**Assessing Clinical-Radiomic Machine Learning Model for Predicting Recurrence of Colorectal Liver Metastases**

For: HINF-4902 Capstone Seminar

Master’s in Health Data Informatics and Analytics

University of Denver, University College

November 17th, 2024

Faculty: Dr. Nicoletta Weide, PhD, CAT, HS-BCP

Academic Director: Dr. Jeffrey Weide, DBA, FACHE, PMP, M.S.

Dean: Michael J. McGuire, M.L.S

**Abstract**

The prevalence of colon cancer (CC) in the United States along with a large percentage of CC patients forming liver metastasis is a concerning matter. Utilizing non-invasive measures to deliver accurate prognosis for patients experiencing colorectal liver metastasis (CRLM) can contribute to better outcomes for those seeking curative treatments. Computed tomography (CT) images enhanced by contrast are a first line method for medical practitioners to understand progression of disease and visualize a future liver remnant (FLR). The Cancer Imaging Archive has recently released clinical data and CT imaging of CRLM patients for the purpose of applying machine learning models to predict clinical outcomes such as recurrence. Using PyRadiomics, 1475 radiomic features were extracted from 429 lesions. Hybrid predictive modeling demonstrated comparable prediction of recurrence rate stratification by Fong Scoring system utilizing ML algorithms such as Random Forest, Multiple Logistic Regression, Support Vector Machines, and K-Nearest Neighbors.

**Table of Contents**

**Background……………………………………………………………………………………………………………………………2**

Importance of the Problem……………………………………………………………….……………………………………….2

Problem Statement…………………………………………………………………….….………………………………………….3

Thesis Statement..…………………………………………………………………….……………………………………………….4

**Approach…….……………………………………………………………………….……………………………………………….4**

Data Collection and Validation……………………………..…………………..……………………………………………….4

Study Design……………………………………………………………………………..……………………………………………….5

Feature Extraction and Image Preprocessing…………………………………..…………………………………………6

Feature Harmonization.…..……………………………………………………..…………………………………………………7

Clinical Variables……………………………………………………………………….…….…………………………………………7

Model Development……………………………………..………………..…………………………………………………………8

**Literature Review…………………………..………………..……………………………………………………………………8**

CRLM Treatment Management……………………….……………….………………………………………………………11

Machine Learning Versus Traditional Statistics………….………………….…………………………………………14 Hybrid ML Models……………………………………………………………..…………………………………………………….20

Recurrence Risk Stratification………………………………………………………………………………………………….23

**Solution…………………………………………………………….…………….………………………………………………….24**

Data Understanding…………………………………………………………………………………………………………..…….24

Data Preparation……………………………………………………………………………………………………………………..25

Model Performance(s)…………………………………….……………………………………………………………………….27

**Discussion……………………………………….………………………………………………………………………………….28**

**Recommendations……………………………………………………………………………………………………………….30**

**Conclusion…………………………………………………………………………………………..…………………………….32**

**References…………………………………………………………………………………………………………………………34**

**Appendix……………………………………………………………………………………………………………………………43**

**Background**

Importance of the Problem

Colorectal cancer is one of the most prevalent cancer diagnoses in the United States, ranking third as a cancer site for both men and women alike (National Cancer Institute Cancer Cases, 2024) . It is also the second leading cause of cancer death for both men and women alike per the American Cancer Society (ACS) (Cancer Facts & Figures, 2024). For patients who develop primary colorectal cancer, studies have indicated that anywhere from 25% to at least 50% will also develop liver metastasis due in part that the portal vein circulates blood from the lower intestines to the liver before recirculation to the rest of the body (Hu, Pan, and Chen 2021).

Recurrence of cancer continues to plague cancer patients and survivors and attempts to predict recurrence have not taken the forefront compared to predicting curative statistics such as overall survival or disease-free survival. Evidently so, more than 59% of cancer patients and survivors report a moderate level of fear when discussing recurrence (Luigjes-Huizer et al. 2022). For enhancing patient quality of care by providing accurate risk of recurrence, proper resource allocation for low or high-risk patients can be streamlined for practitioners and MDT’s. Predictive modeling using recurrence risk will also contribute to determining treatment modality for patients instead of relying heavily on initial presentation. These predictive models hinge on the application of AI and its subset, machine learning on health information to train predictive capabilities.

The exploration of machine learning to predict patient outcomes has grown over the past decade and has branched into utilizing not just clinical data, but leveraging data from pathomics, genomics, and radiomics. Radiomics in particular has grown within the health informatics community as a possible tool to glean informative features from imaging and predict disease presence, pathological grading of tumors, staging of tumors, or even predict presence of tumors non-invasively. Due to the growing interest in artificial intelligence and machine learning in the medical field, medical practitioners have aimed to aggregate deidentified data for health data scientists and informatic specialists. The release of public datasets on cancer data particularly by the NIH Cancer Imaging Archive is framed for machine learning model development and open-source software like Pythons Pyradiomics library or 3D Slicer make building radiomic ML models widely accessible. As per Griethuysen et. al, 2017, “Radiomics aims to quantify phenotypic characteristics on medical imaging through the use of automated algorithms. Radiomic artificial intelligence (AI) technology, either based on engineered hard-coded algorithms or deep learning methods, can be used to develop non-invasive imaging-based biomarkers” (van Griethuysen et al., 2017). But the standardization and application of radiomic machine learning towards CRLM recurrent risk stratification has not been clearly defined or readily applied to justify its inclusion in assisting clinical decision making.

Problem Statement

Colorectal liver metastasis is a significant contributor to cancer incidence for both men and women in the United States, presenting an area of emphasis for practitioners to provide proven curative treatments to maximize disease free survival and minimize recurrence risk. As per Kawaguchi et al (2023), greater recurrence rate of liver metastasis has been stratified for recurrence risk has developed more scrutinous surveillance algorithms but may result in decreasing patient quality of life measure by pandering increased fear and anxiety (Kawaguchi et al, 2023). CRLM is extremely complicated when maximizing curative intent due to sensitivity to individual patient factors, need for multidisciplinary consult, and a great range of differing treatment options ranging from disease mitigation to curative attempts. Therefore developing predictive models to accurately identify recurrence risk will assist MDTs in efficient clinical decision making, accurately inform patients of recurrence risk associated with surgical resection, and contribute in proper allocation of perioperative treatment to enhance curative cancer care for eligible CRLM patients (Simpson et al, 2024).

Thesis Statement

By wielding machine learning algorithms upon a combination of radiomic features and clinical information for CRLM patients prior to liver resection, accurate prediction for stratified recurrence risk using the Fong Score criterion can be utilized to assist MDTs and patients alike, thus enhancing patient selection for proper treatment modality such as surgical resection.

**Approach**

Data Collection and Validation

The public release of CT data pre-operatively along with clinical information for 197 CRLM patients was collected for the application of machine learning models to predict disease free survival, hepatic recurrence, and determine radiomic signatures (features of interest) for the anatomical region of interest. This data is provided by the Cancer Imaging Archive and funded under the National Health Institution, making it a credible source for properly formatted, selective, and de-identified data that follow HIPAA compliances. Given access to this database, CRISP-DM stages of data understanding, data preparation, modeling, and evaluation were followed to ensure reproducibility of approach.

The inclusion criteria for patients were that each patient demonstrates visible tumors on the liver preoperative CT scan that was obtained 6 weeks prior to liver resection and pathologically confirmation of resection. Pathological data of healthy liver tissue (parenchyma) and hepatic tumors must also be included for each patient. The exclusion criteria focused on ensuring patients had greater than 90-day mortality, were followed up on at least in 24 months, and did not receive hepatic artery infusion chemotherapy which can alter CT imaging and pathology. The last exclusion criteria ensured patients had not received local tumor ablation or greater than 3 wedge resections. This data is prepared to ensure that no patient conditions affect radiological features or alter pathology for patients while survival data is collected accordingly.

Study Design

The approach to analyzing recurrence risk for CRLM was developing a retrospective cohort study. Secondary academic research was conducted on the given clinical dataset combined with radiomic features of tumors extracted from CT scans to develop a range of predictive machine learning models. The machine learning algorithms that excel as classifiers to predict probability of recurrence are Random Forest Classifiers, Multiple Logistic Regression, K-Nearest Neighbor, and Support Vector Machines. The target class for prediction in the clinical data is stratification of risk recurrence scored according to comparable hazard ratios indicated by the Fong Score (Fong et al, 199), which was crafted utilizing the clinical dataset. Models were then evaluated with a scaled model score utilizing their performance metrics such as accuracy, precision, recall, F1-score, sensitivity, PPV, NPV, and AUC where applicable. Overall study design followed CRISP-DM methodology to maximize reproducibility of predictive models.

Feature Extraction and Image Preprocessing

Utilizing open-source software to first extract tumor masks was the first step for extracting radiomic features. 3D Slicer is an open-source software that allows for visualization of DICOM files holding CT information for patients and their respective segmentations. 3D Slicer is recommended by the publisher of the CRLM database, as segmentations were created by the use of such software and are extremely utilized in the radiomic community. Existence of tumors and review of segmentations were reviewed by the publisher's radiology team. 3D Slicer also allows for visualization of DICOM SEG files that were provided by TCIA, so visual confirmation of liver, FLR, hepatic arteries, and tumors were manually committed. Tumors were isolated and converted from DICOM metadata to three dimensional masks. Lastly, quantitative features for tumor masks were tabulated for cross validation after feature extraction.

Python along with respective libraries such as Pandas, Radiomics, SimpleITK and PyDicom was utilized for feature extraction of radiomic features. Python is an open-source software that offers extensive libraries ranging from data science, image preprocessing, and specifically radiomic analysis. File structure for DICOM data was confirmed to hold all 197 patients CT imaging and extracted tumor segmentation masks from 3D Slicer. DICOM data was accessed using PyDicom to ensure proper metadata and matching to respective three-dimensional tumor segmentation masks. Feature extraction was conducted with customizable parameter files and standardized by ISBI recommendations within PyRadiomic’s feature extractor function.

Feature Harmonization and Selection

Feature harmonization was conducted upon radiomic data set using Python’s ComBat. Z-score normalization was applied to standardize a wide range of numerical values in radiomic data set and aggregation to average tumor burden was finalized to ensure feature cohesion. Due to large variability within radiomic features, dimensionality reduction was committed in a two-step process. Multicollinearity of features was addressed with variance inflation factor measuring and feature set reduced from original 1475 to 489 radiomic features. Further dimensionality reduction was conducted with Principal Component Analysis to capture greatest variance regarding recurrence risk stratification multi classification.

Clinical Variables

Imputation, normalization, and creation of additional clinical variables were done using the clinicopathological data set provided by the Cancer Imaging Archive. One hot encoding for respective categorical features along with the transformation of percentage-based features to floats was initiated. Imputation was conducted to replace missing values and retain maximum patient characteristics. Z-score normalization conducted on clinicopathological data to prepare for merge with radiomic data. Categorical columns of multiple metastases present and largest tumor diameter greater than five centimeters created using present patient data to follow recurrence risk stratification by Fong. Lastly, radiomic data and clinicopathological data were joined into a single dataset to prepare for model development.

Model Development

Python offers reputable libraries for machine learning development, specifically SciKit Learn and Pytorch, which are staples in the toolkits of data scientists seeking to create machine learning models. SciKit Learn offers feature selection methods ranging from dimensionality reduction to K Best selection, model foundations for linear and non-linear classifiers ranging from Random Forest to Support Vector Machines, and metric reporting tools ranging from confusion matrix to classification reporting. Each model was coded into Python with SciKit Learn and scores were reported utilizing SciKit Learn. Model scores were placed in the scaled model scoring equation to determine overall predictive ability of each machine learning algorithm employed.

**Literature Review**

CRLM Treatment Management

Colorectal liver metastases (CRLM) pose significant challenges in both curative treatment and overall management. Overcoming these barriers necessitates comprehensive strategies and collaboration among various medical specialties. Nearly 20% of patients will have liver metastases upon discovery of CRC, meaning management of primary CRC is delicately balanced with surgical consideration of treating liver metastases (Simpson et al. 2024). Given the great prevalence of colon cancer and a significant percentage forming liver metastasis, the efforts to demonstrate effective curative treatments maximizing disease free survival and minimizing recurrence risk is heavily emphasized in the oncological field.

Treatment course is extremely case by case dependent as qualification for surgery is dependent on tumor size, number of tumors, and bilobar location of tumors when attempting to create a future liver remnant (FLR)(Kathleen De Greef et al. 2016). CRLM patients at initial presentation are usually separated into three groups prior to selection of treatment modality as those with initially resectable cases, resectable of metastases after conversion therapy, and those who will never be resectable (Kathleen De Greef et al. 2016). Treatment options for patients have progressively trended towards curative options such as hepatic resection which has maximized 5 year relapse-free survival rates for 35-60% of resected patients, especially since neoadjuvant chemotherapy has been shown to downsize tumors in preparation for successful resection (Kathleen De Greef et al. 2016). Yet best treatment courses are stringent upon a multidisciplinary approach to determining the treatment modality for greatest disease-free interval along with overall survival after initial presentation (Clark and Smith 2014).

Treatment modality is assessed by a multidisciplinary team of interventional and diagnostic radiologists, surgeons, oncologists, and gastroenterologists for CRLM patients. The benefits of a multidisciplinary team (MDT) range from increased quality of care for cancer patients, but can only be organized in large centers, over burdening practitioners when organizing, and are limited by information exchange channels for communicating during the care process (Patkar et al., 2011) (Berardi et al. 2020). As such, the call for clinical decision support technology to leverage the benefits of MDT’s while minimizing the burden for practitioners is key (Berardi et al. 2020). Practitioners describe these computer programs as integrating pertinent medical history for patients to create valid predictive models and apply precision medicine in a practical way for guiding medical decision making for these teams (Patkar et al., 2011) (Berardi et al. 2020). Predictive models range from focusing on maximizing disease free survival statistics, automating clinical risk values, or even evaluating accuracy of differential diagnoses with advanced neural networks.

Colorectal liver metastases are often complex, leading to treatment challenges in achieving curative outcomes. Traditionally, patients identified with CRLM face limited treatment options post-diagnosis, primarily due to the aggressive nature of the disease and the intrahepatic spread of metastatic lesions (Hu, Pan, and Chen 2021)(Kathleen De Greef et al. 2016). The surgical approach to removing metastases, particularly hepatic resection, remains the only potentially curative strategy. However, the selection criteria for resection are evolving to improve patient outcomes (Kathleen De Greef et al. 2016). Continuous variations in disease patterns dictate that patients diagnosed must be evaluated meticulously regarding resectability, featuring metrics such as tumor burden, liver function, and presence of extrahepatic disease (Kathleen De Greef et al. 2016).

Guidelines advocate a more nuanced approach in selecting candidates for hepatic resection. The criteria for resectability have expanded significantly, allowing for greater surgical intervention from previously deemed unresectable statuses (Kathleen De Greef et al. 2016). Joint efforts of surgical teams, oncologists, and radiologists are vital to discern treatment pathways that maximize the potential for prolonged survival while minimizing recurrence risks (Hu, Pan, and Chen 2021). For instance, classifications consider molecular biomarkers like KRAS and NRAS mutations as critical factors influencing treatment planning and patient selection, indicating not just the need for surgical intervention but the effectiveness of targeting biological agents (Hu, Pan, and Chen, 2021) (Kathleen De Greef et al. 2016).

Central to enhancing patient-centered care in CRLM treatment is the establishment of multidisciplinary teams (MDTs). These teams leverage diverse expertise—from colorectal surgeons to oncologists, and interventional radiologists—to construct comprehensive treatment plans tailored for individual patient needs (Kathleen De Greef et al. 2016). Collaborative consultations allow for seamless transitions from systemic therapy to surgical intervention, complemented by monitoring and follow-up through clinical nurse specialists, crucial for addressing patient concerns holistically (Hu, Pan, and Chen, 2021) (Kathleen De Greef et al. 2016). Studies demonstrate that MDT management leads to improved resection rates and survival rates compared to single-discipline focused approaches (Kathleen De Greef et al. 2016).

The growing utilization of machine learning (ML) models in imaging offers promising advancements in identifying liver metastases more accurately. These prognostic ML models have shown considerable potential in stratifying patients based on imaging studies, effectively predicting those who may benefit from surgical resection interventions (Kambakamba et al. 2021)(Hu, Pan, and Chen 2021). The integration of such predictive analytics can significantly assist multidisciplinary teams in devising comprehensive patient management strategies, optimizing outcomes by pairing surgical options with effective perioperative treatments based on personalized risk assessments.

Machine Learning Versus Traditional Statistics

The integration of machine learning (ML) in healthcare represents a significant evolution in data analysis and predictive modeling. Although traditional statistical methods have been the bedrock of medical research, advances in ML promise enhanced prediction capabilities and the potential for personalized medicine. This section delineates the advantages and disadvantages of ML versus traditional statistical models, their differing reception within the medical community, and the potential that ML continues to demonstrate in cancer research.

Machine learning algorithms excel in handling vast datasets and can discover complex patterns among numerous predictors, often surpassing human capabilities in predictive accuracy. For instance, ML can process millions of patient records, honing its predictive power by analyzing extensive features, which enabled sophisticated diagnoses and treatment predictions without being explicitly programmed for specific relationships(Rajula et al. 2020). Furthermore, ML configurations such as ensemble methods or neural networks allow for exploration of interpersonal and interactive relationships that are more challenging to delineate using conventional methods (Rajula et al. 2020).

In oncology, specific studies have indicated ML's prowess in yielding precise prognosis models. For example, ML techniques have successfully predicted patient outcomes in various cancers, thus marking a paradigm shift in treatment strategies and clinical guidelines (Rajula et al. 2020).

Despite its advantages, several drawbacks impede the widespread acceptance of ML in healthcare. The most persistent issue is interpretability; ML models can produce results that are challenging to comprehend or explain clinically, particularly for methods like deep learning. This opacity creates hesitance among clinicians who prefer models offering clear "clinician-friendly" outputs, such as odds ratios (Rajula et al. 2020). In contrast, traditional statistical models typically present clear metrics that are reliant on a limited number of variables, making results straightforward to interpret within established clinical contexts (Rajula et al. 2020)(Ali et al. 2024).

Another notable concern is the propensity for overfitting ML algorithms, which arises when models learn noise in the training dataset rather than the intended predictive patterns, potentially limiting their applicability to unseen data (Rajula et al. 2020). Furthermore, developing and validating ML models often requires significant computational resources and expertise in algorithm selection and optimization, which can be a barrier to implementation (Rajula et al. 2020)(Ali et al. 2024).

The reception of machine learning within the medical sector has evolved, characterized by both enthusiasm and skepticism. Clinicians acknowledge the potential improvements in diagnostic accuracy and treatment personalization; however, they also express concerns regarding the interpretability of ML outputs. A lack of standardized processes for validating ML methodologies heightens hesitance (Rajula et al. 2020). Overall, while many healthcare professionals view ML as a powerful tool, integrating it into clinical practice will require improved explanation, infrastructural support, and an ongoing dialogue regarding ethical implications (Rajula et al. 2020).

Numerous studies have documented the successful application of ML techniques in cancer research, underscoring its utility in identifying tumor characteristics, predicting treatment responses, and personalizing care plans. Moreover, recent research suggests that ML could facilitate drug discovery pathways and optimize clinical trial designs (Rajula et al. 2020). Its flexibility allows researchers to handle varied data types, integrating imaging results with genetic and clinical information that traditional statistical methods struggle to encompass (Rajula et al. 2020) (Ali et al. 2024).

Machine learning illustrates advanced analytical capabilities that traditionally established statistical methods cannot replicate, particularly concerning data volume and complexity. However, challenges surrounding interpretability, the risk of overfitting, and resource expenditures remain pivotal issues that must be addressed for the successful clinical application of ML. In conclusion, ML continues to exhibit substantial promises in enhancing cancer research outcomes, pushing the boundaries of personalized medicine, and reshaping the diagnostic landscape. Nonetheless, nurturing both ML and traditional statistical approaches might yield the most fruitful collaborations, ensuring robust, interpretable, and clinically applicable predictions (Rajula et al. 2020).

Radiomics

Radiomics has emerged as a potent tool in medical imaging, enabling the extraction of a multitude of features from images that can characterize tumors' phenotypic traits. These quantitative aspects hold significant potential for assisting clinicians in diagnosis, treatment planning, and predicting patient outcomes. Despite its promise, the field faces challenges regarding standardization and reproducibility.

Radiomics utilizes high-throughput computing techniques to derive meaningful features from medical imagery such as CT, MRI, and PET. It enhances personalized medicine by producing quantitative image features that may be associated with clinical outcomes, tumor biology, and treatment efficacy (van Griethuysen et al. 2017). By offering an automated approach to quantify data, radiomics serves as an efficient avenue for uncovering relationships between imaging and malignancy, aiming to optimize clinical decision-making (van Griethuysen et al. 2017).

Efforts to standardize radiomic feature extraction and processing have gained momentum, notably through the establishment of the Image Biomarker Standardization Initiative (IBSI) (Zwanenburg et al. 2020). IBSI objectives include defining consistent nomenclature for radiomic features, creating standard processing protocols, and developing a repertoire of reference values for validation across different software (Zwanenburg et al. 2020). A recent study demonstrated that out of 174 radiomic features, a consensus was achieved for 169, reflecting a strong commitment to improving reproducibility (Zwanenburg et al. 2020). Importantly, these efforts facilitate inter-institutional studies and comparisons by minimizing variability attributed to software discrepancies (Zwanenburg et al. 2020).

An integral component of the standardization movement is the development of open-source software platforms like PyRadiomics, which enable researchers and clinicians to consistently extract features from images. Implemented in Python, PyRadiomics is designed to be user-friendly, offering integration with 3D Slicer—an open-source medical image computing platform (Zwanenburg et al. 2020). The wide availability and extensibility of PyRadiomics not only enhances its adoption in research settings but also ensures that various users can validate and replicate their studies efficiently. Its reliability in extracting pertinent image features has been confirmed, further solidifying the case for its use in clinical radiomics (Zwanenburg et al. 2020).

The application of the ISBI standards within the development of clinical-radiomic machine learning (ML) models underpins the movement towards reproducibility. By establishing general schemes for radiomics image processing, ISBI aims to assure that extracted features are consistent and reliable, thus enhancing the robustness of ML models informed by radiomic data (Zwanenburg et al. 2020). This systematic approach helps address traditional challenges in the field, including the diverse methodologies currently in use, by promoting uniform extraction techniques, which can be pivotal for effective clinical translation (Zwanenburg et al. 2020).

Open-source software such as Python is widely accepted, as the PyRadiomics library is standard for feature extraction (Zwanenburg et al. 2020). 3D Slicer is an open-source software that allows for visualization of DICOM files holding CT information for patients and their respective segmentations, utilized by the original publisher of the data set to create segmentations and segmentations were reviewed by the publisher's radiology team within respective software (Simpson et al, 2024). 3D Slicer also allows for visualization of DICOM SEG files that were provided by TCIA, so visual confirmation of liver, FLR, hepatic arteries, and tumors were manually committed. Tumors were isolated and converted from DICOM metadata to three dimensional masks. Lastly, quantitative features for tumor masks were tabulated for cross validation after feature extraction.

Python along with respective libraries such as Pandas, Pyradiomics, SimpleITK and PyDicom was utilized for feature extraction of radiomic features. Python is an open-source software that offers extensive libraries ranging from data science, image preprocessing, and specifically radiomic analysis. File structure for DICOM data was confirmed to hold all 197 patients CT imaging and extracted tumor segmentation masks from 3D Slicer. DICOM data was accessed using PyDicom to ensure proper metadata and matching to respective three-dimensional tumor segmentation masks. Tumor segmentation masks were then evaluated for region of interest labels, accounting for all patient tumors (Zhang, Guo, and Jin 2023). SimpleITK was then utilized to convert DICOM series for each respective patient into three dimensional volumes in preparation for Pyradiomic feature extraction. CT volumes and tumor segmentation masks were cross validated for each patient by image values such as image sizing, spacing, and origin. Gray intensity ranges for all CT volumes were also checked to prepare for voxel array shift.

Parameter file preparation for customizing feature extraction followed ISBI standards such as normalization with scaling to 500 Hounsfield Units, B-spline interpolation to maintain sharpness of three-dimensional volumes and resampling three-dimensional space to one millimeter on all axes (Zwanenburg et al. 2020). Per recommendation of PyRadiomic developers, bin width for gray level intensities was kept default and all negative gray values were voxel array shifted for only positive values to be present. Image outliers more than three standard deviations accounted for within the feature extraction process (Zhang, Guo, and Jin 2023). Metastatic tumors less than 300 cubic millimeters were excluded from feature extraction to ensure robustness of machine learning models. Features extracted included first-order features, shape-based features, and neighboring gray-tone difference matrix. Three gray level matrices featuring were extracted named co-occurrence matrix, run length matrix, and dependence matrix. All features extracted were applied to the original image, wavelet, and Laplacian of Guassian image filters (Zhang, Guo, and Jin 2023). Lastly, PyRadiomics feature extraction was committed to all labels in tumor segmentation masks matched to its respective three-dimensional CT volume and projected onto a data frame using Pandas.

The fusion of radiomics into clinical practice brings tangible benefits to diagnostic and prognostic capabilities. With ongoing initiatives toward standardization like IBSI, alongside the utilization of reliable open-source tools like PyRadiomics, the capacity of radiomics to contribute meaningfully to clinical applications is becoming increasingly viable. Combined efforts to establish clarity and thoroughness in the extraction and validation processes are crucial next steps in ensuring widespread confidence in radiomic methodologies.

The radiomics workflow consists of five primary steps: image acquisition and pre-processing, image segmentation, feature extraction, feature selection, and model creation and evaluation as per a systematic review by Zhang et al (Zhang, Guo, and Jin 2023).

The image acquisition and pre-processing step involves capturing medical images using modalities such as CT, MRI, PET, and ultrasound. The aim is to ensure the quality and consistency of the images, addressing challenges like variability across different imaging protocols. Normalization techniques are applied to standardize image features before analysis (Zhang, Guo, and Jin 2023).

Image segmentation defines regions of interest (ROIs) by employing manual, semi-automatic, or automatic segmentation techniques. By delineating the tumor boundaries accurately, researchers can ensure that the extracted features represent the targeted area reliably (Zhang, Guo, and Jin 2023).

During the feature extraction phase, quantitative characteristics such as shape, texture, and intensity from the tumor ROI are extracted using specialized software tools including PyRadiomics, LIFEx, and MITK (Zhang, Guo, and Jin 2023). This process transforms imaging data into various first-order and higher-order features, enhancing clinical diagnostic and predictive accuracy (Zhang, Guo, and Jin 2023).

Following extraction, a significant number of features are often redundant or irrelevant. Consequently, methods such as LASSO regression and correlation analysis are utilized to refine the feature set, preserving only those features most pertinent to the study objectives (Zhang, Guo, and Jin 2023). Dimensionality reduction techniques may also be employed to enhance model training and reduce overfitting risks.

Selected features are incorporated into various predictive models, typically leveraging machine learning algorithms such as random forests, support vector machines, and deep learning approaches to assess performance (Zhang, Guo, and Jin 2023). Performance metrics such as ROC curves, AUC, precision, recall, and F1 scores facilitate the evaluation of models' predictive accuracies.

Various cancer types have benefited from implementing this radiomic workflow, demonstrating substantial improvements in diagnostic and prognostic capabilities. Research in cancers such as non-small-cell lung cancer, gastric cancer, colon cancer, cervical cancer, and pancreatic cancer has shown that radiomics can effectively predict treatment responses and clarify prognostic indicators (Zhang, Guo, and Jin 2023).

For instance, the work of Huang et al. focused on developing a radiomics nomogram for early-stage non-small-cell lung cancer, achieving exceptional predictive performance for disease-free survival (Zhang, Guo, and Jin 2023). The versatile nature of the radiomic methods enables exploration across multiple tumor types, integrating extensive datasets to enhance translational research insights.

Overall, researchers conclude that the integration of radiomics into clinical practice could significantly augment diagnostic efforts, facilitate personalized therapies, and provide insights into tumor biology (Zhang, Guo, and Jin 2023). However, the review also emphasizes that successful implementation necessitates rigorous methodologies to address issues of validation, data quality, and the need for standardized protocols.

Challenges exist, particularly concerning the reliability of radiomics features in clinical settings. Variables like data quality, small sample sizes from single institutions, and generalizability across diverse patient populations may limit the practical utility of radiomics (Zhang, Guo, and Jin 2023). Standardization across studies, along with external validation through multicenter trials, is vital to overcome these obstacles. Despite these hurdles, advancing rigorous methodologies offers the potential to embed radiomics solidly within cancer diagnostics and management (Zhang, Guo, and Jin 2023).

In conclusion, while radiomics showcases remarkable potential in revolutionizing cancer research, its practical application must continuously evolve through careful consideration of existing challenges and strategic advancements in methodology.

Hybrid ML Models

Recent advancements in machine learning (ML) and radiomics have initiated a shift towards integrating clinical and imaging data for predicting patient outcomes in cancer patients, including those with colorectal liver metastases (CRLM). Machine learning models have shown comparable predictive results to Cox-proportional models for breast cancer recurrence (Gonzalez-Castro et al, 2023) and survival outcomes in cervical cancer (Matsuo et al, 2018). Utilizing radiomics, hybrid machine learning has shown comparable predictive modeling and identification of image biomarkers across primary colorectal cancer, bladder cancer (Xiong et al, 2023), and liver cancer (Banerjee et al, 2015). This section provides an overview of studies demonstrating the efficacy of utilizing hybrid clinical-radiomic models, compares their methodologies, and highlights common concerns raised in recent literature.

Utilizing both clinical and radiomic features enhances prediction accuracy because each modality encompasses complementary information necessary for assessing disease characteristics. For example, a study investigating muscle-invasive bladder cancer (MIBC) used machine learning-based CT radiomics to integrate quantitative imaging data with clinical variables. This clinical-radiomics model achieved an area under the curve (AUC) of 0.958 in the training cohort, showing superior predictive capability compared to authorized models relying solely on clinical data (AUC of 0.779) (Luo et al. 2024) (Xiong et al. 2024). Similarly, another investigation into CRLM patients revealed that when combining retrospective radiomic profiles with clinical risk factors, the predictive scores were effectively stratified, demonstrating the capability of hybrid models to foretell overall survival and treatment responses (Luo et al. 2024).

Subtracting conventional analytical techniques like Cox regression, several studies have adopted more advanced methodologies. Research on bladder cancer utilized machine learning algorithms such as Support Vector Machines (SVM) for feature selection, thereby enhancing prediction accuracy by handling high-dimensional data produced from imaging modalities, as inter-feature relationships are not necessarily linear (Xiong et al. 2024). Similarly, in a study on CRLM, elastic net regularization and random survival forest modeling were employed for radiomic signature construction, aiming to mitigate challenges of overfitting and multicollinearity typically found in high-dimensional features (Luo et al. 2024). These methodologies represent an evolution from isolated traditional statistical approaches to adjacent inclusion of sophisticated ML strategies, indicating a new paradigm in cancer prognostication.

Despite the promising results of hybrid predictive models, researchers have consistently pointed out limitations that could affect external validity and reproducibility. One primary concern is the small sample size and lack of diversity across cohorts, which may influence the generalizability of models beyond specific institutional contexts (Xiong et al. 2024). Additionally, reliance on specific imaging technology raises the question of whether findings can be applied universally across different devices or techniques (Luo et al. 2024). This is compounded by the variability in manual segmentation of imaging data, which can introduce bias and reduce consistency, highlighting the need for standardized automated segmentation methods to improve reproducibility (Xiong et al. 2024).

The integration of radiomic and clinical features holds great promise for improving predictive modeling in oncology. However, future research should focus on multicenter, larger cohort studies to validate observed findings and develop standardized protocols for data acquisition and analysis. These efforts could leverage the benefits of machine learning while addressing concerns about sample representativeness and operational consistency, thereby facilitating the adoption of hybrid models in clinical practice (Xiong et al. 2024).

The rationale for merging clinical and radiomic data in machine learning models to predict the risk recurrence stratification related to CRLM is potent and is backed by the emerging evidence of enhanced predictive power compared to traditional statistical methods.

Recurrence Risk Stratification

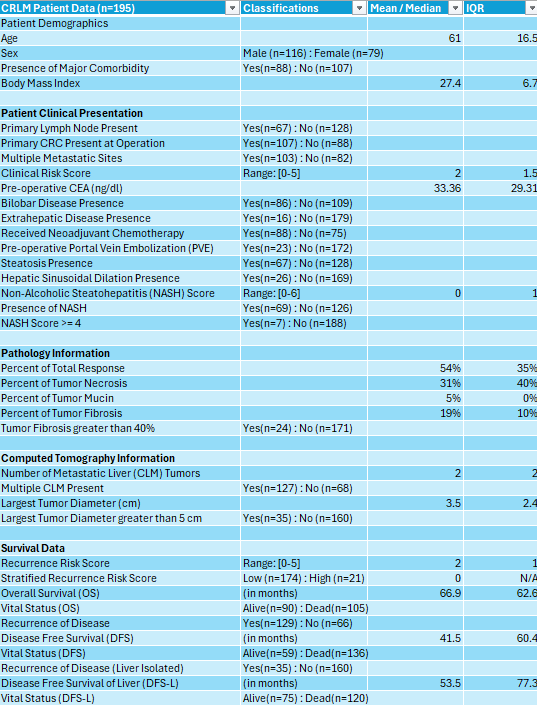
The most utilized survival scoring and stratification for CRLM patients was first developed by Dr. Fong and has been further validated with cross-validating external patient cohorts throughout the early 2000’s (Fong et al, 1999). Significant, independent risk factors have been identified such as threshold for carcinoembryonic antigen levels, disease free survival interval, number of metastatic lesions, nodal status of primary cancer, and diameter of largest tumor. Fong’s scoring system has been validated by nearly six external studies and built upon to include significant recurrence risk stratification between patients ranging from 0-2 and 3-5 (Reissfelder et al, 2009). Recent research by Kawaguchi et al, 2023 also demonstrated that recurrence risk stratification peaks within the first two years post-operatively and levels throughout the survival period. The rationale for utilizing current scoring systems to stratify recurrence risk in predictive modeling is to determine current understanding of CRLM and build upon externally validated studies.

**Solution**

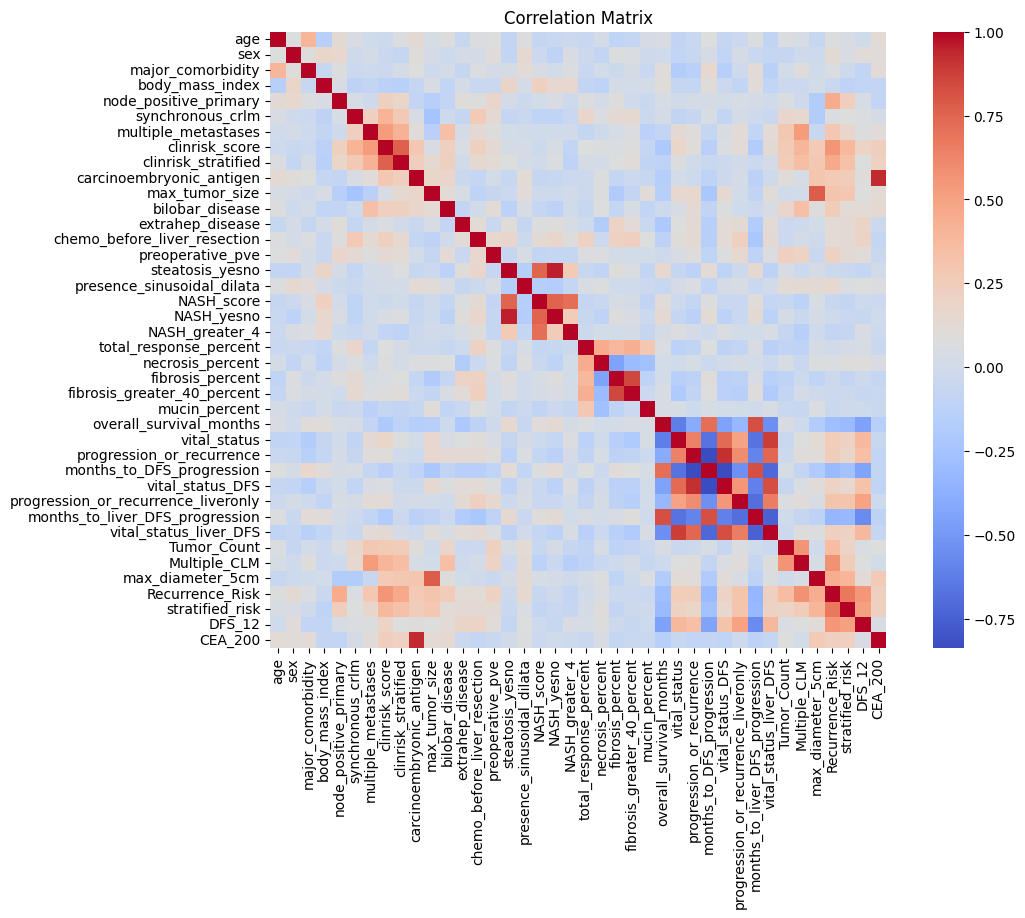
Utilizing libraries for machine learning, specifically SciKit Learn and Pytorch, which are staples in the toolkits of data scientists seeking to create machine learning models, the following models were created and assessed for predictive capabilities. Random Forest, Multiple Logistic Regression, Support Vector Machines, and K-Nearest Neighbors are classical machine learning algorithms that excel in predictive classification, therefore were deployed to predict risk recurrence. SciKit Learn offers feature selection methods ranging from LASSO Regression and Principal Component Analysis (PCA). To ensure computational efficiency of models, the radiomic feature set of 1475 extracted features were reduced to 56 PCA components capturing 95% of informative variance. Metric reporting tools ranging from confusion matrix to classification reporting were conducted for each model and the model score equation was assessed for scaled comparison to other models.

Data Understanding

Twenty patients were randomly selected to cross validate original image shape features with quantitative metrics retrieved from 3D Slicer using NumPy library. Clinical dataset was explored utilizing Pandas libraries locating functions to identify interdependent survival features that would skew predictive modeling aiming to determine recurrence risk rather than overall survival. The data dictionary was verified to hold clinical values indicated from TCIA retrieval and overall characteristics of the patient population shown in figure below. Cox-proportional hazard models were created to ensure comparable cohort trends to new surveillance algorithm patient trends (Kawaguchi et al, 2023) and distribution of literature-indicated independent risk factors for recurrence was analyzed.



Data Preparation

Once tumor radiomic quantitative features were validated, two additional categorical features were manually delineated to supply into the clinical dataset in alignment with the Fong Score and stratified by high and low risk of relative recurrence 24 months post operation (Fong et al, 1999) (Mala et al, 2002) (Reissfelder et al, 2009). Both categories were not included in the original clinical dataset but are features of interest when assessing prediction of stratified recurrence risk, especially regarding the most widely used scoring system utilized by multidisciplinary teams treating CRLM patients. NaN values discovered during data exploration were imputed upon, median replacements for discrete variables and mean replacements for continuous variables, to preserve maximal patient data for machine learning algorithms to utilize. The correlation matrix indicated above was utilized to identify highly collinear clinical data to then remove redundant features prior to model development. Categorical variables were one-hot encoded to allow interpretability within selected models. Statistical analysis was conducted with Wilcoxon-Mann-Whitney where appropriate and Fisher exact test for categorical variables with two-sided p-value < 0.05 considered significant. Lastly, radiomic dataset and clinical dataset were joined by Patient-ID to create a single hybrid clinical-radiomic dataset for model development.

Model Development and Scoring

Utilizing Python’s Sci Kit Learn, each model was coded into with respective library functions and scores were reported for each multi classification predictive model. Model scoring was committed with the indicated figure below. Initial alpha is a negligible constant term, and each x variable indicated is the given average score for a model with respective alpha to x representing coefficient determined from logistic regression model.

Model Performance(s)

Random Forest

Random Forest model was applied, and classification scores were collected. Hyper parameterization of the clinical-radiomic dataset was applied with Grid Search CV to assess the best performing random forest. All metrics such as accuracy, precision, recall, sensitivity, specificity, PPV, NPV were collected for scaled model scoring comparisons. Random Forest machine learning is a powerful classification machine learning model as it mitigates feature redundancy naturally and demonstrates its predictive ability with the highest overall specificity across all models at 1.00. Scaled model score leveled at 0.78, Figure 1.

SVM

The Support Vector model was applied with a linear kernel in this case and classification scores were collected. All metrics such as accuracy, precision, recall, sensitivity, specificity, PPV, NPV were collected for final averaging of model score to compare. Scaled model score demonstrated average accuracy and high negative predictive value, which is significant. Scaled model score leveled at 0.73, Figure 2.

KNN

K Nearest Neighbor model was applied and classification scores were collected. Non parameterization of the K Best feature was needed as KNN determines features based on K value to associate neighboring data points in classification. All metrics such as accuracy, precision, recall, sensitivity, specificity, PPV, NPV were collected for final averaging of model score to compare. KNN with a cross fold utilizing five nearest neighbors demonstrated above average accuracy yet suffered as the other machine learning algorithms in sensitivity. Scaled model score leveled at 0.83, Figure 3.

Multiple Logistic Regression

A multiple logistic regression model was fitted, and classification scores were collected. All features were intra-correlated for collinearity and no redundant features were found post statistical analysis. All metrics such as accuracy, precision, recall, sensitivity, specificity, PPV, NPV were collected for final averaging of model score to compare. Metrics demonstrated below average accuracy along with all other average metrics. Scaled model score was lowest at 0.71, Figure 4.

**Discussion**

Multiple Logistic Regression serves as a baseline model where interpretability is paramount. Although the method can effectively predict binary outcomes, it may not capture complex relationships in data, leading to slightly lower predictive scores relative to more specialized algorithms. KNN excels in non-linear scenarios where the relationship between features is complex, providing good performance in terms of classification but tends to suffer when datasets are too sparse. Its average performance score generally indicates helpful prediction at lower computational costs. SVM often yields higher accuracy rates in cases with clear margin separation in feature space; nevertheless, its performance can decrease significantly when handling unbalanced data sets unless specialized techniques are employed. Random Forest emerges as a highly effective ensemble method that combines the strengths of numerous decision trees, resulting in robustness against overfitting and superior handling of various feature types. Its average scores typically lead the group, demonstrating significant predictive capabilities even in complex datasets.

Evaluating the overall performance of these models reveals that KNN slightly outperformed SVM, Random Forest, and Multiple Logistic Regression for predictive classification of recurrence risk stratification founded upon the Fong Scoring system for clinical radiomics. Thus, integrating various approaches into a hybrid model can potentially mitigate individual weaknesses while enhancing overall prediction accuracy. The combination of multiple modeling algorithms (KNN, Logistic Regression, SVM, Random Forest) allows for a diversified perspective on predictive features, thus harnessing the strengths of each model. Integrating multiple models into a hybrid approach enhances the potential for higher accuracy and improved sensitivity/specificity in predicting recurrence risk, leading to more tailored patient management strategies. The hybrid model can be fine-tuned with additional data, promoting flexibility in model development grounded in updated clinical evidence since training data in this case coasted along the lower boundary of required patient data.

Radiomic feature dimension reduction did not indicate any independent prognostic biomarker for risk recurrence stratification. Image filters such as localized 3D binary patterns, wavelet, Laplace of Gaussian (LoG), and square root were applied. LoG filters demonstrated greater variance contributions to PCA components, but feature levels were unable to be isolated as independent biomarkers with significant relation to the Fong Scoring System.

Prognostic efforts for CRLM continue to evolve as recent surveillance algorithms have found substantial evidence that RAS mutation is a significant risk factor for recurrence, alongside older studies tracking BRAF mutation as well (Kawaguchi et al, 2023). As clinical data set in this instance was created within a single institution, machine learning algorithms call for greater data points for precise training and external validation sets to confirm predictive performance along with recurrence risk stratification.

Another point of discussion is study collection dates for the given CRLM patient data set stemmed from the mid-1990s, when surgical interventions were not nearly as advanced as the current day techniques. New surgical techniques such as the ALPPS procedure, new methodology named portal vein occlusion to enlarge FLR by significant volumetrics, and two-staged hepatectomy that has widened the curative eligibility of surgery for previously considered unresectable patients (Kamba et al, 2021).

**Recommendations**

Although accuracy is key, complex models may render results challenging to interpret, which could hinder their acceptance by clinicians. Additional efforts need to focus on model explainability and providing insights that can influence clinical decision-making. Utilizing various algorithms may introduce complexity and risk of overfitting, especially in smaller datasets if not managed correctly. By predicting recurrence risk accurately, hybrid models present opportunities for more personalized treatment planning, potentially leading to improved outcomes in clinical trials and affecting drug development metrics. Lastly, developing hybrid models opens avenues for collaboration between data scientists, oncologists, and healthcare professionals, encouraging richer insights and practical innovations in CRLM management.

The increasing reliance on patient data raises concerns about data privacy and regulatory compliance. Organizations must ensure strict adherence to ethical practices and regulations. The rapid evolution of machine learning and predictive modeling technologies means that hybrid models could quickly become outdated or less precise than hybrid models that include genomic, proteomics, and even transcriptonomics (Liu et al, 2023). Continuous updates and training for all machine learning algorithms are vital to maintain predictive relevance and justification for precision medicine.

Predictive modeling of recurrence risk in colorectal liver metastases (CRLM) through machine learning (ML) presents several challenges that must be addressed to enhance the robustness and applicability of these models in clinical practice. One significant challenge lies in the heterogeneity of the patient population. CRLM occurs under varying clinical scenarios, with diverse tumor characteristics, treatment histories, and genetic profiles influencing outcomes (Nzenwa et al, 2023). This diversity necessitates the inclusion of an extensive range of features in model training, which can complicate feature selection and may lead to overfitting—where a model performs well on training data but poorly on independent datasets (Xiong et al, 2024).

Additionally, data availability and quality pose a considerable obstacle. High-quality datasets with comprehensive and consistently annotated clinical and imaging data are critical for effective ML model training. However, robust data collection is often hindered by institutional disparities, leading to fragmented and insufficient datasets. This fragmentation impairs the model's ability to generalize across different patient groups and clinical settings, limiting its external validity. Furthermore, incorporating diverse data types, for example, integrating genomic, molecular, and imaging information—adds complexity to the modeling process and requires sophisticated methodologies for multi-modal data integration (Luo et al, 2024).

Interpretability of ML models constitutes another pressing challenge. Many modeling approaches, particularly ensemble methods like Random Forest or black-box algorithms like neural networks, produce complex computations that can obscure clinical relevance and reliability. Clinician acceptance of predictive tools often hinges on the ability to understand and trust their results; hence, enhancing interpretability through model explainability methods becomes essential for fostering confidence and guiding clinical decision-making (Liu et al, 2023). Moreover, regulatory considerations can add yet another layer of complexity. Ensuring compliance with industry standards and ethical guidelines is paramount, particularly when handling sensitive patient data. As predictive models become more integrated into routine clinical workflows, they must adhere to stringent regulatory frameworks to ensure patient safety and uphold trust in technological advances for proper implementation into CDS technology.

**Conclusion**

The performance evaluation of KNN, Multiple Logistic Regression, SVM, and Random Forest indicates that a hybrid approach combining these methodologies holds great promise for predicting recurrence risk in colorectal liver metastases. Support for implementing this solution requires a strategic approach addressing the identified strengths, weaknesses, opportunities, and threats at both macro and micro levels. Continuous improvement, interdisciplinary collaboration, adherence to ethical standards, and focus on model interpretability are critical in advancing hybrid modeling efforts in clinical settings, ultimately contributing to improved patient outcomes and enhanced decision-making in CRLM management.

The continued collection of patient data to train these models will only continue to improve the predictive capability of foundational artificial intelligence models. The inclusion of genomic, proteomic, and radiomic biomarkers holds great promise in hyper specific precision medicine. Deep learning feature crafting has indicated potential in improving radiomic based machine learning models as well, leaving many avenues to continue the pursuit of precision medicine in treating cancer.

While predictive ML modeling for CRLM recurrence presents some challenges for implementation, the promise that such methods hold in transforming clinical practice cannot be understated. As ongoing research addresses these hurdles—through enhanced collaboration among multidisciplinary teams, standardization of practices, and a focus on data quality and model interpretability—the potential impact on patient screening, management, and outcomes will likely grow exponentially. Thus, embarking on this journey, albeit laden with challenges, is critical for advancing precision medicine in oncology.

**References**

Banerjee, Sudeep, David S. Wang, Hyun J. Kim, Claude B. Sirlin, Michael G. Chan, Ronald L. Korn, Aaron M. Rutman, et al. 2015. “A Computed Tomography Radiogenomic Biomarker Predicts Microvascular Invasion and Clinical Outcomes in Hepatocellular Carcinoma.” *Hepatology (Baltimore, Md.)* 62 (3): 792.<https://doi.org/10.1002/hep.27877>.

Bhattacharjya, S., R. Aggarwal, and B. R. Davidson. 2006. “Intensive Follow-up after Liver Resection for Colorectal Liver Metastases: Results of Combined Serial Tumour Marker Estimations and Computed Tomography of the Chest and Abdomen – a Prospective Study.” *British Journal of Cancer* 95 (1): 21–26.<https://doi.org/10.1038/sj.bjc.6603219>.

Buisman, Florian E., Daniele Giardiello, Nancy E. Kemeny, Ewout W. Steyerberg, Diederik J. Höppener, Boris Galjart, Pieter M. H. Nierop, et al. 2022. “Predicting 10-Year Survival after Resection of Colorectal Liver Metastases; an International Study Including Biomarkers and Perioperative Treatment.” *European Journal of Cancer* 168 (June):25–33.<https://doi.org/10.1016/j.ejca.2022.01.012>.

Calderon Novoa, Francisco, Victoria Ardiles, Eduardo de Santibañes, Juan Pekolj, Jeremias Goransky, Oscar Mazza, Rodrigo Sánchez Claria, and Martín de Santibañes. 2023. “Pushing the Limits of Surgical Resection in Colorectal Liver Metastasis: How Far Can We Go?” *Cancers* 15 (7): 2113.<https://doi.org/10.3390/cancers15072113>.

Clark, Margaret E., and Richard R. Smith. 2014. “Liver-Directed Therapies in Metastatic Colorectal Cancer.” *Journal of Gastrointestinal Oncology* 5 (5).<https://doi.org/10.3978/j.issn.2078-6891.2014.064>.

Elmahdy, Mahmoud, and Ronnie Sebro. 2023. “Radiomics Analysis in Medical Imaging Research.” *Journal of Medical Radiation Sciences* 70 (1): 3.<https://doi.org/10.1002/jmrs.662>.

Fajar, Aziz, Riyanarto Sarno, Chastine Fatichah, and Achmad Fahmi. 2022. “Reconstructing and Resizing 3D Images from DICOM Files.” *Journal of King Saud University - Computer and Information Sciences* 34 (6, Part B): 3517–26.<https://doi.org/10.1016/j.jksuci.2020.12.004>.

Fong, Yuman, Joseph Fortner, Ruth L. Sun, Murray F. Brennan, and Leslie H. Blumgart. 1999. “Clinical Score for Predicting Recurrence After Hepatic Resection for Metastatic Colorectal Cancer: Analysis of 1001 Consecutive Cases.” *Annals of Surgery* 230 (3): 309.<https://doi.org/10.1097/00000658-199909000-00004>.

Gholami, Sepideh, Nancy E. Kemeny, Taryn M. Boucher, Mithat Gönen, Andrea Cercek, T. Peter Kingham, Vinod Balachandran, et al. 2020. “Adjuvant Hepatic Artery Infusion Chemotherapy Is Associated with Improved Survival Regardless of KRAS Mutation Status in Patients with Resected Colorectal Liver Metastases: A Retrospective Analysis of 674 Patients.” *Annals of Surgery* 272 (2): 352.<https://doi.org/10.1097/SLA.0000000000003248>.

Gomez, D., V. K. Sangha, G. Morris-Stiff, H. Z. Malik, A. J. Guthrie, G. J. Toogood, J. P. A. Lodge, and K. R. Prasad. 2010. “Outcomes of Intensive Surveillance after Resection of Hepatic Colorectal Metastases.” *The British Journal of Surgery* 97 (10): 1552–60.<https://doi.org/10.1002/bjs.7136>.

González-Castro, Lorena, Marcela Chávez, Patrick Duflot, Valérie Bleret, Alistair G. Martin, Marc Zobel, Jama Nateqi, et al. 2023. “Machine Learning Algorithms to Predict Breast Cancer Recurrence Using Structured and Unstructured Sources from Electronic Health Records.” *Cancers* 15 (10): 2741.<https://doi.org/10.3390/cancers15102741>.

Granata, Vincenza, Roberta Fusco, Federica De Muzio, Carmen Cutolo, Mauro Mattace Raso, Michela Gabelloni, Antonio Avallone, et al. 2022. “Radiomics and Machine Learning Analysis Based on Magnetic Resonance Imaging in the Assessment of Colorectal Liver Metastases Growth Pattern.” *Diagnostics* 12 (5): 1115.<https://doi.org/10.3390/diagnostics12051115>.

Greef, Kathleen De, Christian Rolfo, Antonio Russo, Thiery Chapelle, Giuseppe Bronte, Francesco Passiglia, Andreia Coelho, Konstantinos Papadimitriou, and Marc Peeters. 2016. “Multisciplinary Management of Patients with Liver Metastasis from Colorectal Cancer.” *World Journal of Gastroenterology* 22 (32): 7215.<https://doi.org/10.3748/wjg.v22.i32.7215>.

Griethuysen, Joost JM van, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina GH Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo JWL Aerts. 2017. “Computational Radiomics System to Decode the Radiographic Phenotype.” *Cancer Research* 77 (21): e104.<https://doi.org/10.1158/0008-5472.CAN-17-0339>.

Hu, Dandan, Yangxun Pan, and Gong Chen. 2021. “Colorectal Cancer Liver Metastases: An Update of Treatment Strategy and Future Perspectives.” *Surgery in Practice and Science* 7 (December):100042.<https://doi.org/10.1016/j.sipas.2021.100042>.

Ivey, Gabriel D., Fabian M. Johnston, Nilofer S. Azad, Eric S. Christenson, Kelly J. Lafaro, and Christopher R. Shubert. 2022. “Current Surgical Management Strategies for Colorectal Cancer Liver Metastases.” *Cancers* 14 (4): 1063.<https://doi.org/10.3390/cancers14041063>.

Kambakamba, Patryk, Emir Hoti, Sinead Cremen, Felix Braun, Thomas Becker, and Michael Linecker. 2021. “The Evolution of Surgery for Colorectal Liver Metastases: A Persistent Challenge to Improve Survival.” *Surgery* 170 (6): 1732–40.<https://doi.org/10.1016/j.surg.2021.06.033>.

Kanas, Gena P., Aliki Taylor, John N. Primrose, Wendy J. Langeberg, Michael A. Kelsh, Fionna S. Mowat, Dominik D. Alexander, Michael A. Choti, and Graeme Poston. 2012. “Survival after Liver Resection in Metastatic Colorectal Cancer: Review and Meta-Analysis of Prognostic Factors.” *Clinical Epidemiology* 4 (November):283.<https://doi.org/10.2147/CLEP.S34285>.

Kawaguchi, Yoshikuni, Scott Kopetz, Heather A. Lillemoe, Hyunsoo Hwang, Xuemei Wang, Ching-Wei D. Tzeng, Yun Shin Chun, Thomas A. Aloia, and Jean-Nicolas Vauthey. 2020. “A New Surveillance Algorithm after Resection of Colorectal Liver Metastases Based on Changes Over Time in Recurrence Risk and RAS Mutation Status.” *Journal of the National Comprehensive Cancer Network : JNCCN* 18 (11): 1500.<https://doi.org/10.6004/jnccn.2020.7596>.

Kirimker, Elvan Onur, Alp Togan Kirac, Suleyman Utku Celik, Can Yahya Boztug, Muharrem Berat Kaya, Deniz Balci, and Mehmet Kaan Karayalcin. 2022. “Comparison of Anatomic and Non-Anatomic Liver Resection for Hepatocellular Carcinoma: A Retrospective Cohort Study.” *Medicina* 58 (9): 1305.<https://doi.org/10.3390/medicina58091305>.

Li, Yu, Aydin Eresen, Junjie Shangguan, Jia Yang, Yun Lu, Dong Chen, Jian Wang, Yury Velichko, Vahid Yaghmai, and Zhuoli Zhang. 2019. “Establishment of a New Non-Invasive Imaging Prediction Model for Liver Metastasis in Colon Cancer.” *American Journal of Cancer Research* 9 (11): 2482–92.

Linsalata, Stefania, Rita Borgheresi, Daniela Marfisi, Patrizio Barca, Aldo Sainato, Fabiola Paiar, Emanuele Neri, Antonio Claudio Traino, and Marco Giannelli. 2022. “Radiomics of Patients with Locally Advanced Rectal Cancer: Effect of Preprocessing on Features Estimation from Computed Tomography Imaging.” *BioMed Research International* 2022 (March):2003286.<https://doi.org/10.1155/2022/2003286>.

Liu, Zaoqu, Tian Duan, Yuyuan Zhang, Siyuan Weng, Hui Xu, Yuqing Ren, Zhenyu Zhang, and Xinwei Han. 2023. “Radiogenomics: A Key Component of Precision Cancer Medicine.” *British Journal of Cancer* 129 (5): 741.<https://doi.org/10.1038/s41416-023-02317-8>.

Lu, Cheng, Rakesh Shiradkar, and Zaiyi Liu. 2021. “Integrating Pathomics with Radiomics and Genomics for Cancer Prognosis: A Brief Review.” *Chinese Journal of Cancer Research* 33 (5): 563–73.<https://doi.org/10.21147/j.issn.1000-9604.2021.05.03>.

Luigjes‐Huizer, Yvonne L., Nina M. Tauber, Gerry Humphris, Nadine A. Kasparian, Wendy W. T. Lam, Sophie Lebel, Sébastien Simard, et al. 2022. “What Is the Prevalence of Fear of Cancer Recurrence in Cancer Survivors and Patients? A Systematic Review and Individual Participant Data Meta‐analysis.” *Psycho-Oncology* 31 (6): 879.<https://doi.org/10.1002/pon.5921>.

Luo, Xijun, Hui Deng, Fei Xie, Liyan Wang, Junjie Liang, Xianjun Zhu, Tao Li, et al. 2024. “Prognostication of Colorectal Cancer Liver Metastasis by CE-Based Radiomics and Machine Learning.” *Translational Oncology* 47 (September):101997.<https://doi.org/10.1016/j.tranon.2024.101997>.

Majumder, Shweta, Sharyn Katz, Despina Kontos, and Leonid Roshkovan. 2023. “State of the Art: Radiomics and Radiomics-Related Artificial Intelligence on the Road to Clinical Translation.” *BJR|Open* 6 (1): tzad004.<https://doi.org/10.1093/bjro/tzad004>.

Mala, Tom, Geir Bøhler, Øystein Mathisen, Anstein Bergan, and Odd Søreide. 2002. “Hepatic Resection for Colorectal Metastases: Can Preoperative Scoring Predict Patient Outcome?” *World Journal of Surgery* 26 (11): 1348–53.<https://doi.org/10.1007/s00268-002-6231-x>.

Martin, Jack, Angelica Petrillo, Elizabeth C Smyth, Nadeem Shaida, Samir Khwaja, HK Cheow, Adam Duckworth, et al. 2020. “Colorectal Liver Metastases: Current Management and Future Perspectives.” *World Journal of Clinical Oncology* 11 (10): 761–808.<https://doi.org/10.5306/wjco.v11.i10.761>.

Matsuo, Koji, Sanjay Purushotham, Bo Jiang, Rachel S. Mandelbaum, Tsuyoshi Takiuchi, Yan Liu, and Lynda D. Roman. 2018. “Survival Outcome Prediction in Cervical Cancer: Cox Models vs Deep-Learning Model.” *American Journal of Obstetrics and Gynecology* 220 (4): 381.e1.<https://doi.org/10.1016/j.ajog.2018.12.030>.

Moncada-Torres, Arturo, Marissa C. van Maaren, Mathijs P. Hendriks, Sabine Siesling, and Gijs Geleijnse. 2021. “Explainable Machine Learning Can Outperform Cox Regression Predictions and Provide Insights in Breast Cancer Survival.” *Scientific Reports* 11 (March):6968.<https://doi.org/10.1038/s41598-021-86327-7>.

Nzenwa, I. C., S. Pathak, S. R. Knight, N. G. Mowbray, D. O’Reilly, and R. P. Jones. 2023. “Postoperative Surveillance after Surgery for Colorectal Liver Metastasis: A Cross-Sectional Study.” *Annals of The Royal College of Surgeons of England* 106 (3): 213.<https://doi.org/10.1308/rcsann.2023.0027>.

“‘Role of Radiology in a Multidisciplinary Approach to Patient Care’: Summary of the ESR International Forum 2022.” 2023. *Insights into Imaging* 14 (February):26.<https://doi.org/10.1186/s13244-023-01377-x>.

Schröer, Christoph, Felix Kruse, and Jorge Marx Gómez. 2021. “A Systematic Literature Review on Applying CRISP-DM Process Model.” *Procedia Computer Science*, CENTERIS 2020 - International Conference on ENTERprise Information Systems / ProjMAN 2020 - International Conference on Project MANagement / HCist 2020 - International Conference on Health and Social Care Information Systems and Technologies 2020, CENTERIS/ProjMAN/HCist 2020, 181 (January):526–34.<https://doi.org/10.1016/j.procs.2021.01.199>.

Simpson, Amber L., Alexandre Doussot, John M. Creasy, Lauryn B. Adams, Peter J. Allen, Ronald P. DeMatteo, Mithat Gönen, et al. 2017. “Computed Tomography Image Texture: A Noninvasive Prognostic Marker of Hepatic Recurrence After Hepatectomy for Metastatic Colorectal Cancer.” *Annals of Surgical Oncology* 24 (9): 2482.<https://doi.org/10.1245/s10434-017-5896-1>.

Simpson, Amber L., Jacob Peoples, John M. Creasy, Gabor Fichtinger, Natalie Gangai, Krishna N. Keshavamurthy, Andras Lasso, Jinru Shia, Michael I. D’Angelica, and Richard K. G. Do. 2024. “Preoperative CT and Survival Data for Patients Undergoing Resection of Colorectal Liver Metastases.” *Scientific Data* 11 (1): 172.<https://doi.org/10.1038/s41597-024-02981-2>.

Stark, Gigi F., Gregory R. Hart, Bradley J. Nartowt, and Jun Deng. 2019. “Predicting Breast Cancer Risk Using Personal Health Data and Machine Learning Models.” *PLoS ONE* 14 (12): e0226765.<https://doi.org/10.1371/journal.pone.0226765>.

Xiong, Situ, Zhehong Fu, Zhikang Deng, Sheng Li, Xiangpeng Zhan, Fuchun Zheng, Hailang Yang, et al. 2024. “Machine Learning-Based CT Radiomics Enhances Bladder Cancer Staging Predictions: A Comparative Study of Clinical, Radiomics, and Combined Models.” *Medical Physics* n/a (n/a).<https://doi.org/10.1002/mp.17288>.

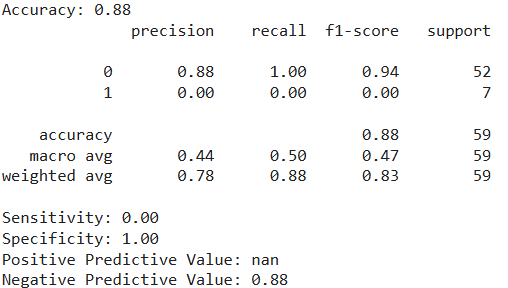
Zhang, Wenchao, Yu Guo, and Qiyu Jin. 2023. “Radiomics and Its Feature Selection: A Review.” *Symmetry* 15 (10): 1834.<https://doi.org/10.3390/sym15101834>.

Zhang, Zhongheng. 2017. “Survival Analysis in the Presence of Competing Risks.” *Annals of Translational Medicine* 5 (3): 47.<https://doi.org/10.21037/atm.2016.08.62>.

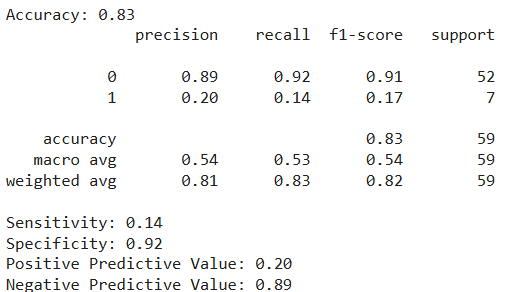
Zhou, Hui, Zhongtao Liu, Yongxiang Wang, Xiaoyong Wen, Eric H. Amador, Liqin Yuan, Xin Ran, et al. 2022. “Colorectal Liver Metastasis: Molecular Mechanism and Interventional Therapy.” *Signal Transduction and Targeted Therapy* 7 (1): 1–25.<https://doi.org/10.1038/s41392-022-00922-2>.

**Appendix**

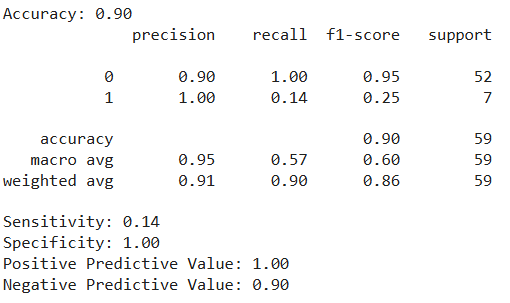
**Figure 1: Random Forest Classifier Classification Report**

****

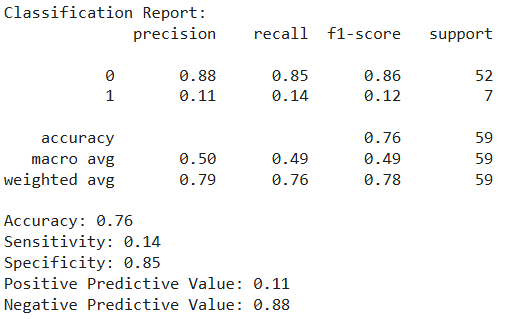
**Figure 2: Support Vector Machine Classification Report**

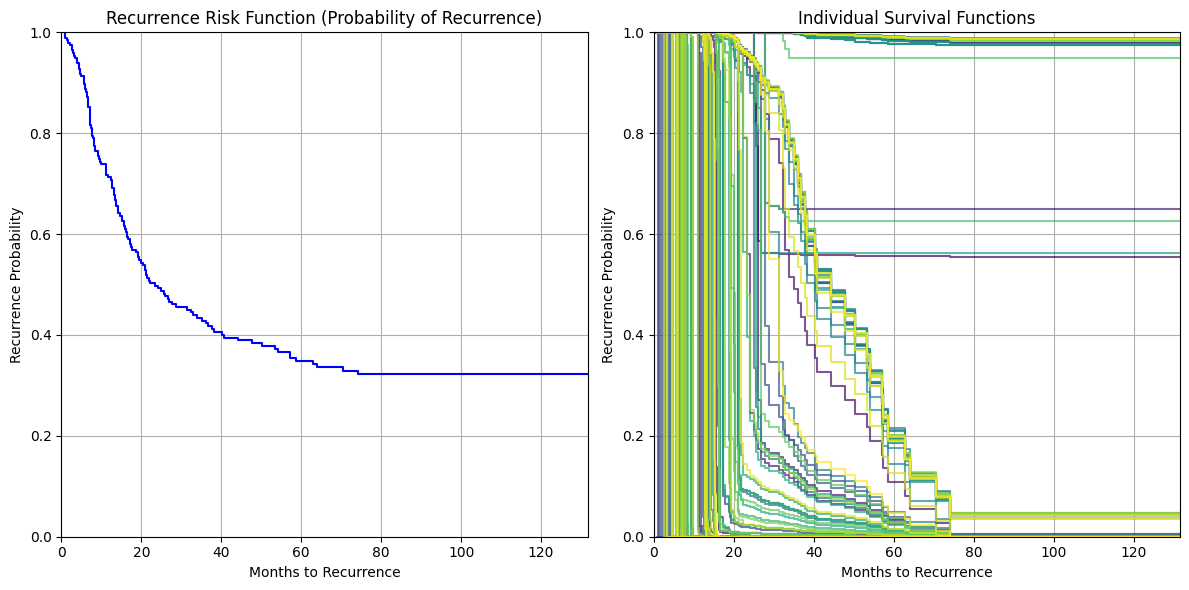
****

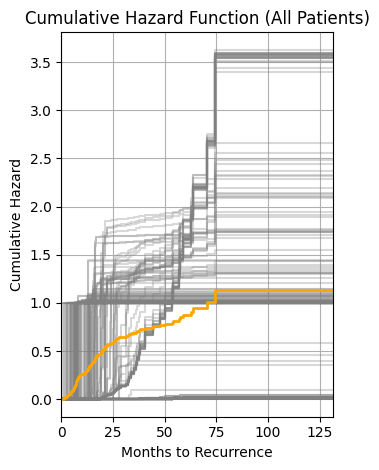
**Figure 3: K-Nearest Neighbors Classification Report**

****

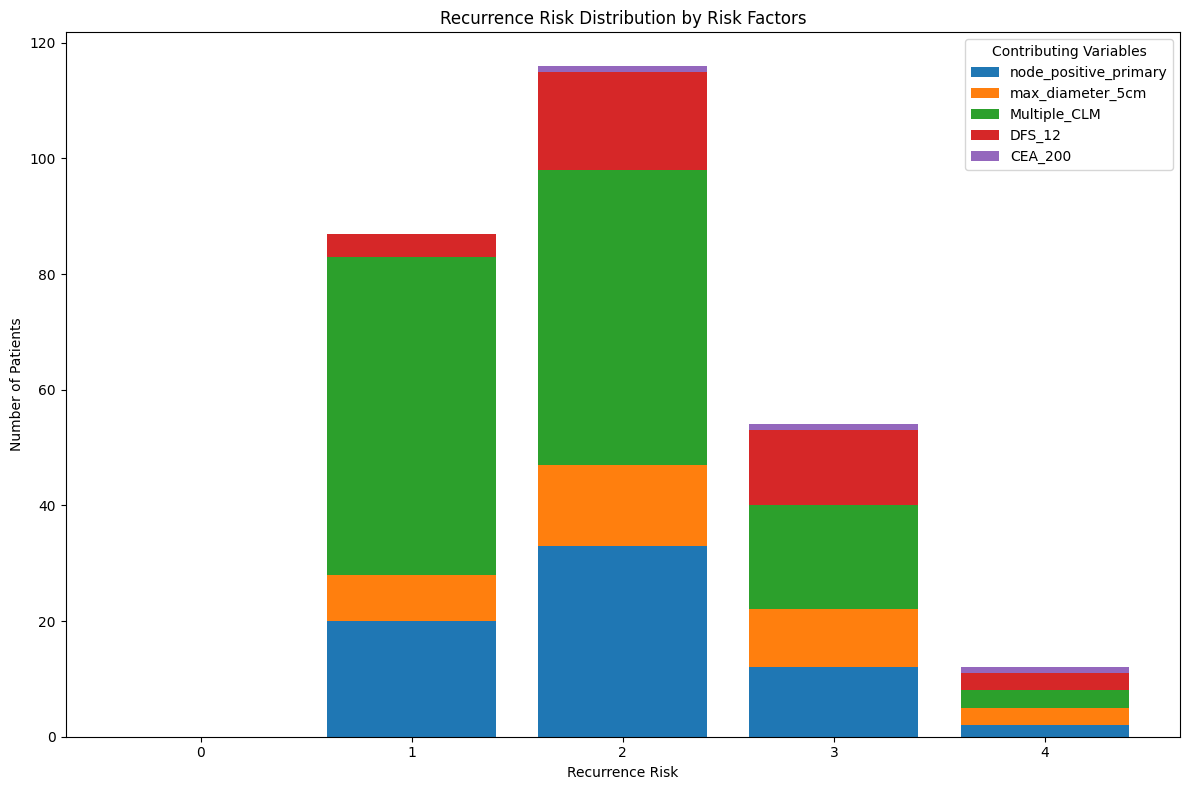
**Figure 4: Logistic Regression Classification Report**

****

**Figure 5: Comparable Survival Curves Using Cox Proportional Hazards**



**Figure 6: Fong Scoring Histogram Distribution**

****