p. 101997, 2024.
- [292] J. Keyl, R. Hosch, A. Berger, O. Ester, T. Greiner, S. Bogner, J. Treckmann, S. Ting, B. Schumacher, D. Albers, P. Markus, M. Wiesweg, M. Forsting, F. Nensa, M. Schuler, S. Kasper et J. Kleesiek, “Deep learning-based assessment of body composition and liver tumour burden for survival modelling in advanced colorectal cancer,” *J. Cachexia Sarcopenia Muscle*, vol. 14, n^o. 1, p. 545–552, 2023.
- [293] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen et K. H. Maier-Hein, “nnU-Net : a self-configuring method for deep learning-based biomedical image segmentation,” *Nat. Methods*, vol. 18, n^o. 2, p. 203–211, févr. 2021.

- [294] Q. Wang, H. Nilsson, K. Xu, X. Wei, D. Chen, D. Zhao, X. Hu, A. Wang et G. Bai, “Exploring tumor heterogeneity in colorectal liver metastases by imaging : Unsupervised machine learning of preoperative CT radiomics features for prognostic stratification,” *Eur. J. Radiol.*, vol. 175, p. 111459, 2024.
- [295] M. Yang, H. Yang, L. Ji, X. Hu, G. Tian, B. Wang et J. Yang, “A multi-omics machine learning framework in predicting the survival of colorectal cancer patients,” *Comput. Biol. Med.*, vol. 146, p. 105516, 2022.
- [296] W. Lu, D. Fu, X. Kong, Z. Huang, M. Hwang, Y. Zhu, L. Chen, K. Jiang, X. Li, Y. Wu, J. Li, Y. Yuan et K. Ding, “Folfox treatment response prediction in metastatic or recurrent colorectal cancer patients via machine learning algorithms,” *Cancer Med.*, vol. 9, n^o. 4, p. 1419–1429, 2020.
- [297] S. L. Sun Tian, Fulong Wang et G. Chen, “Identification of two subgroups of folfox resistance patterns and prediction of folfox response in colorectal cancer patients,” *Cancer Invest.*, vol. 39, n^o. 1, p. 62–72, 2021, pMID : 33258714.
- [298] S. Amniouel et M. S. Jafri, “High-accuracy prediction of colorectal cancer chemotherapy efficacy using machine learning applied to gene expression data,” *Front. Physiol.*, vol. 14, 2024.
- [299] J. Kim, H. Kim, M.-S. Lee, H. Lee, Y. J. Kim, W. Y. Lee, S. H. Yun, H. C. Kim, H. K. Hong, S. Hannenhalli, Y. B. Cho, D. Park et S. S. Choi, “Transcriptomes of the tumor-adjacent normal tissues are more informative than tumors in predicting recurrence in colorectal cancer patients,” *J. Transl. Med.*, vol. 21, n^o. 1, p. 209, mars 2023.
- [300] S. Zheng, H. He, J. Zheng, X. Zhu, N. Lin, Q. Wu, E. Wei, C. Weng, S. Chen, X. Huang, C. Jian, S. Guan et C. Yang, “Machine learning-based screening and validation of liver metastasis-specific genes in colorectal cancer,” *Sci. Rep.*, vol. 14, n^o. 1, p. 17679, juill. 2024.
- [301] C. S. N. Lam, A. A. Bharwani, E. H. Y. Chan, V. H. Y. Chan, H. L. H. Au, M. K. Ho, S. Rashed, B. M. H. Kwong, W. Fang, K. W. Ma, C. M. Lo et T. T. Cheung, “A machine learning model for colorectal liver metastasis post-hepatectomy prognostications,” *Hepatobiliary Surg. Nutr.*, vol. 12, n^o. 4, 2022.
- [302] A. Z. Paredes, J. M. Hyer, D. I. Tsilimigras, A. Moro, F. Bagante, A. Guglielmi, A. Ruzzenente, S. Alexandrescu, E. A. Makris, G. A. Poultsides, K. Sasaki, F. N. Aucejo et T. M. Pawlik, “A novel machine-learning approach to predict recurrence after resection of colorectal liver metastases,” *Ann. Surg. Oncol.*, vol. 27, n^o. 13, p. 5139–5147, déc. 2020.
- [303] Y. Zhai, W. Bai, J. Zhou, Q. Dong et J. Zhang, “Effect of tumour size ratio on liver recurrence-free survival of patients undergoing hepatic resection for colorectal liver

- metastases,” *BMC Cancer*, vol. 22, n°. 1, p. 103, janv. 2022.
- [304] O. Moaven, T. E. Tavalara, C. D. Valenzuela, T. T. Cheung, C. U. Corvera, C. H. Cha, J. A. Stauffer, M. K. K. Niazi, M. N. Gurcan et P. Shen, “Machine learning models for predicting the outcomes of surgical treatment of colorectal liver metastases,” *J. Am. Coll. Surg.*, vol. 236, n°. 4, p. 884–893, avr. 2023.
- [305] E. Vorontsov, A. Tang, C. Pal et S. Kadoury, “Liver lesion segmentation informed by joint liver segmentation,” 2018. [En ligne]. Disponible : <https://arxiv.org/abs/1707.07734>
- [306] K. He, X. Zhang, S. Ren et J. Sun, “Deep residual learning for image recognition,” dans *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2016, p. 770–778.
- [307] F. Wang et H. Liu, “Understanding the behaviour of contrastive loss,” dans *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, p. 2495–2504.
- [308] S. Bakr, O. Gevaert, S. Echegaray, K. Ayers, M. Zhou, M. Shafiq, H. Zheng, W. Zhang, A. Leung, M. Kadoch, J. Shrager, A. Quon, D. Rubin, S. Plevritis et S. Napel, “Data for NSCLC radiogenomics (version 4),” The Cancer Imaging Archive, 2017, data set. [En ligne]. Disponible : <https://doi.org/10.7937/K9/TCIA.2017.7hs46erv>
- [309] C. Nadeau et Y. Bengio, “Inference for the generalization error,” *Mach. Learn.*, vol. 52, p. 239–281, 01 2003.
- [310] A. J. Vickers, B. van Calster et E. W. Steyerberg, “A simple, step-by-step guide to interpreting decision curve analysis,” *Diagn. Progn. Res.*, vol. 3, n°. 1, p. 18, oct. 2019.
- [311] N. Hollmann, S. Müller, K. Eggenberger et F. Hutter, “TabPFN : A transformer that solves small tabular classification problems in a second,” dans *Proc. 11th Int. Conf. Learn. Represent.*, 2023.
- [312] R. C. Pink, K. Wicks, D. P. Caley, E. K. Punch, L. Jacobs et D. R. F. Carter, “Pseudogenes : pseudo-functional or key regulators in health and disease?” *RNA*, vol. 17, n°. 5, p. 792–798, mars 2011.
- [313] P. Zajac, E. Schultz-Thater, L. Tornillo, C. Sadowski, E. Trella, C. Mengus, G. Iezzi et G. C. Spagnoli, “MAGE-A antigens and cancer immunotherapy,” *Front Med (Lausanne)*, vol. 4, p. 18, mars 2017.
- [314] J. L. Weon et P. R. Potts, “The MAGE protein family and cancer,” *Curr Opin Cell Biol.*, vol. 37, p. 1–8, sept. 2015.
- [315] T. Lu, S. Dang, R. Zhu, Y. Wang, Z. Nie, T. Hong et W. Zhang, “Adamts18 deficiency promotes colon carcinogenesis by enhancing β -catenin and p38MAPK/ERK1/2 signaling in the mouse model of AOM/DSS-induced colitis-associated colorectal cancer,” *Oncotarget*, vol. 8, n°. 12, p. 18 979–18 990, mars 2017.

- [316] X. Zhang, H. Zhang, C. Fan, C. Hildesjö, B. Shen et X.-F. Sun, “Loss of CHGA protein as a potential biomarker for colon cancer diagnosis : A study on biomarker discovery by machine learning and confirmation by immunohistochemistry in colorectal cancer tissue microarrays,” *Cancers (Basel)*, vol. 14, n°. 11, mai 2022.
- [317] M. Wang, Z. Miao, H. Cen, J. He et C. Wei, “Long non-coding RNA (LncRNA) CHROMR promotes the expression of the CNNM1 gene by adsorbing hsa-mir-1299 to obtain drug resistance in diffuse large B lymphoma cells,” *Transl Cancer Res*, vol. 11, n°. 5, p. 1362–1371, mai 2022.
- [318] S.-L. Kim, S. T. Lee, I. S. Min, Y. R. Park, J. H. Lee, D.-G. Kim et S.-W. Kim, “Lipo-calin 2 negatively regulates cell proliferation and epithelial to mesenchymal transition through changing metabolic gene expression in colorectal cancer,” *Cancer Sci*, vol. 108, n°. 11, p. 2176–2186, sept. 2017.
- [319] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov et S. Zagoruyko, “End-to-end object detection with transformers,” 2020. [En ligne]. Disponible : <https://arxiv.org/abs/2005.12872>
- [320] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár et R. Girshick, “Segment anything,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2304.02643>
- [321] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger et I. Sutskever, “Learning transferable visual models from natural language supervision,” 2021. [En ligne]. Disponible : <https://arxiv.org/abs/2103.00020>
- [322] R. Girdhar, A. El-Nouby, Z. Liu, M. Singh, K. V. Alwala, A. Joulin et I. Misra, “Imagebind : One embedding space to bind them all,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2305.05665>
- [323] O. Thawkar, A. Shaker, S.S. Mullappilly, H. Cholakal, R. M. Anwer, S. Khan, J. Laaksonen et F. S. Khan, “Xraygpt : Chest radiographs summarization using medical vision-language models,” 2023. [En ligne]. Disponible : <https://arxiv.org/abs/2306.07971>
- [324] N. Wesdorp, M. Zeeuw, D. van der Meulen, I. van ‘t Erve, Z. Bodalal, J. Roor, J. H. van Waesberghe, S. Moos, J. van den Bergh, I. Nota, S. van Dieren, J. Stoker, G. Meijer, R.-J. Swijnenburg, C. Punt, J. Huiskens, R. Beets-Tan, R. Fijneman, H. Marquering, G. Kazemier et on behalf of the Dutch Colorectal Cancer Group Liver Expert Panel, “Identifying genetic mutation status in patients with colorectal cancer liver metastases using radiomics-based machine-learning models,” *Cancers*, vol. 15, n°. 23, 2023.

- [325] S. La Vecchia, S. Fontana, I. C. Salaroglio, D. P. Anobile, S. Digiovanni, M. Akman, N. Jafari, M. Godel, C. Costamagna, C. Corbet, J. Kopecka et C. Riganti, “Increasing membrane polyunsaturated fatty acids sensitizes non-small cell lung cancer to anti-PD-1/PD-L1 immunotherapy,” *Cancer Lett*, vol. 604, p. 217221, sept. 2024.
- [326] P. G. Fonseca et H. D. Lopes, “Calibration of machine learning classifiers for probability of default modelling,” 2017. [En ligne]. Disponible : <https://arxiv.org/abs/1710.08901>

ANNEXE A PARAMÈTRES D'EXTRACTION DE CARACTÉRISTIQUES RADIOMIQUES

L'initiative de normalisation des biomarqueurs d'images (image biomarker standardisation initiative, IBSI) est une collaboration internationale indépendante qui vise à normaliser la manière dont les biomarqueurs d'images sont extraits des images médicales afin de permettre une analyse quantitative fiable et à haut débit des images, également connue sous le nom de radiomique. L'IBSI a établi des lignes directrices pour les rapports de radiomique et les nomenclatures afin de permettre la reproductibilité des résultats. Les lignes directrices de l'IBSI en matière de rapports couvrent plusieurs sujets, notamment : les patients, l'acquisition et le traitement de l'imagerie, la segmentation et le rapport sur les régions d'intérêt, ainsi que l'extraction des caractéristiques radiomiques.

Acquisition d'images

Pour chaque patient, les données d'imagerie consistaient en des examens de TDM avec contraste avant traitement. Les examens tomomodensitométriques ont été réalisés avec différents scanners à 16 et 64 multidétecteurs. Tous les examens tomomodensitométriques comprenaient au moins une phase veineuse portale couvrant la partie supérieure de l'abdomen. Les images avaient un volume transversal de 512×512 .

Dans les deux premières études, les images avaient une résolution axiale moyenne de $0,72 \text{ mm}^2$ (intervalle = $[0,56, 0,98] \text{ mm}^2$) et une épaisseur de coupe moyenne de $2,26 \text{ mm}$ (intervalle = $[0,80, 5,0] \text{ mm}$). Dans la dernière étude, les images avaient une résolution transaxiale moyenne de $0,75 \text{ mm}^2$ (intervalle = $[0,55, 0,98] \text{ mm}^2$) et une épaisseur de coupe axiale de $2,91 \text{ mm}$ (intervalle = $[0,80, 6,5] \text{ mm}$).

Traitement d'images

Toutes les images ont été rééchantillonnées dans toutes les études en voxels isotropes ($1 \times 1 \times 1$). Une interpolation linéaire a été appliquée pour le rééchantillonnage des images. Les valeurs d'intensité des voxels ont été converties en unités Hounsfield (HU). Dans le troisième objectif, étant donné la nature multi-institutionnelle de l'ensemble de données, une opération de fenêtrage supplémentaire limitant les HU à la fenêtre pulmonaire a été effectuée en utilisant les paramètres suivants : niveau de la fenêtre (-600), largeur de la fenêtre (1500).

Segmentation et région d'intérêt Dans toutes les études de cette thèse, une segmentation automatique ou semi-automatique des lésions a été réalisée en utilisant l'algorithme décrit dans chaque étude. Un radiologue expert a ensuite revu et ajusté les lésions.

Paramètres d'extraction des caractéristiques radiomiques

Les paramètres d'extraction des caractéristiques radiomiques couvrent différents aspects, notamment : le type d'image utilisé pour extraire les caractéristiques (image originale ou image filtrée) ; les classes de caractéristiques extraites (caractéristiques de forme 2D ou 3D, statistiques de premier ordre, caractéristiques texturales) et le nombre de caractéristiques extraites dans chaque classe ; les paramètres des classes de caractéristiques, les paramètres de filtrage.

Type d'images :

Dans les deux premières études, l'image originale a été utilisée pour extraire les caractéristiques radiomiques sans filtrage supplémentaire. Dans les troisième et quatrième objectifs, un laplacien de Gauss et une décomposition en ondelettes ont été appliqués aux caractéristiques. L'application des filtres dans les deux dernières études se justifie par l'hétérogénéité des ensembles de données utilisés. Le filtre Laplacien de Gauss permet d'éliminer le bruit des images en se concentrant sur les régions présentant des changements d'intensité importants. La décomposition en ondelettes permet de décomposer l'image en différentes bandes de fréquence, ce qui permet d'effectuer une analyse multi-résolution.

Sigma (filtre Laplacien de Gauss) :

Pour le filtre Laplacien de gauss, le paramètre sigma détermine le niveau de finesse de la texture qui est mis en évidence. Une faible valeur de sigma met l'accent sur les textures fines, qui impliquent des changements sur de courtes distances, tandis qu'une valeur de sigma élevée met en évidence les textures grossières, qui correspondent à des variations de niveaux de gris se produisant sur de plus grandes distances. Dans la troisième étude, les valeurs de sigma utilisées sont (0,1, 0,5, 1, 2, 3, 4 et 5), ce qui conduit à 651 caractéristiques radiomiques Laplaciens de gauss. Dans la quatrième étude, les valeurs sigma utilisées sont (0,1, 0,5, 1, 2, 3), ce qui donne 465 caractéristiques Laplaciens de gauss. La différence dans les valeurs choisies est liée au fait que la taille des lésions dans le dernier objectif est plus petite que dans le troisième, ce qui n'a pas permis d'utiliser des filtres avec un sigma supérieur à 3.

Niveau de décomposition en ondelettes : Une décomposition en ondelettes a été effectuée sur les trois axes de l'image en utilisant les paramètres suivants :

1. niveau de départ de l'élément (niveau de l'ondelette à utiliser comme premier ensemble de décompositions à partir duquel une signature est calculée) : 0 ;
2. niveau (nombre de niveaux de décompositions en ondelettes à partir desquels une signature est calculée) : 1

Classes des caractéristiques radiomiques :

Dans les études de cette thèse, les caractéristiques radiomiques ont été extraites à l'aide de Py-radiomics à partir de chaque type d'image (originale ou filtrée). Les caractéristiques peuvent être réparties dans les classes suivantes : forme, caractéristiques statistiques de premier ordre, et texture. Les caractéristiques texturales sont calculées à partir des matrices suivantes : Gray Level Co-occurrence Matrix (GLCM), Gray Level Dependence Matrix (GLDM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM).

À partir des images originales, 14 caractéristiques de forme 3D ont été extraites, 18 de premier ordre, 24 GLCM, 14 GLDM, 16 GLRLM, 16 GLSZM et 5 NGTDM. L'application de filtres entraîne une augmentation du nombre de caractéristiques texturales de premier ordre et dérivées de la matrice, mais pas du nombre de caractéristiques de forme.

La liste des caractéristiques de forme est :

1. Elongation
2. Flatness
3. LeastAxisLength
4. MajorAxisLength
5. Maximum2DDiameterColumn
6. Maximum2DDiameterRow
7. Maximum2DDiameterSlice
8. Maximum3DDiameter
9. MeshVolume
10. MinorAxisLength
11. Sphericity
12. SurfaceArea
13. SurfaceVolumeRatio
14. VoxelVolume

La liste des caractéristiques statistiques de premier ordre est :

1. 10Percentile
2. 90Percentile
3. Energy
4. Entropy
5. InterquartileRange
6. Kurtosis

7. Maximum
8. MeanAbsoluteDeviation
9. Mean
10. Median
11. Minimum
12. Range
13. RobustMeanAbsoluteDeviation
14. RootMeanSquared
15. Skewness
16. TotalEnergy
17. Uniformity
18. Variance

La liste des caractéristiques GLCM est :

1. Autocorrelation
2. ClusterProminence
3. ClusterShade
4. ClusterTendency
5. Contrast
6. Correlation
7. DifferenceAverage
8. DifferenceEntropy
9. DifferenceVariance
10. Id
11. Idm
12. Idmn
13. Idn
14. Imc1
15. Imc2
16. InverseVariance
17. JointAverage
18. JointEnergy
19. JointEntropy
20. MCC

21. MaximumProbability
22. SumAverage
23. SumEntropy
24. SumSquares

La liste des caractéristiques GLDM est :

1. DependenceEntropy
2. DependenceNonUniformity
3. DependenceNonUniformityNormalized
4. DependenceVariance
5. GrayLevelNonUniformity
6. GrayLevelVariance
7. HighGrayLevelEmphasis
8. LargeDependenceEmphasis
9. LargeDependenceHighGrayLevelEmphasis
10. LargeDependenceLowGrayLevelEmphasis
11. LowGrayLevelEmphasis
12. SmallDependenceEmphasis
13. SmallDependenceHighGrayLevelEmphasis
14. SmallDependenceLowGrayLevelEmphasis

La liste des caractéristiques GLRLM est :

1. GrayLevelNonUniformity
2. GrayLevelNonUniformityNormalized
3. GrayLevelVariance
4. HighGrayLevelRunEmphasis
5. LongRunEmphasis
6. LongRunHighGrayLevelEmphasis
7. LongRunLowGrayLevelEmphasis
8. LowGrayLevelRunEmphasis
9. RunEntropy
10. RunLengthNonUniformity
11. RunLengthNonUniformityNormalized
12. RunPercentage
13. RunVariance

14. ShortRunEmphasis
15. ShortRunHighGrayLevelEmphasis
16. ShortRunLowGrayLevelEmphasis

La liste des caractéristiques GLSZM est :

1. GrayLevelNonUniformity
2. GrayLevelNonUniformityNormalized
3. GrayLevelVariance
4. HighGrayLevelZoneEmphasis
5. LargeAreaEmphasis
6. LargeAreaHighGrayLevelEmphasis
7. LargeAreaLowGrayLevelEmphasis
8. LowGrayLevelZoneEmphasis
9. SizeZoneNonUniformity
10. SizeZoneNonUniformityNormalized
11. SmallAreaEmphasis
12. SmallAreaHighGrayLevelEmphasis
13. SmallAreaLowGrayLevelEmphasis
14. ZoneEntropy
15. ZonePercentage
16. ZoneVariance

La liste des caractéristiques NGTDM est :

1. Busyness
2. Coarseness
3. Complexity
4. Contrast
5. Strength

Discrétisation de l'image : Les valeurs de gris dans la région d'intérêt sont discrétisées avant l'extraction des caractéristiques. Le paramètre de largeur des bins permet de spécifier la taille des bins lors de la création d'un histogramme pour la discrétisation des niveaux de gris dans l'image. Dans nos études, la largeur des bins a été fixée à 25 car les études initiales avec des largeurs de bins et des résolutions d'image variables n'ont pas montré d'amélioration des résultats de la prédiction des biomarqueurs.

Standardisation des caractéristiques radiomiques : Dans chacune des études de la thèse, les caractéristiques radiomiques ont été standardisées en calculant la moyenne μ_{train} et l'écart type σ_{train} sur l'ensemble d'apprentissage, et en appliquant l'équation de normalisation présentée dans l'équation suivante :

$$x_{\text{std}} = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}}} \quad (\text{A.1})$$

avec x et x_{std} sont les valeurs originales et standardisées des caractéristiques, respectivement.

Distance par rapport au voxel voisin : Le paramètre de la distance au voisin est fixé à 1 dans toutes les études. Ce paramètre est pris en compte dans la définition des voxels voisins et affecte les classes de caractéristiques GLCM et NGTDM. Il spécifie la distance maximale à considérer entre le voxel central et le voisin.

Décalage du réseau de voxels : Ce paramètre est fixé à 0 dans toutes les études. Il représente la quantité ajoutée à l'intensité du niveau de gris lors du calcul des caractéristiques de premier ordre, notamment l'énergie, l'énergie totale et le RMS.

GLCM symétrique Ce paramètre booléen a été fixé à True dans toutes les études. Il indique si les cooccurrences doivent être évaluées dans deux directions par angle, ce qui permet d'obtenir une matrice symétrique.

Seuil de dépendance α Ce paramètre, utilisé dans le calcul de la dépendance dans le GLDM, a été fixé à 0 dans toutes les études. Deux voxels voisins avec des intensités de niveau de gris i et j sont considérés comme dépendants si $|i - j| < \alpha$.

ANNEXE B TRAITEMENT DES DONNÉES ET ENTRAÎNEMENT DES MODÈLES

1. Prédiction de CD73 à partir d'images TDM préopératoires

Sélection des hyperparamètres du modèle TabNet : L'optimisation des hyperparamètres du modèle a été réalisée à l'aide d'une grille de recherche dans une approche de validation croisée à cinq niveaux effectuée sur l'ensemble d'apprentissage. Les hyperparamètres de TabNet sont les suivants :

γ : Cet hyperparamètre représente le coefficient de réutilisation des caractéristiques dans les masques. Une valeur proche de 1 rend la sélection des masques moins corrélée entre les couches. Les valeurs sont comprises entre 1,0 et 2,0.

optimizer – *fn* : fonction d'optimisation

optimizer – *params* : paramètres de l'optimiseur, y compris le taux d'apprentissage *lr*

scheduler – *params* : paramètres du planificateur, y compris la taille de chaque étape et les valeurs gamma

scheduler – *fn* : la fonction de planification utilisée

n_{steps} : Le nombre d'étapes dans l'architecture du modèle

n_d : La largeur de la couche de prédiction de décision.

n_a : La largeur de l'intégration de l'attention pour chaque masque. D'après l'article, $n_d=n_a$ est généralement un bon choix. (par défaut = 8)

mask – *type* : la fonction de masquage utilisée pour sélectionner les caractéristiques.

L'espace des paramètres était le suivant :

1. γ : [1.3, 1.5]
2. *optimizer* – *params* : [*lr*=2e-2, *lr*=1e-2]
3. *scheduler* – *params* : [step size :50, gamma :0.9]
4. *optimizer* – *fn* : [Adam]
5. *scheduler* – *fn* : [StepLR]
6. n_{steps} : [3]
7. n_d : [8]
8. n_a : [8]
9. *mask* – *type* : [sparsemax]

Après la recherche sur la grille, les hyperparamètres sélectionnés sont les suivants : $\gamma = 1.3$, $lr = 0.02$, *schedulerparams* : step size = 50, gamma = 0.9, optimization function : Adam, scheduler function : StepLR, $n_{steps} = 3$, $n_d = n_a = 8$, mask type = sparsemax.

Sélection des hyperparamètres des modèles comparatifs : L'optimisation des modèles comparatifs a également été réalisée à l'aide d'une grille de recherche dans un schéma de validation croisée de cinq niveaux sur l'ensemble d'apprentissage.

Pour le classificateur XGBoost, l'espace de recherche était le suivant :

1. max depth : [3, 5, 6],
2. learning rate : [0.01],
3. subsample : (0.6, 1.0), step = 0.1,
4. colsample by tree : range(0.6, 1.0), step=0.1,
5. colsample by level : range(0.6, 1.0), step=0.1,
6. n estimators : [100]

Les hyperparamètres optimaux sélectionnés sont les suivants :

1. subsample : 0.6,
2. n estimators : 100,
3. max depth : 3,
4. learning rate : 0.01,
5. colsample by level : 0.8999999999999999,
6. colsample by tree : 0.8999999999999999

Pour le modèle SVM, l'espace de recherche était le suivant :

1. γ : [0.0001, 0.001, 0.005, 0.01, 0.1, 1, 10],
2. C : [0.1, 1, 10, 100]

Les hyperparamètres optimaux sélectionnés sont les suivants :

1. γ : 10
2. C : 10

Pour le modèle de régression logistique, l'espace de recherche était le suivant :

1. C : [0.001, 0.01, 0.1, 1, 10, 100, 1000]

L'hyperparamètre optimal sélectionné est le suivant : C : 0.001

Pour le modèle de forêt aléatoire, l'espace de recherche était le suivant :

1. n estimators = [(200, 2000), num = 10]
2. max features = ['auto', 'sqrt']

3. max depth = [(10, 110), num = 11]
4. min samples split = [2, 5, 10]
5. min samples leaf = [1, 2, 4]
6. bootstrap = [True, False]

Les hyperparamètres optimaux sélectionnés sont les suivants :

1. n estimators : 200,
2. min samples split : 2,
3. min samples leaf : 4,
4. max features : 'auto',
5. max depth : 30,
6. bootstrap : True

Sélection de la valeur seuil des courbes de Kaplan-Meier : Pour effectuer l'analyse de survie, les probabilités de classe TabNet, appelées rad-CD73, ont été dérivées et stratifiées en rad-CD73^{High} et rad-CD73^{Low} en fixant une valeur seuil égale au tertile inférieur (rad-CD73 > 0,362) sur la base de la distribution du score rad-CD73 (figure B.1), ce qui correspondait au seuil optimal proposé par le logiciel X-tile (0,383).

Cette approche a été adoptée dans toutes les études ultérieures, dans lesquelles nous avons évalué la distribution des scores prédits par les modèles et utilisé le logiciel X-tile pour dériver les seuils optimaux.

2. Prédiction du profil CD3/CMH-I à partir de la TDM préopératoire

Paramètre de suréchantillonnage synthétique des minorités : La technique de suréchantillonnage synthétique des minorités (SMOTE) a été appliquée de manière à obtenir un nombre égal d'instances dans chaque classe. Le suréchantillonnage n'a donc été appliqué qu'à la classe minoritaire avec un taux égal au rapport entre la classe majoritaire et la classe minoritaire (96/64).

Sélection des hyperparamètres de Multi-TabNet et des modèles comparatifs :

Pour chacun des cinq modèles TabNet entraînés pour former Multi-TabNet, une recherche de grille d'hyperparamètres a été effectuée sur l'ensemble d'entraînement à l'aide d'un schéma de validation croisée à cinq niveaux.

Comme pour l'objectif précédent, l'espace des paramètres pour chacun des cinq modèles TabNet était le suivant :

1. γ : [1.3, 1.5]
2. *optimizer - params* : [lr=2e-2, lr=1e-2]

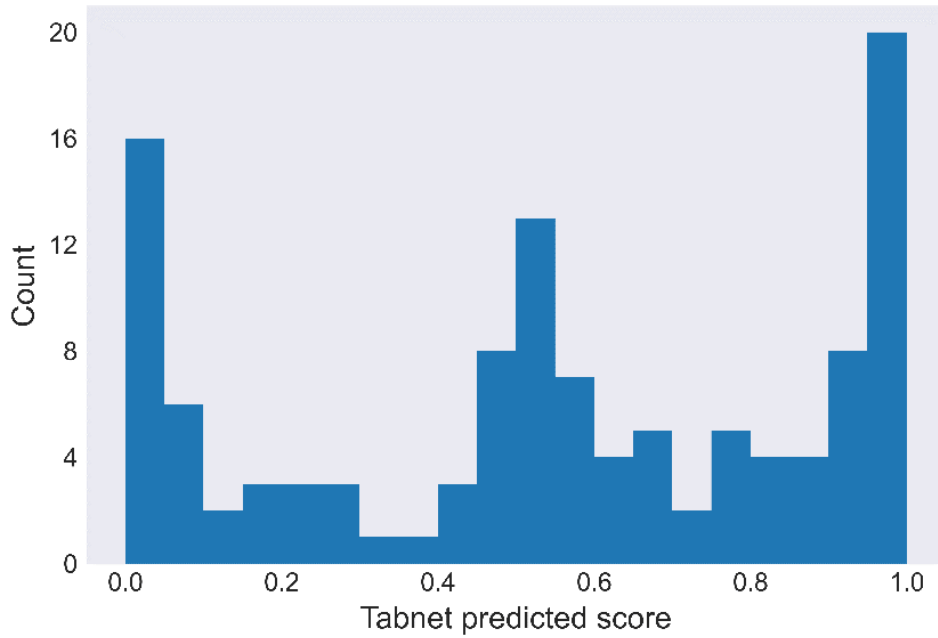


FIGURE B.1 Distribution du score probabiliste TabNet prédit, rad-CD73, parmi les patients. Une distribution trimodale a été observée pour rad-CD73. Le logiciel X-tile a retourné un point seuil basé de 0,383 pour cette distribution (proche de la valeur du premier tertile de 0,362).

3. *scheduler – params* : [step size :50, gamma :0.9]
4. *optimizer – fn* : [Adam]
5. *scheduler – fn* : [StepLR]
6. *n_steps* : [3]
7. *n_d* : [8]
8. *n_a* : [8]
9. *mask – type* : [sparsemax]

Pour chacun des modèles comparatifs, la même approche décrite dans le premier objectif pour sélectionner les hyperparamètres a été suivie. Une recherche en grille a donc été effectuée dans un schéma de validation croisée à 5 niveaux comme pour les modèles TabNet.

Pour le modèle XGBoost, les paramètres optimaux résultants étaient les suivants :

1. subsample : 0.7,
2. n estimators : 100,
3. max depth : 7,

4. learning rate : 0.02,
5. colsample by level : 0.7,
6. colsample by tree : 0.8999999999999999

Pour le modèle SVM, le paramètre optimal est le suivant : $C=0.1$

Pour le modèle de régression logistique, le paramètre optimal est le suivant : $C=100$

3. Prédiction de PD-L1 à partir d'images de TDM

FT - Paramètres du transformateur et des modèles comparatifs :

Les principaux paramètres réglables du FT-Transformer sont les suivants :

1. input dim (int, default : 64) : Ce paramètre représente le nombre de représentations utilisées pour coder chaque colonne. Il est également connu comme la dimension du modèle.
2. n heads (int, default : 8) : Ce paramètre représente le nombre de têtes d'attention par bloc de FT-Transformer.
3. n blocs (int, default : 4) : Ce paramètre représente le nombre de blocs du FT-Transformer.

Compte tenu de la complexité de calcul et des contraintes de mémoire associées au FT-Transformer, l'espace des paramètres a été restreint pour permettre d'entraîner le modèle sur le GPU disponible. Nous avons fait en sorte d'obtenir une représentation riche de la séquence d'entrée tout en respectant les contraintes de mémoire. À cette fin, la dimension d'entrée a été réduite à 32, le nombre de têtes à 2 et le nombre de blocs du FT-Transformer à 1. Les premières expériences réalisées en séparant un sous-ensemble de 20 % de l'ensemble d'apprentissage pour la validation ont montré que ces paramètres permettaient de maximiser la richesse de la représentation des caractéristiques d'entrée tout en tenant compte de la taille de la mémoire du GPU.

La même approche a été adoptée pour l'entraînement des modèles comparatifs. L'autre modèle à base de transformeurs, le Self-Attention and Intersample Attention Transformer (SAINT), a également été entraîné avec 2 têtes d'attention, 1 bloc transformeur et une taille de dimension d'entrée de 32.

Sous-typage histologique basé sur le séquençage de l'ARN

L'optimisation des modèles d'apprentissage automatique formés pour cet objectif a été réalisée à l'aide d'une technique de validation croisée à 5 niveaux imbriquée (nested 5-Fold cross-validation). L'espace des paramètres pour chaque modèle formé est décrit ci-dessous.

Le modèle XGBoost :

1. max depth : [3, 6, 10],
2. learning rate : [0.1, 0.3],
3. subsample : [0.9, 1.0],
4. n estimators : [100, 500]

Le modèle de forêt aléatoire : n estimators : [100, 200, 500]

Le modèle de l'arbre de décision : max depth : [3, 6, 10, None]

Le classificateur AdaBoost : n estimators : [50, 100, 200]

Le classificateur LGBM :

1. max depth : [3, 6, 10],
2. learning rate : [0.1, 0.3],
3. subsample : [0.9, 1.0],
4. n estimators : [100, 500]

Le classificateur CatBoost :

1. depth : [6, 10],
2. n estimators : [100, 200]

4. Prédiction de la récurrence basée sur la multiomique

Compte tenu de la dimensionnalité élevée des données de séquençage de l'ARN impliquant l'expression de 28578 gènes, associée à la complexité computationnelle des modèles basés sur les transformeurs, il a été nécessaire de réduire le nombre de paramètres du modèle FT-Transformer afin de respecter les contraintes de mémoire tout en maximisant le potentiel du modèle dans l'encodage de représentations riches des valeurs d'expression des gènes.

Le nombre de blocs du FT-Transformer et de têtes d'attention a donc été limité à 1 et la dimension d'entrée a été réduite à 8. En outre, le facteur de compression kv du FT-Transformer a été utilisé et fixé à 0,05. Le facteur de compression est utilisé dans le mécanisme d'attention du FT-Transformer pour réduire la longueur de la séquence d'entrée. La longueur de la séquence résultante est calculée comme $k = kv_{facteur-de-compression} \times s$ où s est la longueur de la séquence d'entrée.

La même stratégie a donc été suivie pour les modèles comparatifs basés sur les transformeurs. Les paramètres du modèle TabTransformer ont donc été sélectionnés comme suit :

1. input dim : 8
2. n blocks : 1

3. *n heads* : 1

Les paramètres de TabNet ont également été choisis en utilisant la même stratégie :

1. γ : 1.3
2. *optimizer – params* : lr=1e-2
3. *optimizer – fn* : Adam
4. *n_steps* : 3
5. *n_d* : 8
6. *n_a* : 8
7. *mask – type* : sparsemax

Traitement des données :

Normalisation des features de séquençage de l'ARN :

Pour la normalisation des données d'expression génique, nous avons appliqué une technique de normalisation min-max afin de garantir que les valeurs d'expression des différents gènes se situent dans la plage requise par le FT-Transformer pour l'étape de tokenisation des caractéristiques.

Dans l'approche de validation croisée à cinq niveaux, les paramètres de normalisation $\min(x_{\text{train}})$ et $\max(x_{\text{train}})$ ont été calculés sur les instances des folds d'entraînement et appliqués à toutes les instances. La version normalisée x' d'une caractéristique x est donc :

$$x' = \frac{x - \min(x_{\text{train}})}{\max(x_{\text{train}}) - \min(x_{\text{train}})} \quad (\text{B.1})$$

Standardisation des caractéristiques d'imagerie :

Pour toutes les caractéristiques d'imagerie (caractéristiques radiomiques et caractéristiques d'imagerie dérivées du modèle de fondation), nous avons effectué une standardisation des caractéristiques en calculant la moyenne et l'écart type des caractéristiques sur l'ensemble d'entraînement et en l'appliquant aux instances de l'ensemble d'entraînement et de test.

Pour l'extraction des caractéristiques à l'aide du modèle de fondation, nous avons recadré les régions d'intérêt de taille $50 \times 50 \times 50$. Le choix a été basé conjointement sur la taille des lésions et l'architecture du modèle de fondation.