

CT-Based Radiomic Models for Stage Stratification and Relapse Prediction in Colon Cancer

Manuel Collado

Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas
- Universidad Autónoma de Madrid

María E. Castillo

Alcalá University

M^a Jesús Larriba

Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas
- Universidad Autónoma de Madrid

Cristina Galindo-Pumariño

Ramón y Cajal University Hospital, IRYCIS, Alcalá University

Carmen Lisset Flores

Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas
- Universidad Autónoma de Madrid

Reyes Ferreira

Ramón y Cajal University Hospital, IRYCIS, Alcalá University

Elena Canales Lachén

Ramón y Cajal University Hospital, IRYCIS, Alcalá University

Raquel García

Ramón y Cajal University Hospital, IRYCIS, Alcalá University

José Avendaño-Ortiz

Ramón y Cajal University Hospital and IRYCIS

José Manuel González-Sancho

Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas
- Universidad Autónoma de Madrid

Carolina Pinta

Ramón y Cajal University Hospital, IRYCIS, Alcalá University

Cristina Peña

cristina.pena@csic.es

Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas
- Universidad Autónoma de Madrid

Research Article

Keywords: Colon Cancer, Radiomics, Prognostic value

Posted Date: September 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10802890/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose:

To develop and validate computed tomography (CT)-based radiomic models for differentiating stage II from stage III colon cancer and predicting 5-year relapse risk.

Materials and Methods:

In this retrospective single-center study, 104 patients with colon cancer who underwent preoperative CT between December 2016 and September 2022 were included. Primary tumors were manually segmented in three dimensions, and 105 radiomic features were extracted together with preoperative clinical variables. The dataset was randomly divided into a training cohort (70%) and an independent validation cohort (30%). Model development and optimization were performed using 10-fold cross-validation within the training cohort. Separate machine learning models were developed for stage differentiation ($n = 84$) and 5-year relapse prediction ($n = 104$). Model performance was evaluated in the validation cohort using area under the receiver operating characteristic curve (AUC).

Results:

For stage differentiation, the generalized linear model achieved an AUC of 0.760 in the validation cohort, with a balanced accuracy of 0.744 and a sensitivity of 0.889 for identifying stage II disease. For 5-year relapse prediction, the partial least squares model showed the highest performance, with an AUC of 0.910 and a sensitivity of 0.950. Subgroup analyses demonstrated consistent predictive performance in stage II (AUC, 0.929) and stage III disease (AUC, 0.926).

Conclusion:

CT-derived radiomic features enabled differentiation between stage II and stage III colon cancer and showed prognostic value for predicting 5-year relapse. These findings support the potential role of radiomics as a noninvasive imaging biomarker for preoperative risk stratification in colon cancer.

HIGHLIGHTS

- CT radiomics differentiated stage II from stage III colon cancer.
- CT radiomics predicted 5-year relapse using preoperative imaging biomarkers.
- Radiomic biomarkers may complement conventional staging for individualized preoperative risk stratification in colon cancer.

INTRODUCTION

Colon cancer remains one of the most frequently diagnosed malignancies and a leading cause of cancer-related mortality worldwide (1). Prognosis is largely determined by tumor stage, yet outcomes

vary substantially among patients within the same stage. In stage III disease, adjuvant chemotherapy improves survival, although approximately 25–30% of patients still relapse within five years (2–4). In contrast, most stage II patients have a favorable prognosis, but a subset experiences recurrence and may benefit from systemic treatment (2). Conversely, a proportion of stage III patients have a relatively favorable prognosis and may achieve similar outcomes with shorter durations of adjuvant chemotherapy, thereby reducing treatment-related toxicity (5). This clinical heterogeneity highlights the limitations of current risk stratification strategies and the need for more precise prognostic tools.

The tumor–node–metastasis staging system remains the standard for clinical decision-making; however, it does not fully capture the biological diversity of colon cancer (6, 7). Several molecular and histopathological biomarkers, including microsatellite instability (8) and BRAF mutation (9), have been proposed to refine prognosis, yet their clinical utility remains limited in routine practice (10). As a result, identifying noninvasive biomarkers capable of improving risk assessment before surgery remains an unmet need.

Computed tomography plays a central role in the preoperative evaluation of colon cancer, primarily to detect metastatic disease (11). Its potential to provide prognostic information beyond visual assessment is still underexplored. Radiomics allows extraction of quantitative features from medical images, capturing tumor characteristics related to shape, intensity, and texture that may reflect underlying tumor biology (12–14). Previous studies have shown that radiomic features can be associated with survival, treatment response, and molecular alterations across different tumor types (15–17).

In colon cancer, however, evidence remains limited and heterogeneous. Many studies have evaluated colorectal cancer as a single entity or focused on isolated endpoints such as survival or molecular status (18–22), which restricts their applicability to stage-specific clinical decision-making. In addition, variability in imaging protocols and analytical pipelines has further hindered the translation of radiomic findings into clinical practice, particularly in distinguishing risk within stage II and stage III disease.

The aim of this study was to develop and validate radiomic-based models derived from preoperative computed tomography to (a) differentiate stage II from stage III colon cancer and (b) predict 5-year relapse. By focusing on preoperative data, this approach seeks to provide clinically actionable information at an early stage, with potential implications for treatment selection and risk-adapted management.

MATERIALS AND METHODS

Study Design and Patients

The study population was restricted to patients with sporadic colon cancer. Individuals with familial adenomatous polyposis or fulfilling the Amsterdam criteria for hereditary non-polyposis colorectal cancer were not included. A total of 121 patients with colon cancer who underwent colonoscopy,

preoperative computed tomography, surgery, and adjuvant chemotherapy when indicated were initially identified. Inclusion criteria were age 18 years or older, recent diagnosis of colon cancer, and availability of preoperative computed tomography imaging. Exclusion criteria included inability to provide informed consent, prior colon surgery, pregnancy, or any condition preventing participation in the study or adequate follow-up. After applying these criteria, 104 patients were included in the final analysis.

Patients were followed after diagnosis through scheduled clinical visits, laboratory testing, and imaging studies according to institutional protocols. Additional imaging, including abdominal ultrasound, was performed when clinically indicated. Patients were followed for up to 5 years after diagnosis; those with shorter available follow-up were censored at their last recorded assessment. Overall survival was calculated from the date of diagnosis until death, irrespective of cause, whereas disease-free survival was calculated from diagnosis until the first confirmed recurrence. Patients who did not experience the corresponding event were censored at the date of their last available follow-up. Clinical variables collected included age at diagnosis, sex, tumor location, histopathological characteristics, and tumor stage according to the tumor–node–metastasis classification.

Patients were randomly assigned to a training cohort (70%) and an independent validation cohort (30%), ensuring a comparable distribution of clinical and imaging variables between both sets.

CT Acquisition and Tumor Segmentation

Preoperative contrast-enhanced computed tomography scans as part of routine clinical care between December 2016 and September 2022 were analyzed. All images were processed using the QP-Precision platform (Quibim, Valencia, Spain).

Primary tumors were manually segmented in three dimensions on a slice-by-slice basis by trained imaging technicians in collaboration with experienced radiologists, following institutional segmentation procedures. The entire tumor volume was included for analysis. Figure 1A illustrates the segmentation of the primary lesion in one representative case (patient 82).

Radiomic Feature Extraction

A total of 105 quantitative radiomic features were extracted from each segmented tumor volume using the QP-Precision platform (Quibim, Valencia, Spain) (23). The extracted features were classified into three main categories:

- Shape features, which characterize tumor geometry, including measures such as volume, sphericity, and elongation.
- First-order features, which describe the distribution of voxel intensities, including entropy, skewness, and kurtosis.

- Second-order texture features, which characterize spatial patterns and relationships among voxel intensities. These features were calculated from gray-level co-occurrence, run-length, size-zone, dependence, and neighboring gray-tone difference matrices.

Radiomic features were reviewed to identify outliers and near-zero variance variables. Features with variance < 0.01 were excluded prior to model development.

Study Endpoints and Feature Set

Two primary endpoints were defined:

1. Differentiation between stage II and stage III colon cancer
2. Prediction of 5-year relapse

For stage differentiation, patients with stage I, stage IV, or unknown stage were excluded (final n = 84). For relapse prediction, the full cohort was used (n = 104).

Radiomic features were combined with preoperative clinical variables, including age, sex, and tumor laterality. Pathological variables were excluded from the stage differentiation analysis to ensure a strictly preoperative model.

Data Preprocessing

A k-nearest neighbors (KNN) approach was applied to impute missing values. The proportion of missing values was low and mainly affected tumor laterality and selected pathological variables.

Dimensionality reduction included removal of near-zero variance features and univariate filtering using the Mann–Whitney U test. Variables with effect size > 0.10 and $p < 0.05$ were retained for model development. This process is detailed in Supplementary Material I.

Model Development and Validation

Patients were randomly allocated to model development (70%) and hold-out validation (30%) sets. Model selection and hyperparameter optimization were carried out within the development set, with performance during model training estimated by 10-fold cross-validation.

Several machine learning models were evaluated for each endpoint (Fig. 1B). Their final predictive performance was determined on the 30% hold-out set based on the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity.

For relapse prediction, partial least squares and support vector machine models were further analyzed due to their performance and complementary interpretability.

RESULTS

Patient Characteristics

The study cohort included 104 patients with colon cancer (mean age, 73 years; range, 34–91 years), comprising 49 men and 55 women. Tumors were located in the right colon in 46 patients (44%) and in the left colon in 40 (39%), while 17 patients (16%) had unknown tumor location.

According to tumor stage, 16 patients (15%) were classified as stage I, 40 (39%) as stage II, 44 (42%) as stage III, and 2 (2%) as stage IV; staging was unknown in 2 patients (2%). During follow-up, 19 patients experienced disease relapse (Fig. 1C).

Stage II versus Stage III Differentiation

After excluding patients with stage I, stage IV, or unknown stage, 84 patients were included in this analysis.

Following feature selection, four radiomic variables were retained based on statistical significance and effect size criteria (Fig. 2A). Detailed results, including p-values and effect sizes for all variables, are provided in Supplementary Material II. Most evaluated classification models achieved an area under the curve greater than 0.70, indicating consistent discriminative performance across approaches (Fig. 2B).

Model comparison using accuracy and kappa metrics showed similar performance among classifiers, with the C5.0 model achieving slightly higher accuracy. However, the generalized linear model was selected due to its balance between performance and interpretability (Fig. 2C and 2D).

In the validation cohort, the generalized linear model achieved an area under the curve of 0.760 (Fig. 3A), with a balanced accuracy of 0.744 and a sensitivity of 0.889 for identifying stage II disease (Fig. 3B). The relative importance of the included variables is shown in Fig. 3C.

Prediction of 5-Year Relapse

In the full cohort (n = 104), 28 radiomic features were retained after feature selection (Fig. 4A). Full statistical details for all variables are presented in Supplementary Material III.

Among the evaluated models, partial least squares and support vector machine classifiers showed the best predictive performance (Fig. 4B). The partial least squares model achieved the highest accuracy, with an area under the curve of 0.910 and a sensitivity of 0.950 in the validation cohort (Fig. 5).

Although the partial least squares model showed superior performance, its interpretability was limited due to the use of latent components. In contrast, the support vector machine model showed lower overall performance (area under the curve, 0.730) but enabled identification of the most relevant predictive features, which were predominantly radiomic texture variables (Fig. 6).

Subgroup Analyses by Tumor Stage

Separate analyses were performed for stage II (n = 40) and stage III (n = 44) patients.

For the stage II subgroup, 21 radiomic features were initially retained after univariate filtering (effect size > 0.10 , $p < 0.05$) (Supplementary Fig. 1A; Supplementary Material IV). Given the limited sample size, multicollinearity was assessed using variance inflation factor analysis (Supplementary Fig. 1B). Stepwise exclusion of highly collinear variables resulted in a final model including four radiomic features with acceptable collinearity metrics (Supplementary Fig. 1C). A binomial logistic regression model was then constructed without evidence of multicollinearity (Supplementary Fig. 1D). The model achieved an area under the curve of 0.929 and a specificity of 0.964 for identifying patients without relapse at 5 years (Supplementary Fig. 1E).

For the stage III subgroup, 22 radiomic features were retained after univariate filtering (Supplementary Fig. 2A; Supplementary Material V). After collinearity assessment, seven radiomic features were included in the final model (Supplementary Fig. 2B). The resulting logistic regression model achieved an area under the curve of 0.926, with specificity of 1.000 and sensitivity of 0.714 (Supplementary Fig. 2C).

These subgroup analyses were exploratory due to the limited sample size but demonstrated consistent predictive performance across stages.

DISCUSSION

In this study, we developed radiomic-based models derived from preoperative computed tomography that enabled both differentiation between stage II and stage III colon cancer and prediction of 5-year relapse risk. The generalized linear model achieved moderate accuracy for stage discrimination, whereas the partial least squares model showed high performance for relapse prediction, with consistent results across stage-specific subgroup analyses. Together, these findings suggest that a single imaging-based framework may provide quantitative imaging information related to both tumor stage and prognosis using routinely acquired preoperative imaging.

Beyond model performance, the main value of this work lies in its potential clinical applicability. Current risk stratification in colon cancer is largely driven by postoperative pathological features, which limits its usefulness for guiding decisions before surgery (2, 5). In contrast, the present approach relies exclusively on preoperative data, raising the possibility of incorporating imaging-derived biomarkers earlier in the clinical workflow. In practice, such models could complement the initial staging process by providing additional quantitative imaging information during multidisciplinary evaluation. This may contribute to improved identification of patients with different relapse risk profiles before treatment initiation.

Although the tumor–node–metastasis classification remains the cornerstone of clinical decision-making (6), it does not fully account for the biological heterogeneity observed within each stage (7). The

variability in outcomes among patients with similar staging highlights the need for complementary biomarkers that can refine risk stratification. In this context, radiomics offers a noninvasive approach that may reflect underlying tumor heterogeneity and biological characteristics not captured by conventional staging systems (12, 13).

Radiomic features, particularly those describing tumor texture, were consistently associated with relapse risk in our study. Tumors with more homogeneous imaging patterns were more frequently observed in patients without recurrence, whereas increased heterogeneity appeared to be linked to more aggressive behavior. These findings are consistent with prior studies showing that radiomic features can identify high-risk patients and predict clinical outcomes in colon cancer (18, 20, 24), as well as with work demonstrating their utility in tumor characterization, staging, and the prediction of molecular and histopathological features such as microsatellite instability and perineural invasion (19, 25–27). However, the biological basis of these imaging patterns remains incompletely understood.

Prior radiomics studies in colorectal cancer have explored a range of clinical applications, including stage assessment and outcome prediction, often within heterogeneous cohorts that combine colon and rectal tumors (28–32). In parallel, studies focused specifically on colon cancer have typically addressed individual endpoints, such as stage prediction or prognosis, in isolation (19, 20, 24). This fragmentation may limit the ability of existing approaches to provide a comprehensive framework for risk stratification. In contrast, the present study focuses exclusively on colon cancer and integrates both stage differentiation and relapse prediction within a single analytical approach based on preoperative imaging.

The subgroup analyses further suggest that radiomic models may retain predictive value within individual tumor stages. High area under the curve values were observed in both stage II and stage III cohorts. However, these findings should be interpreted as exploratory due to the limited sample size and the potential risk of overfitting. External validation in larger, multicenter datasets will be necessary to confirm their robustness and generalizability.

This study has several limitations. First, its retrospective and single-center design may introduce selection bias and limit generalizability. Second, the relatively small sample size, particularly in subgroup analyses, constrains the stability of the models. Third, although feature selection and cross-validation strategies were applied, no independent external validation cohort was available. In addition, manual tumor segmentation may introduce variability, and the impact of image acquisition variability and feature extraction reproducibility across institutions and imaging protocols was not formally assessed.

In summary, radiomic features extracted from routine preoperative computed tomography may provide clinically relevant information for both stage differentiation and relapse prediction in colon cancer. If validated in larger prospective studies, these models could complement existing clinical and pathological factors, supporting quantitative imaging-based risk assessment in colon cancer.

Declarations

Clinical Trial Number

Not applicable

Ethics approval and consent to participate

The study was approved by the Comité de Ética de la Investigación con Medicamentos of Hospital Universitario Ramón y Cajal (Approval ID: PI20/00602; Act No. 414) and conducted in accordance with the Declaration of Helsinki. The initial database included 121 CC patients recruited at the same institution, all of whom provided written informed consent to participate in the study. All patients underwent colonoscopy, preoperative CT, surgery, and adjuvant chemotherapy when indicated. Preoperative CT images and postoperative pathological reports were collected for all cases.

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analyzed during the current study are not publicly available due to institutional and data protection restrictions but may be available from the corresponding author on reasonable request and subject to approval by the data owners and the institutional ethics committee.

Competing interests

The authors declare that they have no competing interests.

Funding

This research is supported by PI20/00602 from the Instituto de Salud Carlos III and co-financed by the European Development Regional Fund (FEDER) “A way to achieve Europe” (ERDF); “CIBER de Cáncer” (CB16/12/00273) and “CIBER de Enfermedades Infecciosas” (CB21/13/00084) and and “CIBER Enfermedades Raras” (CB06/07/1029) from the Instituto de Salud Carlos III - FEDER “A way to achieve Europe” (ERDF); CNS2023-144882 from Agencia Estatal de Investigación “Plan de Recuperación, Transformación y Resiliencia; Next Generation EU”; and P2022/BMD7212, Comunidad de Madrid. MJL and JMG-S belong to the Spanish National Research Council (CSIC)'s Cancer Hub. MC was supported by FI21/00132 from Instituto de Salud Carlos III - FEDER “A way to achieve Europe” (ERDF).

Authors' contributions

Following the CRediT taxonomy, the authors' contributions were as follows: conceptualization, MC, MJL, CGP, CDLP, and CP; methodology, MC, CGP, CDLP, and CP; formal analysis, MC, CDLP, and CP; investigation, MC, MEC, CGP, CLF, CDLP, and CP; resources, MEC, RF, EC, RG, JAO, CDLP, and CP; writing—original draft preparation, MC, MEC, MJL, CLF, and CP; writing—review and editing, MC, MJL, CLF, JAO,

JMGS, CDLP, and CP; visualization, MC and CP; supervision, CDLP and CP; project administration, CDLP and CP; funding acquisition, CDLP and CP. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to lab members for help and advice throughout this research. ChatGPT was used for English text editing.

References

1. Filho AM, Laversanne M, Ferlay J, Colombet M, Piñeros M, Znaor A, et al. The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide. *Int J Cancer*. 2025;156(7):1336–46. 10.1002/IJC.35278 PubMed PMID: 39688499.
2. Böckelman C, Engelmann BE, Kaprio T, Hansen TF, Glimelius B. Risk of recurrence in patients with colon cancer stage II and III: a systematic review and meta-analysis of recent literature. *Acta Oncol*. 2015;54(1):5–16. 2014.975839 PubMed PMID: 25430983.
3. Murray NP, Aedo S, Villalon R, López MA, Minzer S, Muñoz L, et al. Effect of FOLFOX on minimal residual disease in Stage III colon cancer and risk of relapse. *Ecancermedicallscience*. 2019;13. 10.3332/. ECANCER.2019.935 PubMed PMID: 31281432.
4. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet*. 2014;383(9927):1490–502. 10.1016/S0140-6736(13)61649-9 PubMed PMID: 24225001.
5. Grothey A, Sobrero AF, Shields AF, Yoshino T, Paul J, Taieb J, et al. Duration of adjuvant chemotherapy for stage III colon cancer. *N Engl J Med*. 2018;378(13):1177–88. doi:10.1056/NEJMoa1713709 PubMed PMID: 29590544.
6. Amin MB, Greene FL, Edge SB, Compton CC, Gershengwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more personalized approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93–9. 21388 PubMed PMID: 28094848.
7. Shackleton M, Quintana E, Fearon ER, Morrison SJ. Heterogeneity in Cancer: Cancer Stem Cells versus Clonal Evolution. *Cell*. 2009;138(5):822–9. 10.1016/j.cell.2009.08.017 PubMed PMID: 19737509.
8. Diaz LA, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, et al. Pembrolizumab versus chemotherapy for microsatellite instability-high or mismatch repair-deficient metastatic colorectal cancer (KEYNOTE-177): final analysis of a randomised, open-label, phase 3 study. *Lancet Oncol*. 2022;23(5):659–70. 2045(22)00197-8 PubMed PMID: 35427471.
9. Tabernero J, Grothey A, van Cutsem E, Yaeger R, Wasan H, Yoshino T, et al. Encorafenib Plus Cetuximab as a New Standard of Care for Previously Treated BRAF V600E-Mutant Metastatic Colorectal Cancer: Updated Survival Results and Subgroup Analyses from the BEACON Study. *J Clin Oncol*. 2021;39(4):273–84. 10.1200/JCO.20.02088 . PubMed PMID: 33503393.

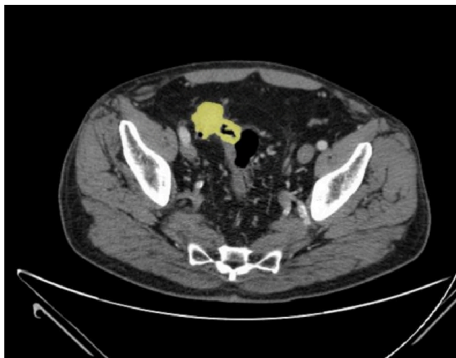
10. Fabregas JC, Ramnaraigh B, George TJ. Clinical Updates for Colon Cancer Care in 2022. Clin Colorectal Cancer. 2022;21(3):198–203. 10.1016/j.clcc.2022.05. .006 PubMed PMID: 35729033.
11. Argilés G, Tabernero J, Labianca R, Hochhauser D, Salazar R, Iveson T, et al. Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. Ann Oncol. 2020;31(10):1291–305. 10.1016/j.annonc.2020.06.022 . PubMed PMID: 32702383.
12. Gillies RJ, Kinahan PE, Hricak H, Radiomics. Images are more than pictures, they are data. Radiology. 2016;278(2):563–77. 10.1148/radiol.2015151169 . PubMed PMID: 26579733.
13. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol. 2016;61(13):R150–66. 10.1088/0031-9155/61/13/. R150 PubMed PMID: 27269645.
14. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. J Nucl Med. 2020;61(4):488–95. 10.2967/JNUMED.118.222893 . PubMed PMID: 32060219.
15. Deng F, Xiao G, Tanzhu G, Chu X, Ning J, Lu R, et al. Predicting Survival Rates in Brain Metastases Patients from Non-Small Cell Lung Cancer Using Radiomic Signatures Associated with Tumor Immune Heterogeneity. Adv Sci (Weinh). 2025;12(10). 10. 1002/ADVS.202412590 PubMed PMID: 39840456.
16. Liao CY, Chen YM, Wu Y, Te, Chao HS, Chiu HY, Wang TW, et al. Personalized prediction of immunotherapy response in lung cancer patients using advanced radiomics and deep learning. Cancer Imaging. 2024;24(1). 10.1186/S40644-024-00779-4 . PubMed PMID: 39350284.
17. He B, Dong D, She Y, Zhou C, Fang M, Zhu Y, et al. Predicting response to immunotherapy in advanced non-small-cell lung cancer using tumor mutational burden radiomic biomarker. J Immunother Cancer. 2020;8(2). 10. 1136/JITC-2020-000550 PubMed PMID: 32636239.
18. Caruso D, Polici M, Zerunian M, Del Gaudio A, Parri E, Giallorenzi MA, et al. Radiomic Cancer Hallmarks to Identify High-Risk Patients in Non-Metastatic Colon Cancer. Cancers (Basel). 2022;14(14). 10. 3390/CANCERS14143438 PubMed PMID: 35884499.
19. Li Y, Eresen A, Lu Y, Yang J, Shangguan J, Velichko Y, et al. Radiomics signature for the preoperative assessment of stage in advanced colon cancer. Am J Cancer Res. 2019;9(7):1429–38. PubMed PMID: 31392079.
20. Dai W, Mo S, Han L, Xiang W, Li M, Wang R, et al. Prognostic and predictive value of radiomics signatures in stage I-III colon cancer. Clin Transl Med. 2020;10(1):288–93. 2.31 PubMed PMID: 32508036.
21. Negreros-Osuna AA, Parakh A, Corcoran RB, Pourvaziri A, Kambadakone A, Ryan DP, et al. Radiomics Texture Features in Advanced Colorectal Cancer: Correlation with BRAF Mutation and 5-year Overall Survival. Radiol Imaging Cancer. 2020;2(5). 10. 1148/Rycan.2020190084 PubMed PMID: 33778733.
22. MG L, AM del R NSSJL, C S, RB H, et al. CT textural analysis of hepatic metastatic colorectal cancer: pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes. Abdom Imaging. 2015;40(7):2331–7. 10.1007/S00261-015-0438-4 . PubMed PMID: 25968046.

23. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics. *Radiol Med.* 2021;126(10):1296–311. 10.1007/S11547-021-01389-X PubMed PMID: 34213702.
24. JM C, KM C, JC JC. Differences in Liver Parenchyma are Measurable with CT Radiomics at Initial Colon Resection in Patients that Develop Hepatic Metastases from Stage II/III Colon Cancer. *Ann Surg Oncol.* 2021;28(4):1982–9. 10.1245/S10434-020-09134-W PubMed PMID: 32954446.
25. Zerunian M, Nacci I, Caruso D, Polici M, Masci B, De Santis D, et al. Is CT Radiomics Superior to Morphological Evaluation for pN0 Characterization? A Pilot Study in Colon Cancer. *Cancers (Basel).* 2024;16(3). 10.3390/CANCERS16030660 PubMed PMID: 38339411.
26. Golia Pernicka JS, Gagniere J, Chakraborty J, Yamashita R, Nardo L, Creasy JM, et al. Radiomics-based prediction of microsatellite instability in colorectal cancer at initial computed tomography evaluation. *Abdom Radiol (NY).* 2019;44(11):3755–63. 02117-W PubMed PMID: 31250180.
27. Guo T, Cheng B, Li Y, Li Y, Chen S, Lian G, et al. A radiomics model for predicting perineural invasion in stage II-III colon cancer based on computer tomography. *BMC Cancer.* 2024;24(1). 10.1186/S12885-024-12951-X . PubMed PMID: 39367321.
28. Lv L, Xin B, Hao Y, Yang Z, Xu J, Wang L, et al. Radiomic analysis for predicting prognosis of colorectal cancer from preoperative 18F-FDG PET/CT. *J Transl Med.* 2022;20(1). 10.1186/S12967-022-03262-5 . PubMed PMID: 35109864.
29. Badic B, Hatt M, Durand S, Jossic-Corcos C, Le, Simon B, Visvikis D, et al. Radiogenomics-based cancer prognosis in colorectal cancer. *Sci Rep.* 2019;9(1):9743. 10.1038/s41598-019-46286-6 . PubMed PMID: 31278324.
30. Li Y, Gong J, Shen X, Li M, Zhang H, Feng F, et al. Assessment of Primary Colorectal Cancer CT Radiomics to Predict Metachronous Liver Metastasis. *Front Oncol.* 2022;12. 10.3389/FONC.2022.861892 PubMed PMID: 35296011.
31. M T, F GM, F I, DB FS. CT-Based Radiomics Analysis Before Thermal Ablation to Predict Local Tumor Progression for Colorectal Liver Metastases. *Cardiovasc Intervent Radiol.* 2021;44(6):913–20. 10.1007/S00270-020-02735-8 . PubMed PMID: 33506278.
32. Rocca A, Brunese MC, Santone A, Avella P, Bianco P, Scacchi A, et al. Early Diagnosis of Liver Metastases from Colorectal Cancer through CT Radiomics and Formal Methods: A Pilot Study. *J Clin Med.* 2021;11(1). 10.3390/JCM11010031 PubMed PMID: 35011771.

Figures

FIGURE 1

A



B

Classification models tested
Logistic regression
Gradient Boosting with Component-wise Linear Models
GLM net
Random Forest
Support Vector Classification
Classification And Regression Tree
Conditional Interference Tree
C5.0
Partial least squares discrimination analysis

C

Variable	N=104
Age (years), mean (range)	73 (34-91)
Sex	
• Male	49 (47%)
• Female	55 (53%)
N at diagnosis	
• N0	65 (63%)
• Other value	39 (37%)
Histological Grade	
• Low	91 (88%)
• Moderate	5 (5%)
• High	8 (7%)
Vascular permeability	
• 0/No	75 (72%)
• 1/ Yes	28 (27%)
• Unknown	1 (1%)
Tumor budding	
• Absent/No/Low	85 (82%)
• Moderate/High/Marked	18 (17%)
• Unknown	1 (1%)
Growth pattern	
• Expansive	24 (23%)
• Infiltrative	79 (76%)
• Unknown	1 (1%)
Tumor Laterality	
• Right	46 (44%)
• Left	40 (39%)
• Recto-sigma	1 (1%)
• Unknown	17 (16%)
Stage	
• I	16 (15%)
• II	40 (39%)
• III	44 (42%)
• IV	2 (2%)
• Unknown	2 (2%)
Relapse	
• Yes	19 (18%)
• No	85 (82%)
Situation	
• Death	12 (11%)
• Alive	90 (87%)
• Unknown	2 (2%)

Figure 1

Clinical and imaging data of the cohort of colon cancer patients. A) In yellow segmentation of CT images of the primary tumor of patient 82 using Quibim Precision® platform. B) Classification models tested in this study. C) Clinicopathological characteristics of the patient cohort.

FIGURE 2

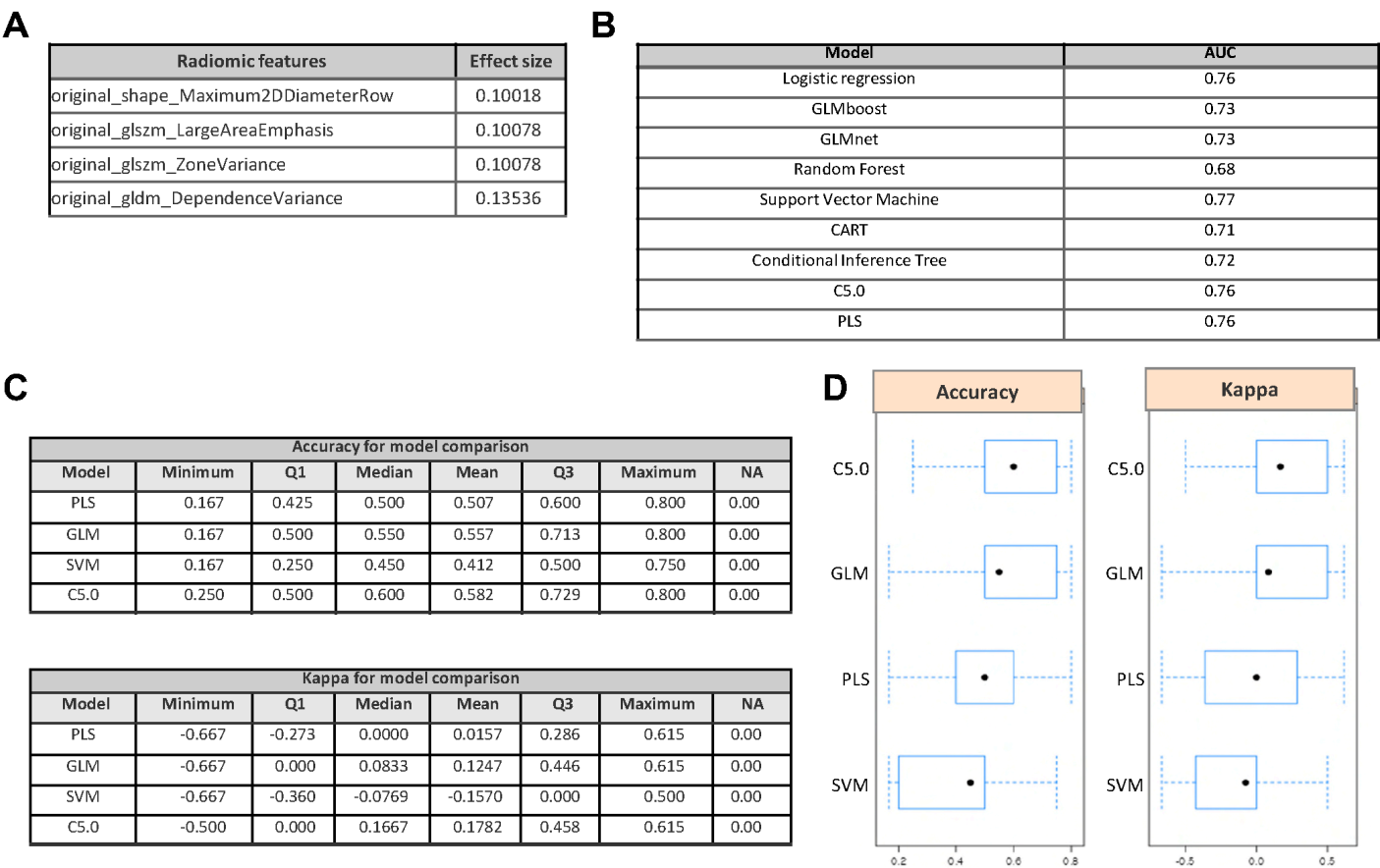


Figure 2

Radiomic data and model performance for the stratification of stage II and stage III colon cancer patients. A) Radiomic features included in the patient stratification analysis. B) AUC values of the different classification models tested for stage II vs. stage III stratification. C) Comparison of model performance using accuracy and kappa metrics for the endpoint of stage II vs. stage III stratification. D) Boxplots comparing classification models based on accuracy and kappa metrics for the endpoint of stage II vs. stage III stratification.

FIGURE 3

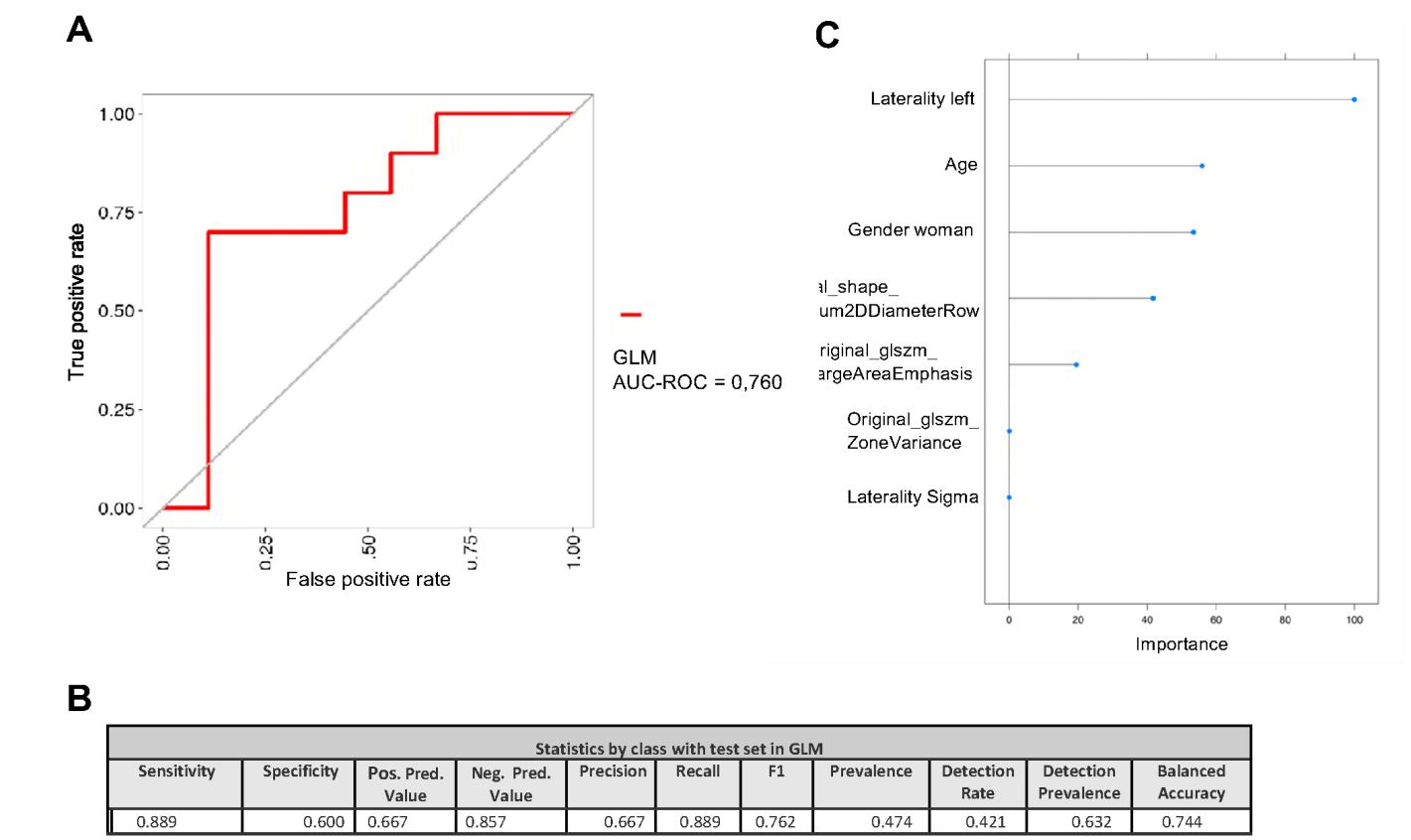


Figure 3

Preoperative radiomic data stratify colon cancer patients into stage II and stage III. A) ROC curve for the GLM model in the stratification between stages II and III. B) Class statistics for the test set using the GLM model for patient stratification. C) Variable importance in the GLM model for patient stratification.

FIGURE 4

A

Radiomic features	p-value	Effect size
diagnostics_Mask-interpolated_Mean	0.017	0.3601
diagnostics_Mask-interpolated_Maximum	0.010	0.3869
original_shape_LeastAxisLength	0.005	0.4217
original_shape_Maximum2DDiameterColumn	0.038	0.3139
original_shape_Maximum2DDiameterRow	0.012	0.3775
original_shape_Maximum3DDiameter	0.018	0.3561
original_shape_MeshVolume	0.007	0.4110
original_shape_MinorAxisLength	0.038	0.3133
original_shape_SurfaceArea	0.009	0.3922
original_shape_VoxelVolume	0.007	0.4110
original_firstorder_10Percentile	0.016	0.3641
original_firstorder_90Percentile	0.021	0.3494
original_firstorder_Energy	0.003	0.4511
original_firstorder_Maximum	0.010	0.3869
original_firstorder_Mean	0.017	0.3601
original_firstorder_Median	0.016	0.3628
original_firstorder_RootMeanSquared	0.017	0.3601
original_firstorder_TotalEnergy	0.003	0.4511
original_glrlm_GrayLevelNonUniformity	0.010	0.3882
original_glrlm_LongRunHighGrayLevelEmphasis	0.037	0.3159
original_glrlm_RunLengthNonUniformity	0.016	0.3655
original_glrlm_RunVariance	0.049	0.2972
original_glszm_LargeAreaEmphasis	0.010	0.3909
original_glszm_LargeAreaHighGrayLevelEmphasis	0.003	0.4444
original_glszm_ZoneVariance	0.010	0.3909
original_gldm_DependenceNonUniformity	0.017	0.3601
original_gldm_GrayLevelNonUniformity	0.008	0.3989
original_gldm_LargeDependenceHighGrayLevelEmphasis	0.022	0.3454

Note. $H_a: \mu_{NO} \neq \mu_{SI}$

B

Model	AUC
Logistic regression	0.58
GLMboost	0.62
GLMnet	0.62
Random Forest	0.69
Support Vector Machine (SVM)	0.73
CART	0.50
Conditional Inference Tree	0.50
C5.0	0.65
PLS	0.91

Figure 4

Radiomic data and model performance for the prediction of 5-year relapse in colon cancer patients. A) Radiomic features included in the 5-year relapse prediction analysis. B) AUC values of the different classification models tested for 5-year relapse prediction.

FIGURE 5

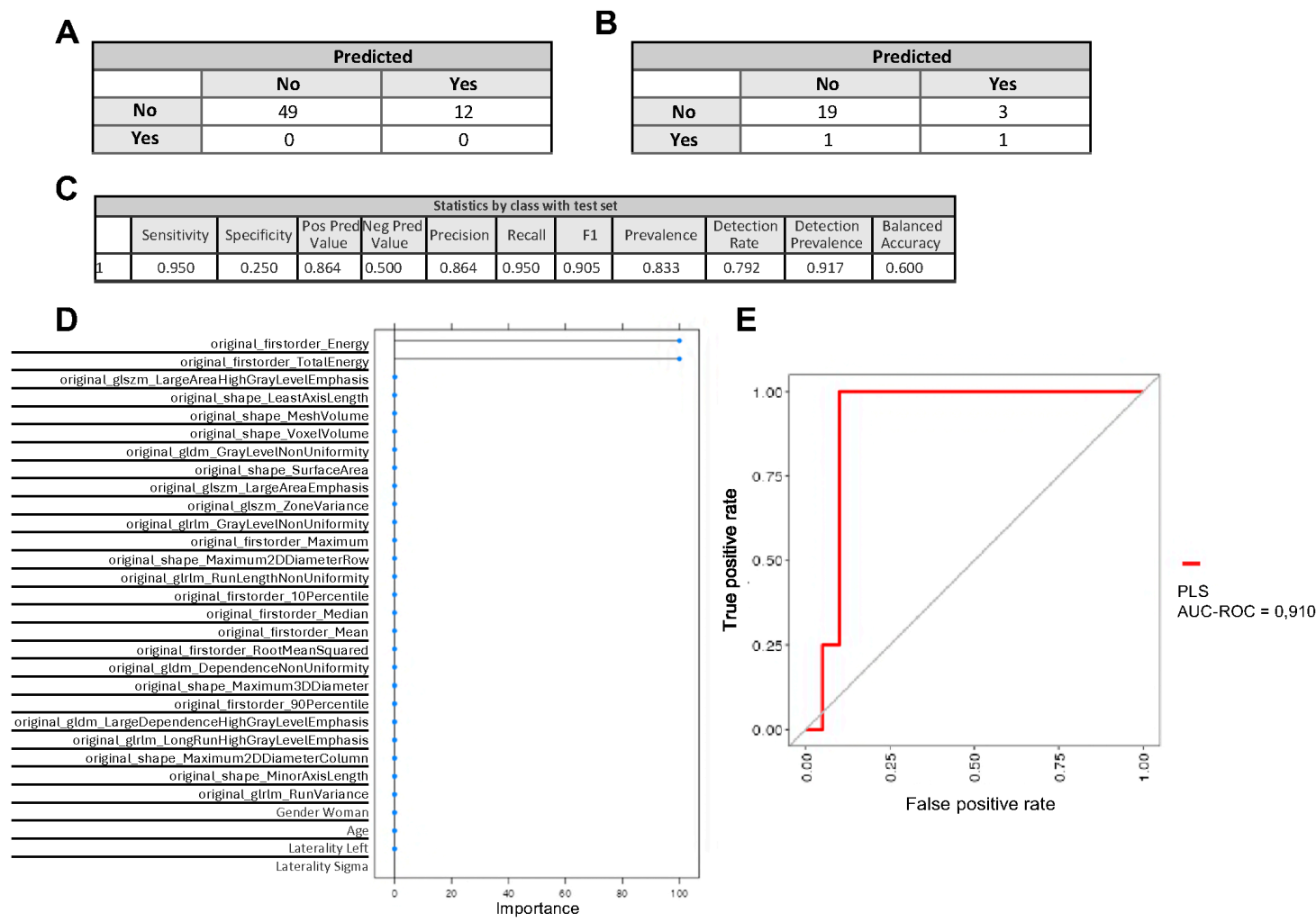


Figure 5

Partial Least Squares Discriminant Analysis (PLS) predicts 5-year relapse in colon cancer patients. A) Confusion matrix for the training set using the PLS model for 5-year relapse prediction. B) Confusion matrix for the test set using the PLS model for 5-year relapse prediction. C) Class statistics for the test set in the PLS model for 5-year relapse prediction. D) Variable importance in the PLS model for 5-year relapse prediction. E) ROC curve for 5-year relapse prediction using the PLS model.

FIGURE 6

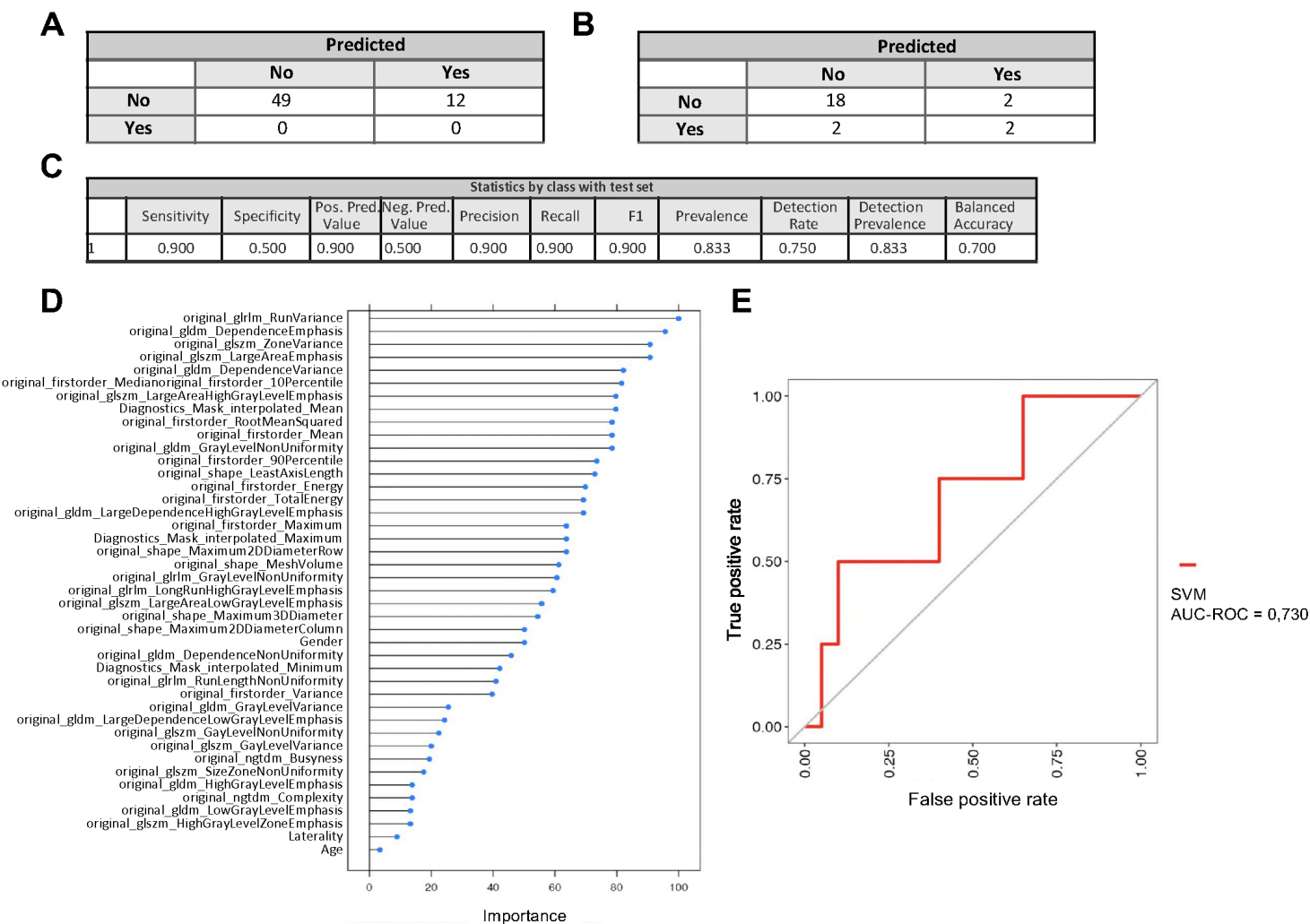


Figure 6

Support Vector Machine Analysis (SVM) predicts 5-year relapse in colon cancer patients. A) Confusion matrix for the training set using the SVM model for 5-year relapse prediction. B) Confusion matrix for the test set using the SVM model for 5-year relapse prediction. C) Class statistics for the test set in the SVM model for 5-year relapse prediction. D) Variable importance in the SVM model for 5-year relapse prediction. E) ROC curve for 5-year relapse prediction using the SVM model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [C.etalSupplementalMaterial.ods](#)
- [C.etalSupplementaryFigure1.odp](#)
- [C.etalSupplementaryFigure2.odp](#)

- [C.etalGraphicalAbstract.pptx](#)