Longitudinal MRI-based radiomics: Estimating rectal cancer response to chemoradiation

Alice Chen

Introduction

Colorectal cancer is the most prominent gastrointestinal cancer in the Western world, and the fourth most common cancer diagnosed in the United States (Kapiteijn & van de Velde, 2002; Seigel, Miller, & Jemal, 2019). Due to the likelihood of local recurrence in rectal cancer, a combination of neoadjuvant chemoradiation (NAC) and total mesorectal excision (TME) is the standard treatment for locally advanced rectal cancer (Kapiteijn & van de Velde, 2002; Gosselink et al., 2005; Stewart & Dietz, 2007; Shaverdian et al., 2017). This technique has been employed since the 1980s when Heald et al. found it reduce recurrence rates to 2.7% percent in 1986. Since then, many studies have corroborated these low recurrence rates; however, TME is a highly invasive procedure, sometimes involving the removal of the entire rectum, with many possible surgical complications, and its adverse effects on quality of life such as a possible colostomy. Additionally, up to 27% of patients achieve a pathological complete response on chemoradiation alone, thus creating a demand for the estimation of patient response during and after treatment (Kapiteijn & van de Velde, 2002; Shaverdian et al., 2017). Doing so would provide physicians with the opportunity to identify the patients who were highly responsive to chemoradiation, and thus, may be spared from the potential adverse effects of TME (Kapiteijn & van de Velde, 2002; Shaverdian et al., 2017).

Apparent diffusion coefficient (ADC) maps may be calculated from DWI and provide information regarding the levels of diffusion and tumor cellularity in tissue. The mean ADC of a region of interest can be calculated from these ADC maps and is an established oncologic biomarker (Traverso et al., 2019; Blazic, Maksimovic, Gajic & Saranovic, 2015). While the slopes of ADC values over a course of NAC have been shown to correlate with tumor regression of rectal cancer in the past, other studies have produced conflicting results, and the correlation of

mean ADC and rectal cancer response to NAC remains unclear (Curvo-Semedo et al., 2011; Blazic, Maksimovic, Gajic & Saranovic, 2015; Shaverdian et al., 2017). Radiomics provides a potential solution; it is an emerging technology in which quantitative radiomics features are extracted from medical images in the process described in **Figure 1** (Lambin et al., 2017).

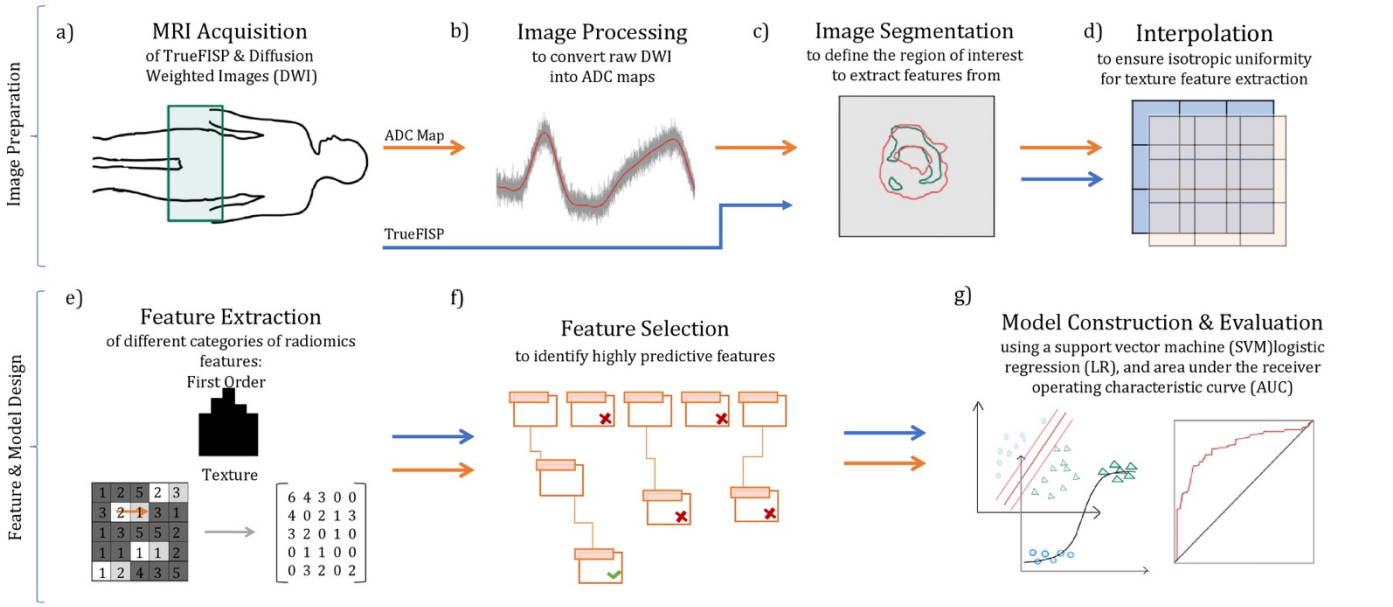

Figure 1. Radiomics workflow for TrueFISP and ADC maps. After a) imaging, b) image processing is necessary to convert DWI imaging into ADC maps. The region of interest (ROI) is then defined during c) image segmentation before d) image interpolation ensures isotropic uniformity for the extraction of texture features. e) Several categories of features may then be extracted, including first-order and texture features. f) Feature selection identifies important features for the g) construction of a model, which is typically evaluated using the area under the receiver operating characteristic curve (AUC) (Ibrahim et al., 2016; IBSI; van Griethyusen, 2017). Image created by student.

Radiomics features extracted from various types of MRIs have been correlated with rectal cancer response to NAC, including T2-weighted-imaging and ADC-map based features (Dinapoli et al., 2016; Traverso et al., 2019). However, ADC-map based radiomics features have been seen to be highly sensitive due to differences image pre-processing, and previous studies have failed to consider radiomics features extracted from true fast imaging with steady-state precession, or TrueFISP MRI: an anatomical MRI commonly used for rectal cancer NAC planning purposes (Duerk, Lewin, Wendt, & Petersilge, 1998; Ho, Liu, & Narra, 2008; Traverso

et al., 2019). Additionally, functional and anatomical MRI based radiomics features have yet to be compared while analyzing longitudinal changes in features during NAC as well. Therefore, this study sought to extract and compare features from longitudinal functional ADC-maps and anatomical TrueFISP MRI to identify 1) the superior MRI from which to extract highly estimative features, 2) the most estimative feature categories, and 3) the most advantageous times to image patients for the estimation of patient response to NAC, while using the mean ADC as a baseline for comparing estimative ability.

Methods

Patients and Image Registration

The longitudinal ADC maps and TrueFISP MRI of 10 rectal cancer patients who had received NAC with oral chemotherapy, capecitabine, at UCLA between April 2015 and November 2017 were analyzed. For each patient, rigid image registration was used to transform the gross tumor volume contour from treatment planning images onto ADC maps and TrueFISP MRI from five times throughout his/her NAC using commercial software (MIM® 6.6.5, MIM Software Inc.). These five times refer to a dose between those defined in **table 1**. Each patient's response to NAC was defined based on their Dworak grades, which were assigned post-surgery by a pathologist: those with grades of 0 and 1 were considered non-responders, while 2, 3, and 4 were considered responders.

Table 1. Times of image acquisition during NAC. Time x refers to the image being taken between the two doses of chemoradiation displayed.

Time	Dose	Time	Dose
1	1-10	4	19-29
2	9-15	5	25-39
3	14-21		

Mean ADC Analysis

The mean ADC values at times 1 and 5, as well as the delta, or difference in, mean ADC between the two time points were calculated for each patient. The area under the receiver operating characteristic curve (AUC) was used to compare the mean ADC between responders and non-responders and to evaluate the estimative ability of the mean ADC for estimating response to NAC (Curvo-Semedo et al., 2011; Blazic, Maksimovic, Gajic & Saranovic, 2015; Parmar et al., 2015).

Radiomics Feature Extraction

A total of 92 radiomics features were extracted from the ADC maps and TrueFISP MRI from each time using PyRadiomics ver. 2.2.0 (van Griethuysen et al., 2017). Of these features, there were 18 first order, 23 gray level co-occurrence matrix (GLCM), 16 gray level run length matrix (GLRLM), 16 gray level size zone matrix (GLSZM), 5 neighboring gray tone difference matrix (NGTDM), and 14 gray level dependence matrix (GLDM) features. A description of these features may be found in **Table 2.** Each feature from each time was considered a unique, single time (ST) feature, and delta features were calculated as the differences between all ten combinations of the five time points, where "5min1" represents features values from time point 5 minus values from time point 1. A total of 1,380 radiomics features were evaluated for each MRI type.

Table 2. Summary of first order and texture features. Gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighboring gray tone difference matrix (NGTDM), and gray level dependence matrix (GLDM) features are all texture features.

Feature Category	Description
First Order	Describe the distribution of voxel intensities using simple statistics.
Gray Level Co-Occurrence Matrix (GLCM)	Describe the probabilities of two gray level values occurring in adjacent voxels.
Gray Level Run Length Matrix (GLRLM)	Quantifies gray level runs, or lengths of consecutive pixels with the same gray level value.
Gray Level Size Zone Matrix (GLSZM)	Quantifies gray level zones, or groups of voxels with the same gray level values.
Neighboring Gray Tone Difference Matrix (NGTDM)	Quantifies the difference between a gray value and the average gray value of its neighbors within a certain number of voxels.
Gray Level Dependence Matrix (GLDM)	Quantifies gray level dependence, or the number of connected voxels within a certain number of voxels that are dependent on the center voxel.

Radiomics Feature Evaluation

The AUC and boxplots were first calculated and graphed to observe the degree of separation in feature values between responders and non-responders for the radiomics features with the top testing accuracies.

Two methods were used to evaluate the radiomics features. The first of which was a logistic regression (LR), where the radiomics features were evaluated in a "leave-one-out" (LOO) method. For each feature, a generalized LR to discern responders from non-responders was calculated based on nine patients' feature and response data in order to test the model's ability to estimate the tenth patient's response. Each feature was tested ten times, "leaving out" a different patient each time. The second method involved a support vector machine (SVM): each

feature was also evaluated in two, nested LOO methods: the first left the tenth patient out for the final check similarly to the LR method, while the inner LOO left out each of the nine patients in order to train the box constraint for the SVM.

The testing accuracy for each feature was calculated as the proportion of "left-out" patients whose response was correctly estimated by the LR or SVM, and the average training accuracy was calculated as the average proportion of the nine patients whose responses were accurately predicted by the regression line. A threshold of 0.8 was used on the testing and training accuracies of each of the radiomics features to identify the high accuracy features alone. The frequencies of each time during NAC, category of radiomics feature, and delta versus single time features were recorded. High-performing features for both LR and SVM evaluation were identified to identify the most predictive features. All calculations and evaluations were performed on MatlabR2015a.

Results and Discussion

The mean ADC values in responders and non-responders were compared using the AUC. The low AUCs and the lack of a clear separation in the mean ADCs of responders and non-responders, as seen in **Figure 2**, indicate the weak, estimative ability of the mean ADC values from the beginning and end of NAC.

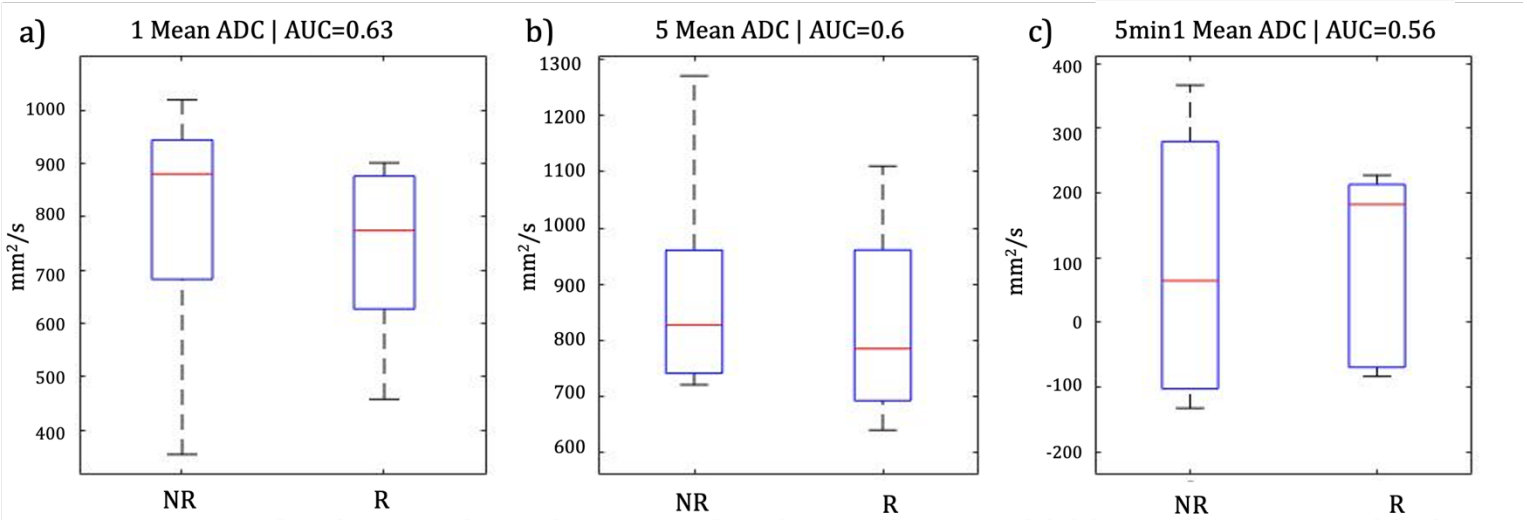

Figure 2. Mean ADC values for responders and non-responders. a) time 1, b) time 5, and c) delta 5 minus 1. Times 1 and 5 refer to images taken toward the beginning and end of the NAC course.

Similar AUCs of the mean ADC value have been observed in previous studies, with pre- and post-NAC mean ADC values having had AUCs of 0.51 and 0.55, respectively (Curvo-Semedo et al., 2011). However, while the results in **Figure 2** indicate the mean ADC from the beginning of NAC to hold the highest estimative abilities, previous studies have seen post-NAC and delta ADC values to hold higher AUCs of up to 0.828 when compared to the pre-NAC ADC (Blazic, Maksimovic, Gajic & Saranovic, 2015). These varying results may be due to varying definitions of responders and non-responders, as well as differences in ROI definition and placement (Curvo-Semedo et al., 2011).

The AUCs and boxplots of each feature were calculated and graphed to compare the degree of separation of features values. Most of the single time features had overlap in feature values when comparing responders and non-responders, however, one TrueFISP-based GLCM feature had a complete separation in responder and non-responder feature values, as observed in **Figure 3.**

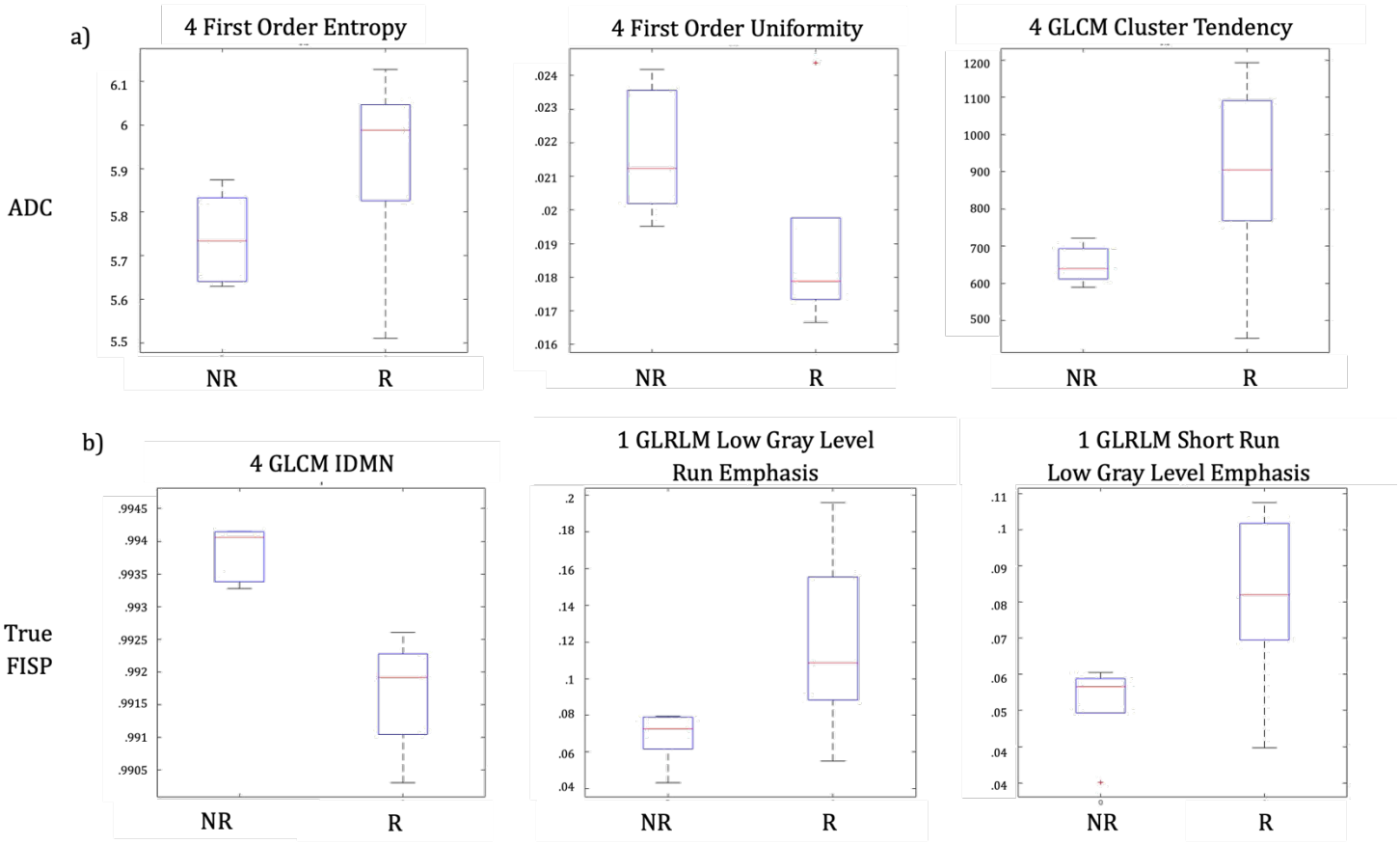

Figure 3. Top single time radiomics features for responders (R) and non-responders (NR). a) ADC map-based and b) TrueFISP based.

With the sole feature with complete separation being TrueFISP-based and a GLCM feature, these boxplots suggest TrueFISP MRI and GLCM features to be superior to the mean ADC, ADC map-based features, and other feature categories. However, single time features were predominantly observed to have significant overlap in responder and non-responder feature values, suggesting poor estimative abilities of features extracted from MRIs at individual times during NAC. Therefore, delta features, or differences in feature values between different times in NAC, were analyzed with the same method, yielding **Figure 4.**

The complete separation of TrueFISP-based delta features suggests that the changes in TrueFISP-based feature values between several times hold superior predictive abilities when compared to ST features and ADC-map-based features. Overall, delta features exhibited better separation in feature values in both ADC-map-based and TrueFISP-based features, suggesting the strength in observing changes in feature values during NAC rather than features at individual times. Once more, ADC map-based features suffered from an overlap in feature values when comparing responders and non-responders. However, the top few TrueFISP-based features all demonstrated a complete separation in responder and non-responder feature values. Additionally, these top features were also all GLCM features, supporting GLCM features' superior estimative abilities in comparison to other feature categories. The strength of GLCM features has been observed in previous studies as well, including GLCM's robustness to image noise (Gatta et al., 2019). The presence of first order and GLSZM features among the top testing accuracy features also suggest these feature types' strength in estimating rectal cancer response to NAC as well.

All radiomics feature were then evaluated using the LR and SVM LOO methods. However, due to the limited sample size and the considerable number of radiomics features, it was likely that some of the features achieved a high testing accuracy by chance. Thus, a threshold of 0.8 was used on the testing and training accuracies of all features. This yielded a remainder of 137 features, 14 ADC-based and 123 TrueFISP-based, for the LR method, and 82 features, 15 ADC-based and 67 TrueFISP-based features, for the SVM method. **Table 3.** depicts the AUCs and testing and training accuracies of six of the top ADC-map- and TrueFISP- based features.

Table 3. Testing and training accuracy and AUCs for top six radiomics features. Accuracies and AUCs are shown for the top ADC map-based features by a) LR and b) SVM evaluation and TrueFISP-based features by c) LR and d) SVM evaluation.

a)

Feature	GLCM Cluster Tendency	GLCM Cluster Shade	First Order Total Energy	First Order Skewness	First Order Energy	GLRLM Run Entropy
Time	4	4 min 3	4 min 2	5 min 3	4 min 2	4
Test	80%	80%	80%	80%	80%	80%
Train	90%	90%	87.8%	87.8%	86.7%	84.4%
AUC	0.8	0.88	0.84	0.84	0.88	0.8

b)

Feature	GLCM IDMN	GLCM Cluster Tendency	GLCM Sum Squares	GLCM Joint Average	First Order Mean Abs. Deviation	GLRLM Gray level Nonuniformity Norm.
Time	4	4 min 1	4 min 1	3 min 1	4 min 2	3 min 2
Test	100%	100%	100%	100%	100%	100%
Train	100%	100%	100%	100%	100%	100%
AUC	1	1	1	1	1	1

c)

Feature	GLCM IMC1	First Order Skewness	GLCM Sum Entropy	GLCM Cluster Tendency	GLRLM Run Entropy	First Order Energy
Time	2 min 1	5 min 3	4	4	4	4 min 2
Test	90%	90%	90%	90%	90%	80%
Train	90%	90%	90%	90%	90%	96.7%
AUC	0.84	0.84	0.8	0.8	0.8	0.88

d)

Feature	GLCM IDMN	First Order Mean Abs. Deviation	GLRLM Run Length Nonuniformity	GLRLM Low Gray Level Run Emph.	GLSZM Large Area Emph.	GLSZM Zone Variance
Time	4	4 min 2	3 min 2	3 min 2	4 min 1	3 min 2
Test / Train	100%	100%	100%	90%	90%	90%
Train	100%	100%	99%	99%	90%	90%
AUC	1	1	1	1	1	1

The LR and SVM evaluations revealed similar patterns: the highest accuracies of the ADC-map-based features were 0.8 and 0.9, respectively, while there were several TrueFISP-based features (8 for LR and 2 for SVM) with perfect AUCs and accuracies. These results further confirm the superiority of TrueFISP-based features when compared with ADC-map-based features. Notably, some of the features observed in **Figures 3 and 4** were also present in **Table 3**, including many first order and GLCM features. These GLCM features included the inverse difference moment normalized (IDMN), a measure of the local homogeneity of an image; cluster tendency, a measure of the groupings of voxels with similar gray-levels; and sum squares, a measure of the distribution of neighboring intensity level pairs (van Griethuysen, 2017). Additionally, several features performed highly under both LR and SVM evaluation: GLCM

IDMN and first order mean absolute deviation were two TrueFISP-based features that maintained perfect accuracies under both evaluations, and ADC-map-based GLRLM run entropy and GLCM cluster tendency performed well in both models as well. Lastly, the majority of the features observed in **Table 3.** Were delta features, further supporting the strength of observing changes in feature values throughout NAC instead of relying on features extracted from one time during NAC. These results are indicative of the predictive abilities of these select features, as well as the superiority of TrueFISP-based features, GLCM , GLRLM, and first-order features, and delta features. Also notable were the highly selected features, as listed and defined in **table 5.**

Table 5. The mostly highly selected radiomics features by both LR and SVM evaluation.

Radiomics Feature		**Feature Description**
GLRLM	Short Run High Gray Level Emphasis	Measures the joint distribution of short run lengths with high gray levels
GLSZM	Gray Level Variance	Measures the variance in gray level intensities for the gray level zones
First Order	Root Mean Squared	The square root of the mean of the gray level intensities squared
GLCM	Joint Average	The mean gray level intensity of the *i* distribution
GLCM	Cluster Tendency	Measures the groupings of voxels of similar gray levels

Among these features, GLRLM short run high gray level emphasis and GLSZM gray level variance were each selected 9 times among the total pool of high-accuracy features, First order root mean squared was selected 7 times, and GLCM joint average and cluster tendency were each selected 6 times. The average frequency of each of the original 92 radiomics features was roughly 2.3, indicating these features to hold superior classifying ability when compared to the remaining features.

To further characterize highly estimative radiomics features, the high accuracy TrueFISP features of both LR and SVM methods were analyzed by observing the frequencies and relative frequencies of times during NAC, feature category, and delta versus single time features. These frequencies can be seen in **Table 5, 6, and 7**.

Table 5. Frequency table of the feature categories among TrueFISP-based features with high (>0.8) training and testing accuracies by LR and SVM evaluation. Frequencies and relative frequencies of each feature category among the total 123 LR high accuracy TrueFISP features and 67 SVM high accuracy TrueFISP-features.

Feature Category	First Order	GLCM	GLRLM	GLSZM	GLDM	NGTDM
LR	31 / 25%	28 / 23%	24 / 20%	18 / 15%	17 / 14%	5 / 4%
SVM	18 / 27%	14 / 21%	18 / 27%	7 / 10%	4 / 6%	6 / 9%

Feature frequencies for GLCM, GLSZM, GLDM, and NGTDM all experienced a decrease when compared to its original proportion among the originally extracted features. However, first order and GLRLM features had significant increase in relative frequencies compared to its original proportion, which were 19% and 17%, respectively, before, suggesting the utility of these categories of features for predicting rectal cancer response to NAC. Additionally, while GLCM features did not experience in increase in relative frequency, they did have higher relative frequencies in comparison to the other feature categories, further suggesting their estimative abilities for rectal cancer NAC response estimation.

Table 6. Frequency table of the times among TrueFISP-based features with high (>0.8) training and testing accuracies by LR and SVM evaluation. Frequencies and relative frequencies of each time among the total 123 LR high accuracy TrueFISP features and 67 SVM high accuracy TrueFISP-features.

Time	1	2	3	4	5
LR	66 / 50%	47 / 39%	58 / 48%	41 / 34%	11 / 10%
SVM	51 / 76%	22 / 33%	31 / 46%	17 / 25%	7 / 10%

Table 7. Frequency table of delta and ST features among high accuracy (>0.8) TrueFISP-based features. Relative frequencies of each type among the a) 123 high accuracy features by LR evaluation and b) 67 high accuracy features by SVM evaluation.

Type	Delta	ST
LR	100 / 81%	23 / 19%
SVM	52 / 78%	15 / 22%

When comparing the relative frequencies of times during NAC, times 1 and 3 prevailed among the high accuracy TrueFISP features by both LR and SVM evaluation, suggesting times near the beginning and the middle of NAC to be optimal for MRI-acquisition for radiomics-based treatment response estimation. Additionally, the vast majority of high accuracy TrueFISP features by LR and SVM evaluation were delta features, once more supporting the importance of observing changes in feature values rather than features from individual times themselves for predicting rectal cancer response to NAC.

Conclusion

Radiomics is a novel technology involving the high-throughput extraction of endless quantitative features from medical images (Lambin et al., 2017). Many radiomics features have been shown in the past to correlate features extracted from several different imaging modalities with clinical endpoints in a variety of cancers, rectal cancer (Lambin et al., 2012; Bulens et al., 2019). This study was the first to compare the estimative powers of radiomics features derived from functional ADC maps and anatomical TrueFISP MRI.

First order and texture radiomics features from ADC maps and TrueFISP MRI of 10 rectal cancer patients were extracted and evaluated for patient response estimative ability.

Unfortunately, due to the image registration and the availability of only one, pre- or early-treatment tumor contour per patient, this study failed to evaluate shape radiomics features. Overall, TrueFISP-based features and first order and GLCM features were revealed to have superior training and testing accuracies, with several features achieving perfect training and testing accuracies. With the identification of these highly estimative features, radiologists can utilize not only their own qualitative observations of medical images, but also clinical-decision support systems (CDSS) that incorporate some of these quantitative features, further improving treatment individualization (Lambin et al., 2017). Additionally, times 1 and 3 were most prevalent among features with high training and testing accuracies. With a knowledge of the most advantageous times for image acquisition, physicians and patients may be spared from unnecessary imaging, which costs both time and resources. While MRI-guided radiation therapy (MRgRT) machines enable physicians to take an MRI while delivering a fraction of radiation therapy, a full MRgRT session can take 45 minutes which limits the number of patients that can be treated during the day. On the other hand, a normal external-beam radiation therapy session does not exceed 30 minutes and can take as little as 15 minutes (Botman et al., 2019).

Future studies ought to analyze of features extracted from medical imaging of rectal cancer patients in multi-center facilities. This study analyzed features from UCLA rectal cancer patients, the results may not be applicable to the general population of rectal cancer patients (Dinapoli et al., 2016). Additionally, future studies should not only use TrueFISP-based first order and texture features but also other radiomics features, MRI parameters, and additional clinical information and imaging times for the construction of a portable patient response estimation model for rectal cancer.

References

Blazic I., Maksimovic R., Gajic M. & Saranovic D. (2015). Apparent diffusion coefficient measurement covering complete tumor area better predicts rectal cancer response to neoaduvant chemoradiotherapy. *Croatian Medical Journal, 56*(5), 460-469. https://dx.doi.org/10.3325%2Fcmj.2015.56.460

Botman, R., Tetar, S. U., Palacios, M. A., Slotman, B. J., Lagerwaard, F. J., & Bruynzeel, A. M. E. (2019). The clinical introduction of MR-guided radiation therapy from a RTT perspective. *Clinical and Translational Radiation Oncology, 18,* 140-145. https://doi.org/10.1016/j.ctro.2019.04.019

Bulens, P., Couwenberg, A., Intven, M., Debucquoy, A., Vandecaveye, V., van Cutsem, E., . . . Haustermans, K. (2019). Predicting the tumor response to chemoradiotherapy for rectal cancer: Model development and external validation using MRI radiomics. *Radiotherapy and Oncology*. https://doi.org/10.1016/j.radonc.2019.07.033

Curvo-Semedo, L., Lambregts, D. M. J., Maas, M., Thywissen, T., Mehsen, R. T. … Beets-Tan, R.G.H. (2011). Rectal cancer: Assessment of complete response to preoperative combined radiation therapy with chemotherapy—Conventional MR volumetry versus diffusion-weighted MR imaging. *Radiology, 260*(3), 734-743. https://doi.org/10.1148/radiol.11102467

De Cecco, C., Ganeshan, B., Ciolina, M., Rengo, M., Meinel, F., Musio, D., . . . Laghi, A. (2015). Texture analysis as imaging biomarker of tumoral response to neoadjuvant chemoradiotherapy in rectal cancer patients studied with 3T magnetic resonance. *Investigative Radiology*, *50*(4), 239-245. https://doi.org/10.1097/RLI.0000000000000116

Dinapoli, N., Casà, C., Barbaro, B., Chiloiro, G. V., Damiani, A., Matteo, M. D., . . . Valentini, V. (2016). Radiomics for rectal cancer. *Translational Cancer Research*, *5*(4). https://doi.org/10.21037/tcr.2016.06.08

Duerk, J. L., Lewin, J. S., Wendt, M., & Petersilge, C. (1998). Remember true FISP? a high SNR, near 1-second imaging method for T2-like contrast in interventional MRI at .2 T. *Journal of Magnetic Resonance Imaging*, *8*(1), 203-208. https://doi.org/10.1002/jmri.1880080134

Gatta, R., Vallati, M., Dinapoli, N., Masciocchi, C., Lenkowicz, J…. Valentini, V. (2019). Towards a modular decision support system for radiomics: A case study on rectal cancer. *Artificial Intelligence in Medicine, 96*, 145-153. https://doi.org/10.1016/j.aratmed.2018.09.003

Gosselink M.P., Busschbach, J.J., Dijkhuis, C.M., Stassen, L.P., Hop, W.C. & Schouten W.R. (2005). Quality of life after total mesorectal excision for rectal cancer. *Colorectal Disease, 8*(1). https://doi.org/10.1111/j.1463-1318.2005.00836.x

Ho, M.-L., Liu, J., & Narra, V. (2008). Magnetic resonance imaging of rectal cancer. *Clinics in Colon and Rectal Surgery*, *21*(3), 178-187. https://doi.org/10.1055/s-2008-1080997

Kapiteijn, E., & van de Velde, C. J. H. (2002). The role of total mesorectal excision in the management of rectal cancer. *Surgical Clinics of North America*, *82*(5), 995-1007. https://doi.org/10.1016/S0039-6109(02)00040-3

Lambin, P., Leijenaar, R. T.H., Deist, T. M., Peerlings, J., de Jong, E. E.C., van Timmeren, J., . . . Walsh, S. (2017). Radiomics: The bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, *14*, 749-762. https://doi.org/10.1038/nrclinonc.2017.141

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., . . . Aerts, H. J. (2012). Radiomics: Extracting more information from medical images using advanced

feature analysis. *European Journal of Cancer*, *48*(4), 441-446. https://doi.org/10.1016/j.ejca.2011.11.036

Lee, S. L., Ravi, A., Morton, G., Loblaw, A., Tseng, C.-L., Haider, M., . . . Chung, H. T. (2019). Changes in ADC and T2-weighted MRI-derived radiomic features in patients treated with focal salvage HDR prostate brachytherapy for local recurrence after previous external-beam radiotherapy. *Brachytherapy*, *18*(5), 567-573. https://doi.org/10.1016/j.brachy.2019.04.006

MATLAB. (2015). *Version 8.5.0 (R2015a)*. Natick, Massachusetts: The MathWorks Inc.

Ng, F., Ganeshan, B., Kozarski, R., Miles, K. A., & Goh, V. (2013). Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: Contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology*, *266*(1), 177-184. https://doi.org/10.1148/radiol.12120254

Shaverdian, N., Yang, Y., Hu, P., Hart, S., Sheng, K., Lamb, J., . . . Lee, P. (2017). Feasibility evaluation of diffusion-weighted imaging using an integrated MRI-radiotherapy system for response assessment to neoadjuvant therapy in rectal cancer. *The British Journal of Radiology*, *90*(1071). https://doi.org/10.1259/bjr.20160739

Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics, 2019. *Cancer Journal for Clinicians*, *69*(1), 7-34. https://doi.org/10.3322/caac.21551

Traverso, A., Kzmierski, M., Shi, Z., Kalendralis, P., Welch, M.…Wee, L. (2019). Stability of radiomic features of apparent diffusion coefficient (ADC) maps for locally advanced rectal cancer in response to image pre-processing. *Physica Medica, 61*, 44-51. https://doi.org/10.1016/j.ejmp.2019.04.009

van Griethuysen, J. J.M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., . . . Aerts, H. J. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, *77*(21), e104-e107. https://doi.org/10.1158/0008-5472.CAN-17-0339