Quantitative Tumor Heterogeneity MRI Profiling Improves Machine learning-based Prognostication in Patients with Metastatic Colon Cancer

Cite this article as: Dania Daye, MD PhD; Azadeh Tabari, MD; Hyunji Kim; Ken Chang; Sophia C. Kamran, MD; Theodore S. Hong, MD; Jayashree Kalpathy-Cramer, PhD; Michael S. Gee, MD, PhD, Quantitative Tumor Heterogeneity MRI Profiling Improves Machine learning-based Prognostication in Patients with Metastatic Colon Cancer, *European Radiology*, doi: 10.1007/s00330-020-07673-0

Quantitative Tumor Heterogeneity MRI Profiling Improves Machine learning-based Prognostication in Patients with Metastatic Colon Cancer

Dania **Daye**, MD PhD[1#]; Azadeh **Tabari**, MD[1#]; Hyunji **Kim**[1,2]; Ken **Chang**[1]; Sophia C. **Kamran**, MD[3]; Theodore S. **Hong**, MD[3]; Jayashree **Kalpathy-Cramer**, PhD[1]; Michael S. **Gee**, MD, PhD[1]

1. Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
2. Massachusetts Institute of Technology, Boston, MA, USA
3. Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

# Authors contributed equally

Corresponding Author

Dania Daye, MD, PhD

Email: ddaye@mgh.harvard.edu

Address: Massachusetts General Hospital

55 Fruit Street, GRB #290, Boston, MA 02114

ABSTRACT

OBJECTIVES: Intra-tumor heterogeneity has been previously shown to be an independent predictor of patient survival. The goal of this study is to assess the role of quantitative MRI-based measures of intra-tumor heterogeneity as predictors of survival in patients with metastatic colorectal cancer.

METHODS: In this IRB-approved retrospective study, we identified 55 patients with stage 4 colon cancer with known hepatic metastasis on MRI. 94 metastatic hepatic lesions were identified on post-contrast images and manually volumetrically segmented. A heterogeneity phenotype vector was extracted from each lesion. Univariate regression analysis was used to assess the contribution of 110 extracted features to survival prediction. A random forest-based machine learning technique was applied to the feature vector and to the standard prognostic clinical and pathologic variables. The dataset was divided into a training and test set at a ratio of 4:1. ROC analysis and confusion matrix analysis were used to assess classification performance.

RESULTS: Mean survival time was 39±3.9 months for the study population. A total of 22 textures features were associated with patient survival ($p<0.05$). The trained random forest machine learning model that included standard clinical and pathological prognostic variables resulted in an area under the ROC curve of 0.83. A model that adds imaging-based heterogeneity features to the clinical and pathological variables resulted in improved model performance for survival prediction with an AUC of 0.94.

CONCLUSIONS: MRI-based texture features are associated with patient outcomes and improve the performance of standard clinical and pathological variables for predicting patient survival in metastatic colorectal cancer.

Keywords:

1. colorectal cancer;
2. radiomics;
3. MRI

Key points:

- MRI-based tumor heterogeneity texture features are associated with patient survival outcomes.
- MRI-based tumor texture features complement standard clinical and pathological variables for prognosis prediction in metastatic colorectal cancer.
- Agglomerative hierarchical clustering shows that patient survival outcomes are associated with different MRI tumor profiles.

Abbreviations:

CRLM: colorectal liver metastases;

GLCM: gray level co-occurrence matrix;

GLDM: gray level dependence matrix features.

GLRLM: gray level run time length matrix;

GLSZM: gray level size zone matrix;

MRI: magnetic resonance imaging;

MSI: microsatellite instability;

NGTDM: neighboring gray tone difference matrix;

Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related death in the United States and one of the most common cancers in the Western world [1]. Approximately 140,000 patients are diagnosed with CRC annually in the US [2]. 25% of patients have colorectal liver metastases (CRLM) at the time of diagnosis, with another 50-60% of patients later developing metachronous CRLM [3]. Surgical resection is most effective in achieving long-term survival in patient with CRLM. However, up to 30% of CRC patients who undergo surgical resection experience a subsequent relapse within three years, with a median survival of 12 months. A variety of potential treatment strategies are available in the setting of hepatic oligometastatic disease, including multiple liver directed options such as surgical resection, percutaneous image-guided therapies, and stereotactic radiotherapy, with treatment options continuing to evolve and improve. To better tailor treatment decisions, better methods of tumor characterization and risk stratification are needed to personalize patient management as proposed by the important areas of research put forth by the American Society of Clinical Oncology [2].

At present, the gold standard for clinical prognostication in patients with CRC is based on the American Joint Committee on Cancer (AJCC) staging system, which includes the invasion extent of the primary tumor (T stage), lymph node status (N stage), and distant spread (M stage). Currently, the role of imaging is restricted to assessment of presence and size of lesions for staging purposes. Intra-tumor heterogeneity is an emerging tumor characteristic assessing differences in voxel signal intensities with a lesion that has been correlated with poor patient prognosis [4, 5]. Tumor heterogeneity metrics reflect regional changes within tumors, and it likely that these features enable imaging to provide additional prognostic information beyond lesion size.

Recently, machine learning-based approaches through analyzing large numbers of sub-visual image features (termed 'radiomics') have been identified as a new method for lesion characterization and for the quantitative assessment of intra-tumor heterogeneity. Radiomics analysis refers to the extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support. Many applications of machine learning have been reported in cancer diagnosis [6-10]. Advances in computing power have given radiomics, the potential for more accurate characterization of metastatic lesions that could positively help guide management strategies. Radiomic features of focal liver lesions may provide information about underlying tumor biology [6, 11] that potentially could provide additional prognostic information in patients with metastatic CRC. The objective of this study was to investigate the association between radiomic features of CRLMs and patient survival and to design a prognostic model integrating radiomics for the prediction of patient outcomes.

Methods

Study Design

This retrospective study was approved by the Institutional Review Board (IRB), which waived the requirement for patient informed consent, and was conducted in accordance with the Health Insurance Portability and Accountability Act guidelines for research.

This study retrospectively identified patients with biopsy-proven metastatic stage IV colorectal lesions to the liver treated with standard chemotherapy protocols (including FOLFOX and

FOLFIRI) as well as radiotherapy at a single institution during the years 2007 to 2013 based on a query of a database of colorectal cancer patients undergoing radiation therapy. Colorectal cancer cases were identified through review of pathology reports in the electronic health record. Date of diagnosis was taken as the date of colonoscopy biopsy or surgical pathology report of the resected tumor. TNM staging information was collected from the clinical or pathology report at the time of primary tumor diagnosis. Imaging follow-up was obtained through the picture archiving and communication system (PACS). All patients underwent contrast-enhanced MRI as part of their routine diagnostic workup. Inclusion criteria consisted of (a) histopathologically confirmed colorectal adenocarcinoma; (b) presence of biopsy-proven hepatic metastases on the portal venous phase post-contrast T1-weighted MRI subtraction sequence within 6 months prior to starting treatment (with at least one lesion measuring > 1 cm in longest diameter), (c) patients treated by an oncologist in the MGH Cancer Center to ensure availability of follow-up data. The portal venous phase was chosen as metastatic colorectal tumors are thought to be best visualized during that phase of imaging, in contrast to hepatomas which are typically better visualized on the arterial phase.

MR protocol and acquisition parameters

All the liver MRI examinations of the abdomen were performed on either a 3T magnet (n=8; Discovery 750MR GE Medical Systems and Magnetom Trio, Siemens Healthcare) or a 1.5T MRI scanner (n=47; Signa HDx, GE Medical Systems and Magnetom Avanto, Siemens Healthcare) using a body coil positioned over the abdomen. Slice thickness was 4 mm and matrix size was 320x256. Contrast enhanced volumetric T1-weighted fat-suppressed sequences (LAVA or VIBE) were obtained at multiple timepoints post-contrast as per standard clinical protocol. For this study, the portal venous phase (60-70 seconds post-contrast) images were selected for

quantitative analysis. Intravenous gadolinium contrast (gadopentetate/Magnevist, Bayer; gadoterate/Dotarem, Guerbet; or gadoxetate/Eovist, Bayer) was administered at the standard approved clinical dose (0.1 mmol/kg for gadopentetate and gadoterate, 0.025 mmol/kg for gadoxetate) by power injector at a rate of 1 mL/second. Of note, these MRI examinations were all conducted prior to the restriction on the use of linear gadolinium agents.

Data Collection

Standard clinical and pathologic prognostic variables extracted from the medical record included: patient age, patient gender, tumor KRAS status, tumor microsatellite instability (MSI), primary tumor site, additional metastasis sites and chemotherapy regimen. The images were reviewed by a radiologist to identify the hepatic metastatic lesions that met the inclusion criteria for the study. All metastatic hepatic lesions greater than 1 cm in greatest diameter were identified. The assessed outcome for the study was all-cause mortality within the study period. Follow-up ended with the most recent clinic visit or most recent imaging study before the study end period or with the patient's death.

Image Segmentation and Texture Analysis

Volume-based hepatic lesion segmentation was performed using an open source volumetric image analysis software platform (3D slicer; http://www.slicer.org) [12]. The brightness and contrast levels were user-selected to optimize visualization of the liver mass. All measurements were performed on the portal venous phase T1-weighted fat-suppressed post-

contrast axial images. MRI images were evaluated by a radiologist for the presence of metastatic hepatic lesions. All metastatic hepatic lesions greater than 1 cm in greatest diameter were included. Manual delineations and semi-automatic segmentations methods were applied to ensure the reproducibility of radiomic features resulting from the impact of segmentation methods. A multiparametric imaging phenotype vector was extracted from each lesion by using quantitative texture analysis. All images were normalized and an inhomogeneity correction algorithm was applied prior to extracting quantitative features, to address differences in data acquisition techniques.

A total of 110 radiomic features were extracted from all 94 hepatic metastatic lesions from the patients that met the inclusion criteria for the study. Those consisted of first-order features, 2D and 3D shape features, gray level co-occurrence matrix (GLCM) features, gray level size zone matrix (GLSZM), gray level run time length matrix (GLRLM) features, neighboring gray tone difference matrix (NGTDM) features and gray level dependence matrix features (GLDM). These features have been previously described [13]. A summary of the image analysis algorithm is included in Figure 1. Intensity standardization was first applied to all images to address any inter-subject intensity variations prior to computing the radiomics features. The extracted radiomics vector defined the MR fingerprint of each tumor. Hierarchical clustering of the fingerprints was subsequently performed and a dendrogram was generated using Matlab R2019b (Mathworks Inc.) to assess for inherent MR fingerprint differences between patients who remained alive at the end of the study and those who were deceased.

Statistical Analysis and Machine Learning

Kaplan-Meier survival analysis was performed to assess for patient survival estimates during the study period. Univariate logistic regression analysis was used to assess the contribution of each of the computed texture features to the prediction of patient mortality. A P-value < 0.05 on the Wald test was considered statistically significant. All statistical analyses and machine learning models were performed using Stata (Stata Statistical Software: Release 14. StataCorp. 2015. StataCorp LP).

A random forest-based machine learning model was used to predict patient survival outcome during the study period. The dataset was divided into a training and testing set at a ratio of 4:1. Recursive feature elimination was used for feature selection. A model that included all features exhibited the best performance. A total of 3 models were assessed: a model that includes all the radiomic features, a model that includes only the clinical and pathology features as the current clinical management, and a model that includes the combination of both the radiomic features and the clinical and pathology features. If a patient had more than one tumor, all tumors were included in either in the training or testing set to avoid overfitting. The same training and testing sets were used for all 3 models. The depth of the random forest algorithm was set to be the square root of the number of features used. ROC analysis and confusion matrix metrics were used to assess classification performance. The machine learning model was implemented using Python 3.8.

Results

Patient population

A total of 55 consecutive patients with 94 metastatic tumors who met the inclusion criteria were included in the study (45% female, Average age 55.1 ±10.6 years). 72% (40/55) of the patients had KRAS-positive disease. 40% (22/55) of patients had high microsatellite instability. 67% (37/55) of the primary colorectal cancer lesions were in the descending or sigmoid colon; 29% (16/55) were in the ascending colon and 4% (2/55) were in the transverse colon. 50.5 % (28/55) of the patients had extrahepatic metastases, while the remaining had metastatic disease limited to the liver only. 90% of the study population was treated with FOLFIRI or FOLFOX-based chemotherapy regimen and all received radiation treatment of their hepatic disease. A summary of the patient characteristics is included in Table 1.

Survival analysis

Patient survival data was available for up to 95 months. The mean follow-up duration was 34.5 ± 19.6 months. 45% of patients were deceased at the end of the study. The median overall survival was 37 months (95% CI: 35-52) months. Survivor function at 12, 36 and 60 months were 97%, 51% and 37% respectively. The Kaplan-Meier curve illustrating the survival estimates is included in Figure 2 with the number at-risk in the study population at each time point.

Radiomic Features

A total of 110 radiomic features were extracted from each segmented tumor including first-order features, 2D and 3D shape features, gray level co-occurrence matrix (GLCM) features, gray level size zone matrix (GLSZM), gray level run time length matrix (GLRLM) features, neighboring gray tone difference matrix (NGTDM) features and gray level dependence matrix features (GLDM). Univariate logistic regression analysis was performed to assess the contribution of each feature to the prediction of the survival outcome variable in the patient population. 22 of the texture features were predictive of patient survival on univariate logistic regression analysis ($P<0.05$ on Wald test.). Eight first order features, six GLCM features, two GLDM features, one GLRM features, four GLZM features and one NGTM feature contributed significantly to survival prediction. A summary of the logistic regression results for the statistically significant features are included in table 2.

An agglomerative hierarchical cluster analysis was used to assess if the MRI tumor fingerprints result in patient clustering into two separate groups with different survival outcomes. Indeed, hierarchical clustering resulted into two distinct clusters of signatures where tumors were from predominantly deceased patients in the first cluster (~75%) while tumors in the second cluster were from predominantly living patients at the end of the study (~68%). All tumors from the same patient belonged to the same cluster. The resulting dendrogram of the hierarchical clustering analysis is included in Figure 3.

Machine Learning for Survival Prediction

We included all radiomic features in a random forest-based machine learning model to predict survival outcome in the study population. The dataset was divided into a training (n=76 tumors) and testing (n=18 tumors) set at a ratio of 4:1. All radiomics features were included in the final models. A random forest-based machine learning algorithm that included all radiomic features resulted in an area under the curve (AUC) of 0.93 in the testing dataset (Figure 4A). Confusion matrix analysis on the testing dataset revealed a sensitivity of 50%, a specificity of 100%, a negative predictive value of 63% and a positive predictive value of 100%.

Subsequently, we built 2 additional random forest-based machine learning models. The first included clinical and pathology variables currently used for clinical prognostication. The variables consisted of the stage of the disease at time of primary tumor diagnosis, tumor KRAS mutation status, satellite instability, colon cancer primary site, chemotherapy regimen used in addition to patient age and gender. The second model included all the radiomics features in addition to the above clinical and pathology features. When applied to the testing dataset, the latter model that combines both radiomics and standard clinical and pathology variables exhibited improved performance for survival prediction (AUC = 0.94) when compared to the clinical and pathology variables that are the current gold standard for clinical prognostication (AUC =0.83) (Figure 4B). When assessing the model that includes both the radiomic and clinical/pathology variables, confusion matrix analysis on the testing dataset resulted in a sensitivity was 50%, a specificity of 100%, a negative predictive value of 59% and a positive predictive value of 100%.

In Figure 5, we include two representative MRI images illustrating visually striking differences in tumor heterogeneity between a patient who survived 26 months and another who succumbed

at 8 months. Interestingly, both patients had a similarly sized metastatic tumors in the right hepatic lobe and the main difference on imaging was a difference in tumor heterogeneity.

Discussion

This work represents a preliminary study assessing the role of MRI-based radiomic features in predicting patient survival in metastatic colorectal cancer. The analysis of tumor heterogeneity parameters derived from T1-weighted contrast-enhanced liver MRI of CRC patients showed that several features of intratumoral heterogeneity are associated with patient survival. In addition, our results demonstrate that a radiomics-based multivariate classifier that includes the distinct MRI fingerprint of metastatic tumors can complement and improve patient prognosis prediction when compared to the current clinical gold standard that consists of clinical, pathology and genetic variables. Radiomics-based tumor profiling may be beneficial for informing treatment choices and tailoring personalized therapy decisions for patients with CRLM.

Inter- and intra-tumor heterogeneity have been implicated in the development of therapy resistance after an initial response and in the development of metastatic disease [6, 14-18]. Clonal heterogeneity often occurs before tumor progression is apparent on clinical imaging and is the most substantial obstacle to patient treatment with targeted therapy [19, 20]. Radiomics-based MRI profiling, as presented in this study, may capture intra-tumor heterogeneity. At present, current biopsy methods are unable to capture intra-tumor heterogeneity due to the small volume of tissue sampled, while standard cross-sectional imaging interpretation only documents presence and size of metastatic lesions [17]. Methods to identify biologically aggressive metastases could guide treatment selection and potentially improve patient prognosis [21]. This

is particularly relevant given the number of systemic molecule-targeted agents, as well as lesion-directed therapies, that are in clinical practice for metastatic disease [17].

The field of texture analysis provides quantitative measures of the spatial arrangement and distribution of gray-level intensities in a selected region of interest (ROI) [22, 23]. Tumor texture analysis on CT and MRI has been previously used to quantify tumor heterogeneity in breast cancer, lung cancer, glioblastoma multiforme, colon cancer and prostate cancer [4, 24-31], and has been shown to improve tumor staging and therapy response assessment [32-35]. A recent study showed that texture parameters derived from T2-weighted images of primary rectal cancer have added value as imaging biomarkers of tumor response to neoadjuvant chemotherapy [34].

Our study adds to the existing evidence in the literature by showing that MRI-based quantitative heterogeneity measures within hepatic metastases are associated with patient survival in metastatic colorectal cancer and may add information beyond genetic mutation status to optimize prognosis prediction. This study demonstrates that imaging-based heterogeneity features improve survival prediction of the patients with stage IV colorectal cancer when added to clinical and pathology variables. In addition, our analysis is based on volumetric analysis of the entire tumor volume, rather than analysis of a single 2D image slice that has been used in prior studies. Volumetric analysis allows more complete lesion evaluation and is less prone to sampling errors associated with single slice analysis.

This study has several limitations. This is a retrospective proof-of-concept study with a relatively small patient population consisting of patients that has undergone radiation as part of their treatment. A future prospective study in a larger and broader patient population will be needed to validate our findings. Although separate training and testing sets were used, the study was trained and tested in the same patient population. Validation of the results in an independent

external dataset will be needed to ensure the generalizability of our findings. Finally, tumor delineation is an important aspect of the radiomics workflow. Although volumetric whole lesion analysis was performed, the tumors were manually contoured and potential variation in tumor contouring could affect the extracted feature values. In this study, this issue was mitigated by having one radiologist assess tumor contouring on every slice to ensure accuracy. In the future, automatic segmentation software tools can help to minimize variations in tumor contouring by humans and also decrease overall time for lesion analysis.

Our preliminary results showed that radiomic features reflecting tumor heterogeneity may play a role in predicting patient survival in patients with metastatic colorectal lesions. Defining and validating imaging-based heterogeneity features based on the appearance of metastatic lesions has the potential to improve treatment response prediction and assessment for individual lesions, allowing for an effective targeted therapy and moving towards a personalized treatment in patients with colorectal cancer. Further prospective multicenter trials are needed to achieve a large-scale validation of our results in clinical practice.

Funding

The authors state that this work has not received any funding.

Compliance with Ethical Standards

Guarantor:

The scientific guarantor of this publication is Dania Daye, MD, PhD

Conflict of Interest:

The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and Biometry:

One of the authors has significant statistical expertise.

Informed Consent:

Written informed consent was waived by the Institutional Review Board.

Ethical Approval:

Institutional Review Board approval was obtained.

Methodology

- retrospective
- cross sectional study
- performed at one institution

REFERENCES:

1. Van Cutsem E, Cervantes A, Nordlinger B, Arnold D: Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2014, 25:iii1-iii9.
2. Siegel R, DeSantis C, Virgo K, Stein K, Mariotto A, Smith T, Cooper D, Gansler T, Lerro C, Fedewa S et al: Cancer treatment and survivorship statistics, 2012. CA: A Cancer Journal for Clinicians 2012, 62(4):220-241
3. Mella J, Biffin A, Radcliffe AG, Stamataki JD, Steele RJC: **Population-based audit of colorectal cancer management in two UK health regions**. *British Journal of Surgery* 1997, **84**(12):1731-1736.
4. Davnall F, Yip CSP, Ljungqvist G, Selmi M, Ng F, Sanghera B, Ganeshan B, Miles KA, Cook GJ, Goh V: **Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice?** *Insights into Imaging* 2012, **3**(6):573-589.
5. García-Figueiras R, Baleato-González S, Padhani AR, Luna-Alcalá A, Vallejo-Casas JA, Sala E, Vilanova JC, Koh D-M, Herranz-Carnero M, Vargas HA: **How clinical imaging can assess cancer biology**. *Insights into Imaging* 2019, **10**(1).
6. Gillies RJ, Kinahan PE, Hricak H: **Radiomics: Images Are More than Pictures, They Are Data**. *Radiology* 2016, **278**(2):563-577.
7. Peng Y, Jiang Y, Yang C, Brown JB, Antic T, Sethi I, Schmid-Tannwald C, Giger ML, Eggener SE, Oto A: **Quantitative Analysis of Multiparametric Prostate MR Images: Differentiation between Prostate Cancer and Normal Tissue and Correlation with Gleason Score—A Computer-aided Diagnosis Development Study**. *Radiology* 2013, **267**(3):787-796.
8. Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, Wick A, Eidel O, Schlemmer H-P, Radbruch A *et al*: **Radiogenomics of Glioblastoma: Machine Learning–based Classification of Molecular Characteristics by Using Multiparametric and Multiregional MR Imaging Features**. *Radiology* 2016, **281**(3):907-918.
9. Zhang Y-D, Wang J, Wu C-J, Bao M-L, Li H, Wang X-N, Tao J, Shi H-B: **An imaging-based approach predicts clinical outcomes in prostate cancer through a novel support vector machine classification**. *Oncotarget* 2016, **7**(47):78140-78151.
10. Daye D, Staziaki PV, Furtado VF, Tabari A, Fintelmann FJ, Frenk NE, Shyn P, Tuncali K, Silverman S, Arellano R *et al*: **CT Texture Analysis and Machine Learning Improve Post-ablation Prognostication in Patients with Adrenal Metastases: A Proof of Concept**. *CardioVascular and Interventional Radiology* 2019, **42**(12):1771-1776.
11. Ganeshan B, Miles KA, Young RCD, Chatwin CR: **In Search of Biologic Correlates for Liver Texture on Portal-Phase CT**. *Academic Radiology* 2007, **14**(9):1058-1068.
12. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M *et al*: **3D Slicer as an image computing platform for the Quantitative Imaging Network**. *Magnetic Resonance Imaging* 2012, **30**(9):1323-1341.

13. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin J-C, Pieper S, Aerts HJWL: **Computational Radiomics System to Decode the Radiographic Phenotype**. *Cancer Research* 2017, **77**(21):e104-e107.
14. Bedard PL, Hansen AR, Ratain MJ, Siu LL: **Tumour heterogeneity in the clinic**. *Nature* 2013, **501**(7467):355-364.
15. Turner NC, Reis-Filho JS: **Genetic heterogeneity and cancer drug resistance**. *The Lancet Oncology* 2012, **13**(4):e178-e185.
16. Morris LGT, Riaz N, Desrichard A, Şenbabaoğlu Y, Hakimi AA, Makarov V, Reis-Filho JS, Chan TA: **Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival**. *Oncotarget* 2016, **7**(9):10051-10063.
17. Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, Ahronian LG, Mussolin B, Kwak EL, Buscarino M, Lazzari L *et al*: **Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in Colorectal Cancer**. *Cancer Discovery* 2015, **6**(2):147-153.
18. O'Connor JPB, Rose CJ, Waterton JC, Carano RAD, Parker GJM, Jackson A: **Imaging Intratumor Heterogeneity: Role in Therapy Response, Resistance, and Clinical Outcome**. *Clinical Cancer Research* 2014, **21**(2):249-257.
19. Gerlinger M RA, Horswell S, et al.: **Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing**. *New England Journal of Medicine* 2012, **367**(10):976-976.
20. Strickler JH, Banks KC, Nagy RJ, Lanman RB, Talasaz A, Corcoran RB, Kopetz S: **Blood-based genomic profiling of circulating cell-free tumor DNA (ctDNA) in 1397 patients (pts) with colorectal cancer (CRC)**. *Journal of Clinical Oncology* 2017, **35**(4_suppl):584-584.
21. Simone JV: **Personalized Medicine**. *Oncology Times* 2008, **30**(9):4.
22. Brown RA, Frayne R: **A comparison of texture quantification techniques based on the Fourier and S transforms**. *Medical Physics* 2008, **35**(11):4998-5008.
23. Castellano G, Bonilha L, Li LM, Cendes F: **Texture analysis of medical images**. *Clinical Radiology* 2004, **59**(12):1061-1069.
24. Al-Kadi OS, Watson D: **Texture Analysis of Aggressive and Nonaggressive Lung Tumor CE CT Images**. *IEEE Transactions on Biomedical Engineering* 2008, **55**(7):1822-1830.
25. Eliat P-A, Olivié D, Saïkali S, Carsin B, Saint-Jalmes H, de Certaines JD: **Can Dynamic Contrast-Enhanced Magnetic Resonance Imaging Combined with Texture Analysis Differentiate Malignant Glioneuronal Tumors from Other Glioblastoma?** *Neurology Research International* 2012, **2012**:1-7.
26. Lopes R, Ayache A, Makni N, Puech P, Villers A, Mordon S, Betrouni N: **Prostate cancer characterization on MR images using fractal features**. *Medical Physics* 2010, **38**(1):83-95.
27. Holli K, Lääperi A-L, Harrison L, Luukkaala T, Toivonen T, Ryymin P, Dastidar P, Soimakallio S, Eskola H: **Characterization of Breast Cancer Types by Texture Analysis of Magnetic Resonance Images**. *Academic Radiology* 2010, **17**(2):135-141.
28. Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, Davatzikos C: **Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme**. *Magnetic Resonance in Medicine* 2009, **62**(6):1609-1618.
29. Ganeshan B, Burnand K, Young R, Chatwin C, Miles K: **Dynamic Contrast-Enhanced Texture Analysis of the Liver**. *Investigative Radiology* 2011, **46**(3):160-168.
30. Guo Q, Shao J, Ruiz VF: **Characterization and classification of tumor lesions using computerized fractal-based texture analysis and support vector machines in digital mammograms**. *International Journal of Computer Assisted Radiology and Surgery* 2008, **4**(1):11-25.
31. Woods BJ, Clymer BD, Kurc T, Heverhagen JT, Stevens R, Orsdemir A, Bulan O, Knopp MV: **Malignant-lesion segmentation using 4D co-occurrence texture analysis applied to dynamic**

contrast-enhanced magnetic resonance breast image data. *Journal of Magnetic Resonance Imaging* 2007, **25**(3):495-501.

32. Alic L, van Vliet M, van Dijke CF, Eggermont AMM, Veenland JF, Niessen WJ: **Heterogeneity in DCE-MRI parametric maps: a biomarker for treatment response?** *Physics in Medicine and Biology* 2011, **56**(6):1601-1616.
33. Assefa D, Keller H, Ménard C, Laperriere N, Ferrari RJ, Yeung I: **Robust texture features for response monitoring of glioblastoma multiforme on T1-weighted and T2-FLAIR MR images: A preliminary investigation in terms of identification and segmentation**. *Medical Physics* 2010, **37**(4):1722-1736.
34. O'Connor JPB, Rose CJ, Jackson A, Watson Y, Cheung S, Maders F, Whitcher BJ, Roberts C, Buonaccorsi GA, Thompson G *et al*: **DCE-MRI biomarkers of tumour heterogeneity predict CRC liver metastasis shrinkage following bevacizumab and FOLFOX-6**. *British Journal of Cancer* 2011, **105**(1):139-145.
35. Ganeshan B, Abaleke S, Young RCD, Chatwin CR, Miles KA: **Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage**. *Cancer Imaging* 2010, **10**(1):137-143.

TABLE LEGENDS:

Table 1: Patient characteristics (SD: standard deviation).

Table 2: Logistic Regression analysis for survival prediction. Only radiomic features with statistically significant Wald test are shown. P-value of the Wald test in included in the last column. (C.I: confidence interval)

FIGURE LEGENDS:

Figure 1: Image Analysis pipeline used in this study. The pipeline consisted of tumor segmentation, feature extraction and random forest-based machine learning for survival prediction. T1-weighted contrast-enhanced MRI images in the portal venous phase used for analysis. All tumors were identified by a radiologist.

Figure 2: Kaplan-Meier survival estimate of the patients enrolled in the study. Follow-up data was available for 95 months. The mean survival time was 37 months.

Figure 3: Hierarchical Clustering of radiomic features from all patients enrolled in the study. A dendrogram revealed two distinct radiomic fingerprints in the extracted data. Predominantly deceased patients had a different radiomic signature when compared to those who were predominantly alive by the end of the study.

Figure 4: Receiver operator curve analysis of machine learning models. A) A random forest-based model of the radiomic features applied to the testing set for patient survival prediction revealed an area under the curve (AUC) of 0.93. B) A model that combines both clinical/pathology features with radiomic features for survival prediction (AUC=0.94) has improved performance compared to a model that only include clinical/pathology features (AUC=0.83).

Figure 5: Representative images illustrating differences in tumor heterogeneity between 2 patients with different survival outcomes. A) 44 year old male with a 4.2 cm KRAS-negative contrast-enhancing hepatic metastasis in the right hepatic lobe. This patient survived 36 months following imaging. B) 68 year old female with a 4.4 cm KRAS-positive contrast-enhancing hepatic metastasis in the right hepatic lobe. This patient succumbed at 8 months following imaging.

Figure 1

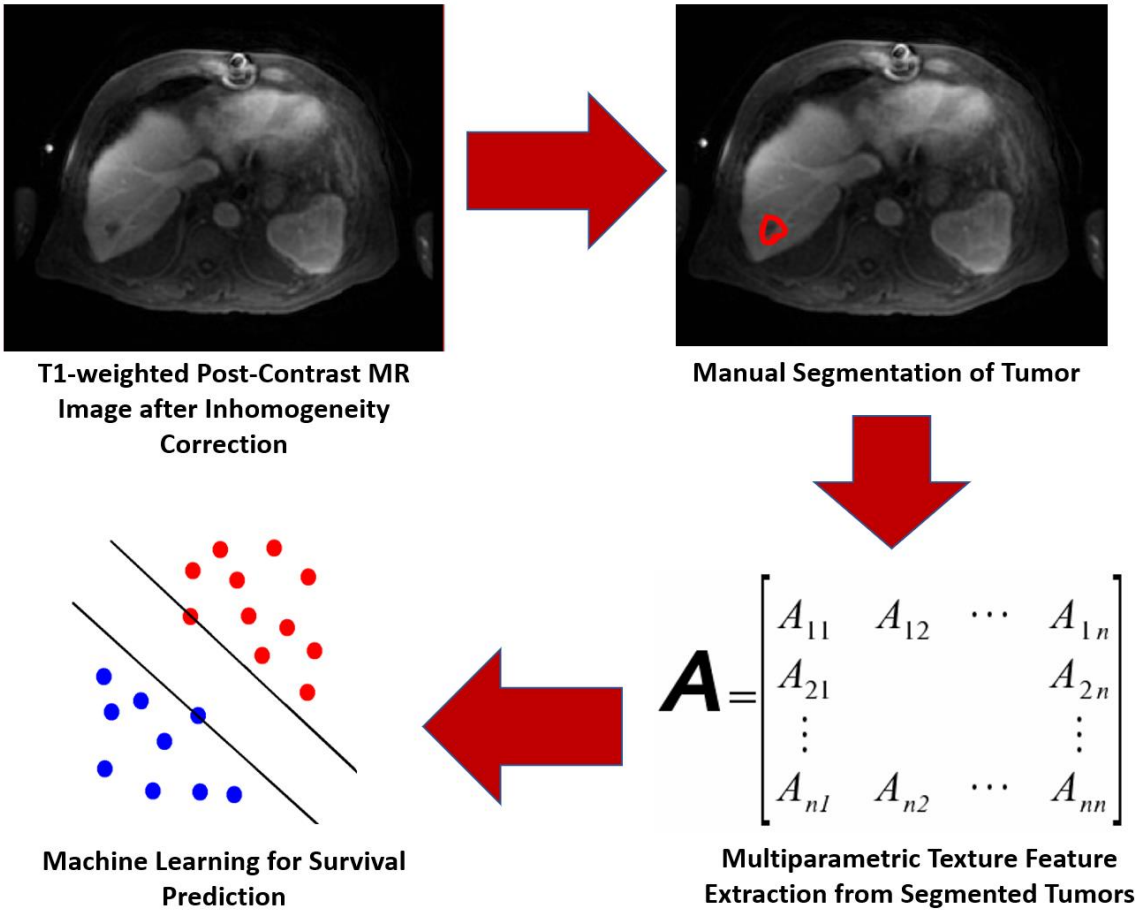


Figure 2

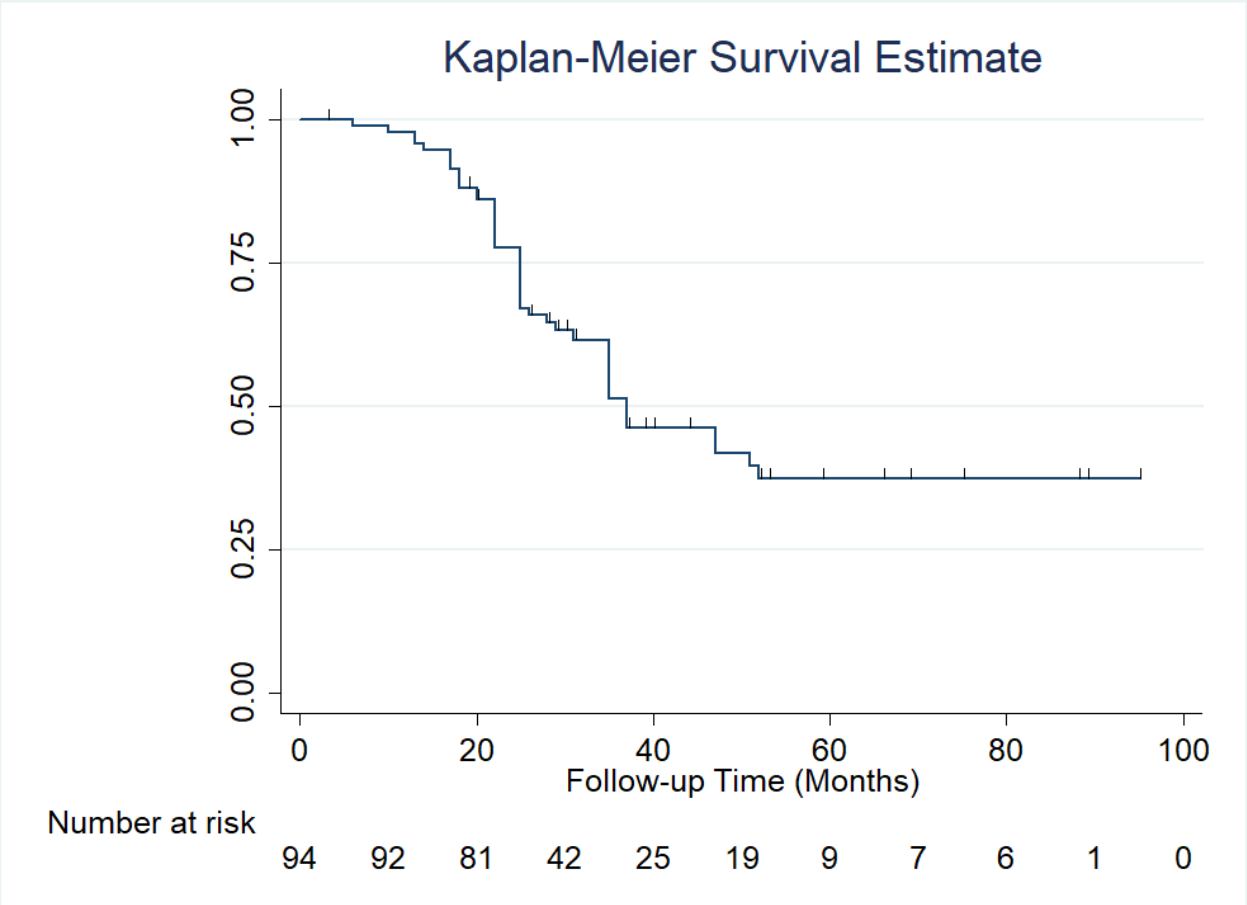


Figure 3

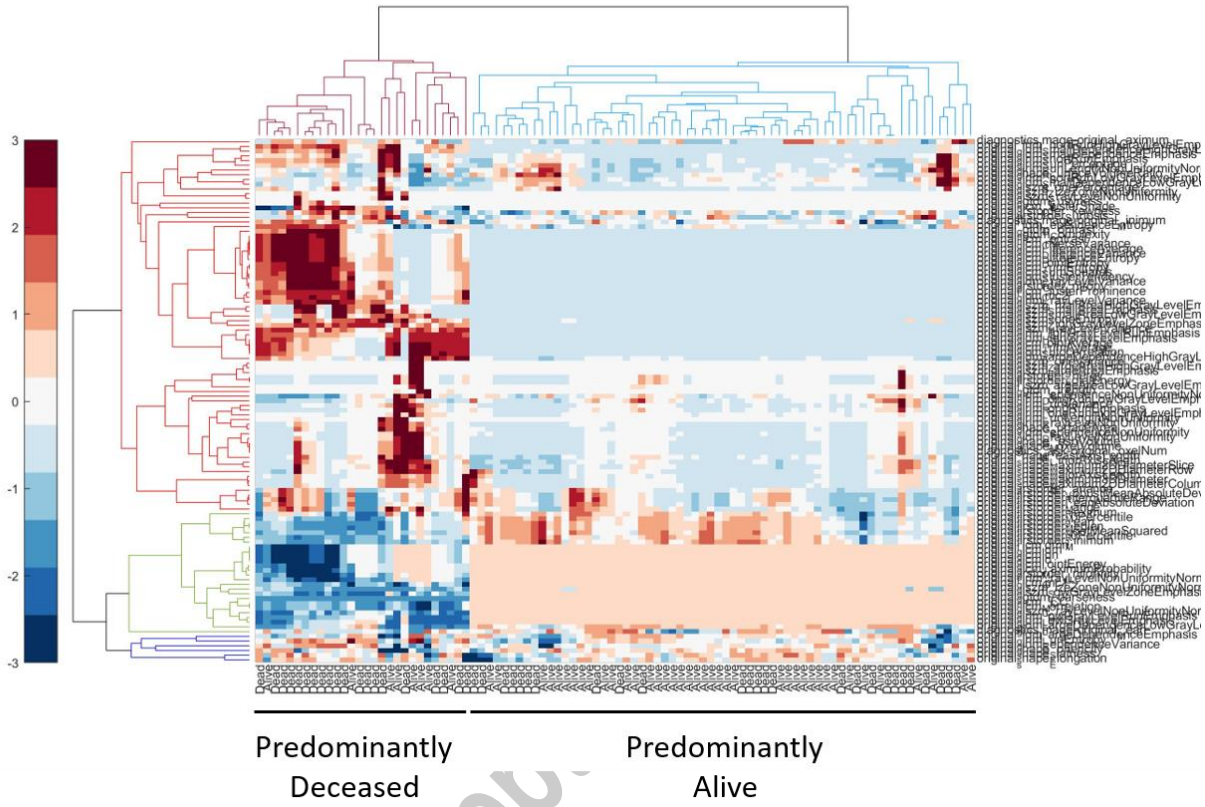
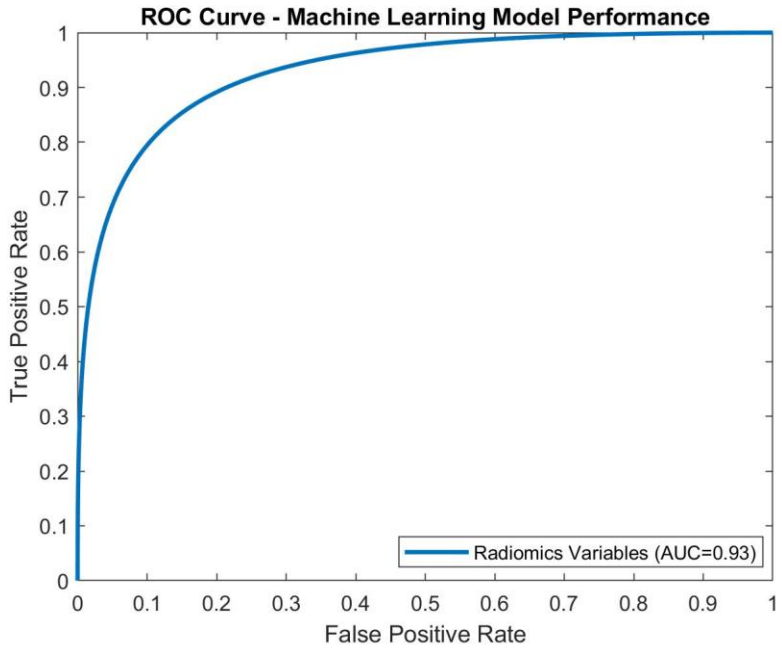


Figure 4

A



B

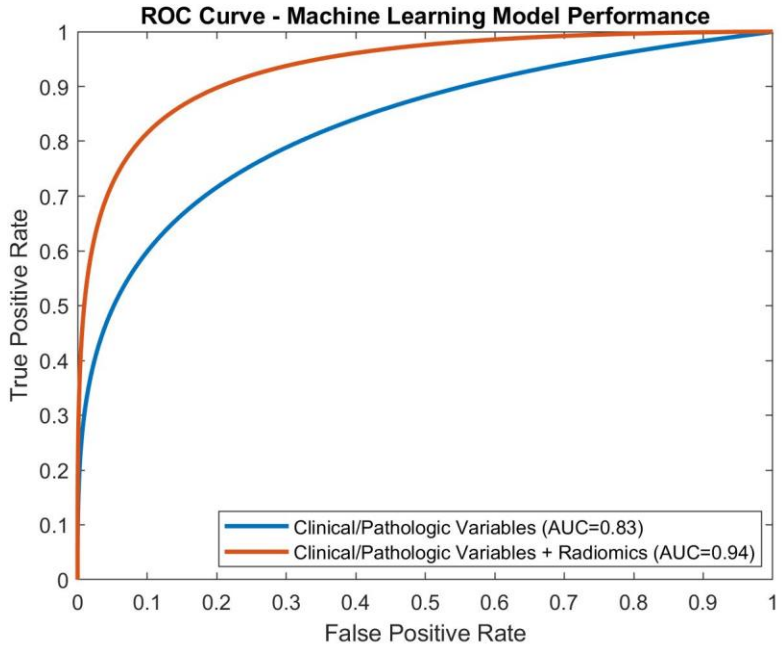
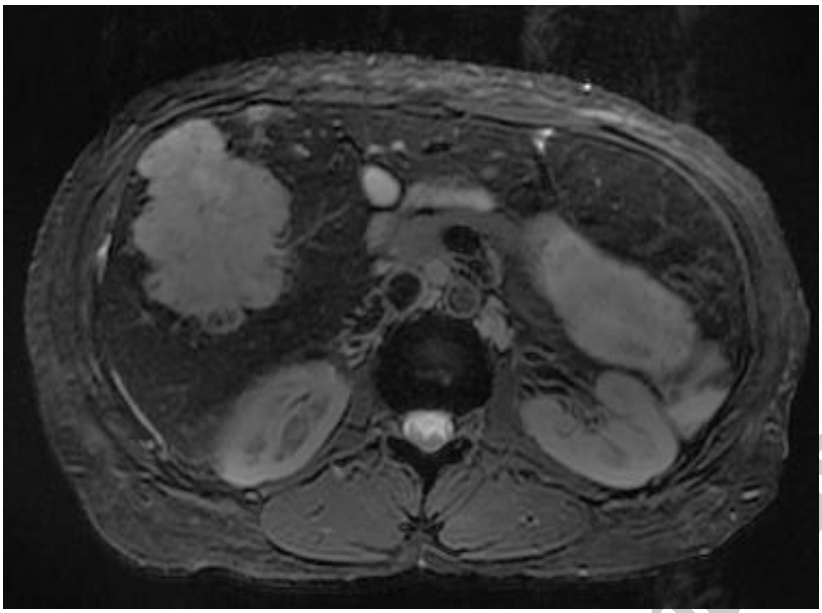


Figure 5

A



B

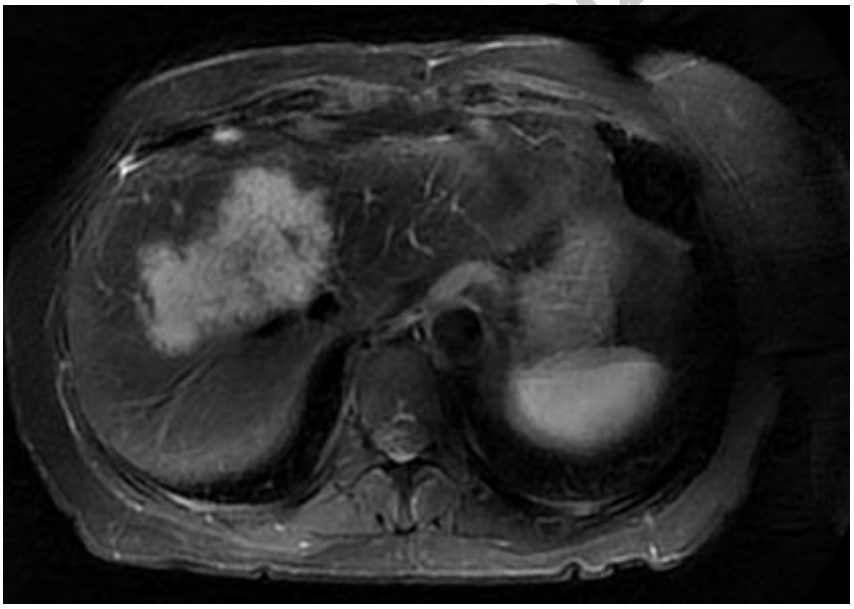


Table 1

Patient Demographics (n=55)	
Age (years)	55.1 ±10.6
Gender	
Male	45%
Female	55%
KRAS Status	
Positive	72%
Negative	28%
Microsatellite Instability (MSI)	
Stable	60%
High	40%
Colon Cancer site	
Right-sided	29%
Left-sided	67%
Transverse	4%
Extent of Metastatic Disease	
Hepatic Metastases Only	49.5%
Extrahepatic Metastatic Disease	50.5%
Chemotherapy Regimen	
FOLFOX	61%
FOLFIRI	29%
Other	10%

Table 2

Feature	Coefficient (B)	95% C.I.		P-Value
First Order – Minimum	-5.5	(-8.8 -	-2.2)	0.001
First Order - 10th Percentile	-0.8	(-1.4 -	-0.1)	0.020
First Order - 90th Percentile	-1.3	(-2.3 -	-0.3)	0.009
First Order – Maximum	-1.7	(-2.9 -	-0.4)	0.009
First Order – Mean	-1.1	(-1.9 -	-0.3)	0.008
First Order – Median	-1.1	(-1.9 -	-0.3)	0.010
First Order – RMS	-1.5	(-2.5 -	-0.5)	0.003
First Order – Uniformity	-6.1	(-10.5 -	-1.6)	0.007
GLCM- LD	-39.1	(-67.6 -	-10.5)	0.007
GLCM-LDM	-39.1	(-67.6 -	-10.5)	0.007
GLCM- LDMN	-98.3	(-169.9 -	-26.8)	0.007
GLCM- LDN	-58.8	(-101.7 -	-16.0)	0.007
GLCM- Joint Energy	-5.0	(-8.7 -	-1.4)	0.007
GLCM - Maximum Probability	-7.8	(-13.7 -	-1.8)	0.011
GLDM- Dependence Entropy	1.9	(0.3 -	3.5)	0.020
GLDM - Dependence Variance	2.1	(0.2 -	4.1)	0.032
GLRM - Normalized Non-Uniformity	-5.2	(-8.6 -	-1.9)	0.002
GLSZM - Normalized Non-Uniformity	-2.4	(-4.6 -	-0.1)	0.042
GLSZM - High Gray Level Zone Emphasis	1.3	(0.5 -	2.2)	0.003
GLSZM - Low Gray Level Zone Emphasis	-2.9	(-4.9 -	-1.0)	0.004
GLSZM-Normalized Size Zone Non-Uniformity	-1.2	(-2.2 -	-0.2)	0.024
NGTDM – Coarseness	-0.9	(-1.7 -	-0.1)	0.028