Validation of Optimal DCE-MRI Perfusion Threshold to Classify At-risk Tumor Imaging Voxels in Heterogeneous Cervical Cancer for Outcome Prediction

Zhibin Huang[1,2], **Kevin A. Yuh**[1,2], **Simon S. Lo**[3], **John C. Grecula**[2], **Steffen Sammet**[4,5], **Christina L. Sammet**[6], **Guang Jia**[10], **Michael V. Knopp**[4], **Qiang Wu**[7], **Norman J. Beauchamp III**[8], **William TC Yuh**[4,8], **Roy Wang**[2], and **Nina A. Mayr**[2,9]

[1]Department of Radiation Oncology, East Carolina University, Greenville, NC

[2]Department of Radiation Oncology, Ohio State University, Columbus, OH

[3]Department of Radiation Oncology, Case Western Reserve University, Cleveland, OH

[4]Department of Radiology, Ohio State University, Columbus, OH

[5]Department of Radiology, University of Chicago, Chicago, Il

[6]Department of Medical Imaging, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Il

[7]Department of Biostatistics, East Carolina University, Greenville, NC

[8]Department of Radiology, University of Washington, Seattle, WA

[9]Department of Radiation Oncology, University of Washington, Seattle, WA

[10]Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA

Abstract

Purpose—To classify tumor imaging voxels at-risk for treatment failure within the heterogeneous cervical cancer using DCE MRI and determine optimal voxel's DCE threshold values at different treatment time points for early prediction of treatment failure.

Material and Method—DCE-MRI from 102 patients with Stage IB2–IVB cervical cancer was obtained at 3 different treatment time points: before (MRI 1) and during treatment (MRI 2 at 2–2.5 weeks and MRI 3 at 4–5 weeks). For each tumor voxel, the plateau signal intensity (SI) was derived from its time-SI curve from the DCE MRI. The optimal SI thresholds to classify the at-

---

Corresponding author's contact information: To whom correspondence should be addressed: Nina A. Mayr MD, Department of Radiation Oncology, UW Medicine, University of Washington, 1959 NE Pacific Street, Seattle, WA 98195-6043 (ninamayr@uw.edu).

risk tumor voxels was determined by the maximal area under the curve using ROC analysis when varies SI value from 1.0 to 3.0 and correlates with treatment outcome.

Results—The optimal SI thresholds for MRI 1, 2 and 3 were 2.2, 2.2 and 2.1 for significant differentiation between local recurrence/control, respectively, and 1.8, 2.1 and 2.2 for death/survival, respectively.

Conclusion—Optimal SI thresholds are clinically validated to quantify at-risk tumor voxels which vary with time. A single universal threshold (SI = 1.9) was identified for all 3 treatment time points and remain significant for the early prediction of treatment failure.

Index term

Tumor heterogeneity imaging; at-risk tumor voxels; low DCE perfusion threshold; poor perfusion regions in heterogeneous tumor; early outcome prediction; cervical cancer

Introduction

Tumor biological heterogeneity is a well-known phenomenon, which critically influences treatment outcome (1, 2). Heterogeneous therapy responsiveness occurs not only among different patients, even with the same clinical stages, but also at different sub-regions within the same tumor (3). The ultimate treatment failure is likely more adversely influenced by those at-risk sub-regions (tumor voxels) within the heterogeneous tumor, which possess unfavorable biological properties. Evaluation of tumor heterogeneity can be challenging. Tissue biopsy has been used to evaluate tumor heterogeneity, but this method is invasive, prone to sampling error and cannot investigate the entire heterogeneous tumor volume. Furthermore, tumor biological properties may change within short periods of time during cytotoxic treatment. Therefore, an innovative way to non-invasively and readily assess heterogeneity throughout the entire tumor bed at different treatment time points is critically needed. The ability to quantify the at-risk tumor component, contributing to the ultimate therapy failure would profoundly impact cancer therapy strategy, particularly during early treatment.

Dynamic contrast-enhanced (DCE) MRI has been reported to be valuable for the evaluation of tumor microvascular structure and function (4–8). DCE MRI indirectly reflects tumor perfusion status and the effectiveness in the delivery of chemotherapeutic drugs (9–11) and oxygen (hypoxia) (12) to the tumor cells. Therefore the DCE-based perfusion biomarkers may provide important information related to the early responsiveness or resistance to an ongoing treatment, which critically influence the ultimate treatment outcomes: local recurrence (LR) vs. control (LC) and death of disease (DOD) vs. disease-specific- survival (DSS) (13–16). Poor perfusion status of the entire heterogeneous tumor, as judged by the low mean value of the plateau signal intensity (SI) of the DCE curve for the whole tumor, has been reported to correlate significantly with poor treatment outcome in cervical cancer (16–19). Although the mean SI value provides useful information related to the functional/ biological properties of the entire tumor, it does not account for tumor heterogeneity and does not specifically target the high-risk tumor sub-regions that are likely contributing to treatment failure.

Clinical MRI has been utilized in the management of cervical cancer at variable treatment time points and but DCE-MRI has not been routinely used for assessing functional tumor heterogeneity for personalized cancer treatment strategy. With the improved high spatial and temporal resolution, tumor anatomical and functional imaging studies including DCE-MRI are now available in the clinical setting to assess the tumor biological microenvironment throughout the entire heterogeneous tumor with a precision at the level of each tumor imaging voxel (20–22). Therefore, clinical DCE-MRI provides new opportunities to evaluate each tumor voxel's perfusion status within the whole tumor and to characterize tumor heterogeneity metrics by the heterogeneous distribution of the tumor voxels' SI values throughout the entire tumor mass. However, functional tumor imaging in the clinical setting has not routinely targeted at-risk components within the heterogeneous tumor because there are no established DCE imaging metrics to reliably classify the at-risk tumor voxels that are likely resistant to cytotoxic therapy and contribute to the unfavorable treatment outcome. Therefore a clinically validated tumor voxels' SI threshold to define at-risk tumor voxels would have tremendous implication for clinical translation for personalized risk assessment particularly for early prediction of treatment failure when the adaptive therapy is most effective.

In cervical cancer, the well-established relationship among tumor vascularity, hypoxia and radiation therapy response provides a unique tumor-biology based hypoxia model to apply serial DCE perfusion imaging in clinical patients, explore DCE functional biomarkers and classify the at-risk components within the heterogeneous tumor. Since advanced cervical cancer is typically treated with radiation/chemotherapy without surgical resection, it provides a unique opportunity for early sequential imaging of the entire tumor mass, and for assessment of the tumor perfusion status before and during treatment (23). Furthermore, it is an ideal tumor model with readily confirmable clinical endpoints which is critical for clinical validation research. In contrast to other tumors, cervical cancer requires a relatively short patient follow-up time to confirm the ultimate treatment outcomes: local tumor control and survival. Most treatment failures occur within two years, and essentially all within five years after treatment completion.

We hypothesize that, within the entire heterogeneous tumor mass, tumor voxels with critically low DCE perfusion values below a certain SI cut-off threshold value significantly contribute to treatment failure, more than those with perfusion values above such threshold. The at-risk component of the tumor derived from the DCE MRI provides additional essential information which can be a useful adjunctive to the current established cancer treatment paradigm for personalized early response monitoring, risk assessment and early prediction for the ultimate treatment failure.

The purpose of the current study is to 1) quantify the tumor heterogeneity with DCE-MRI; 2) determine the optimal SI thresholds for best classification of the at-risk tumor voxels, representing tumor components with unfavorable biological properties at three different treatment time points using Receiver Operating Characteristics (ROC) Analyses; 3) validate clinically the efficacy of these high-risk tumor components for the differentiation between favorable and unfavorable outcomes; and 4) investigate if a single SI threshold value can be

applied universally at various treatment time points to differentiate ultimate treatment outcomes.

Materials and Methods

Patient population, clinical information and treatment end points

Three serial MRI were performed before and during treatment in 102 patients with advanced cervical cancer. Patient's pertinent clinical information is summarized in Table 1. Pre-treatment clinical evaluations included routine initial clinical work-up following FIGO guidelines (24). Patients were treated with standard external beam RT with 6 to 24 MeV photons beams, concurrent weekly cisplatin-based chemotherapy and standard-of-care brachytherapy. The dose of external beam RT ranged from 39.6 to 66.6 (mean, 47.8) Gy in conventional 1.8–2.0 Gy fractions. External beam radiation included the pelvis radiation and field reductions. All except four patients received brachytherapy, with Point A doses ranging from 14 to 61 (mean 39.5) Gy. Patients were followed up routinely with 3–6 monthly intervals to assess treatment outcomes: local tumor control (LC) and disease-specific survival (DSS) outcomes. LC was defined as no tumor recurrence or persistence within pelvis during the follow-up period. For DSS analysis, death due to causes other than cervical cancer was considered a censored observation. Mean follow-up was 6.8 years (range: 0.2 to 9.4 years).

MRI acquisition

Serial MRI examinations were performed prospectively at three time points: before RT, early during RT (at 20–25 Gy of pelvic RT), and mid-way during RT (at 40–50 Gy) with 1.5 Tesla scanners. The imaging protocols were detailed in a prior published study (14).

Each MR study consists of pre-contrast T1- and T2 weighted, post contrast and DCE images. Precontrast imaging included sagittal T2-weighted images (TR: 4000–5600 ms, TE: 100–120 ms, FOV: 30–40 cm, matrix: 256×192–256, slice thickness:4 mm, gap: 1 mm, NEX: 2), axial T2-weighted images (TR:2500–3200 ms, TE: 100–120 ms, FOV: 30–40, matrix: 256×192–256, slice thickness: 7 mm, gap:1 mm, NEX: 1–2), and axial T1-weighted images (TR: 350–700, TE: 5–17, FOV: 30–40, matrix: 256×192–256, slice thickness: 7 mm, slice gap: 1 mm, NEX:2–4) for optimal tumor localization and delineation (Figs. 2 and 3). The postcontrast MRI used the same pulse sequence for the precontrast T1-weighted SE image.

For the DCE MRI studies, an intravenous bolus of 1 mmol/kg commercially available Gadolinium chelated MR contrast agent was initiated after completion of 3–5 baseline acquisitions at an injection rate equal to or greater than 5 mL/s. Two DCE imaging protocols were applied. With the initial limitation of the temporal resolution of 3 second, only T1-weighted FSE image (TR: 150 ms, TE: 18 ms, FOV: 30–40 cm, matrix size: 128×192) was available during early phase of our study using single slice through the epicenter of the tumor (1 cm), where most of the hypoxic tumor cell located. Later a 3-D gradient echo pulse sequence with the improved spatial and temporal resolution was available for the DCE MRI volume studies (TR: 12.0 – 13.7 ms, TE: 4.2 – 5.0 ms, NEX= 1, angle = 30°, FOV = 30 – 40

cm, matrix size: $160 - 192 \times 256$, partitions: 12, slab: 9.6 cm, slice thickness: 8 mm, and temporal resolution: 26 to 28 seconds.

Image analysis

The region-of-interest (ROI) of the tumor area was delineated on each T2-weighted imaging slice by three independent reviewers. The tumor 3D anatomical volume was derived by summation of all the voxels within the tumor ROI slices for each imaging time point as described previously (25, 26). Tumor ROIs derived from the T2-weighted images were mapped into the DCE data file. Using the first-pass DCE method for DCE imaging (27), a time-signal intensity DCE curve was generated for each tumor voxel and the plateau signal intensity (SI) was derived using the relative plateau SI to the average precontrast SI prior to the bolus arrival (14). Tumor Voxel SI histograms, first developed in our lab (Figs 2 and 3), were derived from each of the three DCE MRI studies at different time points for each individual tumor based upon total tumor voxels and their SI distribution as described in detail earlier (14). Based upon the tumor voxels' SI distribution on the Voxel SI histogram, tumor heterogeneity can be readily displayed and assessed by the range of variation of SI values within the entire tumor voxel population.

Statistical analysis

All statistical analyses were performed with SPSS (SPSS 19, IBM SPSS Inc. Chicago, IL). The at-risk tumor voxels can be classified as those voxels with SI lower than a given SI threshold and readily identified on the Voxel-Histogram. The total number of low DCE at-risk voxels within a heterogeneous tumor is influenced by the choice of a specific SI threshold value to classify the at-risk tumor voxels populations which is directly proportional to the value of the SI threshold (Fig 1). Consequently, the higher the applied SI threshold value, the more tumor voxels will be classified as at-risk tumor voxels. Similarly, when a lower SI threshold is chosen to define at-risk tumor voxels, a smaller portion (lower number) of the total tumor voxels will be considered at-risk.

A. Determination of optimal SI threshold at different treatment time points by ROC analysis—For MRI obtained at a specific treatment time point, the optimal SI threshold value to classify the at-risk tumor voxels and differentiate between favorable (LC or DSS) and unfavorable (LR or DOD) treatment outcome was determined using Receiver Operating Characteristics (ROC) analyses and Mann-Whitney rank sum tests. Because only single index of SI is applied, we used the maximum AUC to determine the optimal value of SI for the best performance for a MR study at a given time. For a given SI threshold value (such as SI= 2.1 depicted in figure 1), the number of at-risk voxels below a threshold (figs 2 and 3) is defined for all the 102 patients to correlate the patients' outcomes (102 pairs of variables). ROC curve for a specific SI threshold is derived as a plot of sensitivity as a function of specificity, which provides the best classification for differentiate at-risk tumor voxels and ultimate clinical outcomes with a balance between the specificity and the sensitivity. The area under curve (AUC) of ROC curves based upon a selected SI is defined under a balanced classification of sensitivity and specificity. The maximum AUC of ROC curves corresponding to the optimal SI threshold value could be used to separate patients into two groups, yielding the best correlation between the variable (risk voxel) and clinical

outcome. For the current research, 21 different SI threshold values (from SI=1.0 to SI =3.0 in 0.1 increments) were applied to each of the 102 tumors to classify their respective at-risk tumor voxels populations (example in Fig. 1 using SI threshold of 2.1 to classify at-risk voxels with SI < 2.1). This resulted in 21 different populations of at-risk voxels for each of the 102 tumors. For all the 102 tumors, each of their 21 populations of at-risk tumor voxels was correlated with the 102 patients' respective clinical outcomes (LC and DSS) for statistical analyses. For example, when a specific SI threshold value (from the 21 possible options) is applied to the 102 tumors, 102 pairs of data points can be generated, which consist of 102 populations of at-risk tumor voxels and their corresponding binary outcome (yes or no for either LC or DSS). This data set is used to generate an ROC curve that is pertinent to this specific SI threshold value. Therefore 21 different ROC curves are generated using 21 different SI values from the DCE MRI obtained at a given treatment time point (Figure 4, obtained at 2 weeks into treatment time point). For each of these 21 ROC curves derived from the DCE MRI at a given time point, the respective AUC and the corresponding zero derivative (AUC_{dev}) can be determined through the ROC analysis. These 21 AUC and AUC_{dev} were then compared to determine the maximal AUC or the zero point of the AUC_{dev}. The optimal SI threshold value is derived from interpolating the SI value that corresponds to the one that is most close to the maximal AUC or the zero point of the AUC_{dev}.

B. Determination of a single universal SI threshold that can be readily applied for all the treatment time points to effectively classify at-risk voxels and remain significant for early outcome prediction—We further evaluated a possible single SI values within the range of the entire optimal SI threshold values for either LC or DSS at all 3 time points to establish a universal SI threshold that can be efficaciously applied for all the 3 time points. Such single universal SI threshold may not be as powerful as those optimal thresholds specific for different treatment time points, but it remains comparatively effective for all 3 time points for significant differentiation between favorable and unfavorable outcomes (LC or DSS). For this purpose, a multiple logistic regression of the clinical outcome and the numbers of at-risk tumor voxels at all 3 time points was performed to generate new sets of variables that weighted all 3 studies (28). ROC curves were generated based upon these regression results using different SI values within the range of the entire optimal SI threshold values (1.8–2.1). The AUC and AUC_{dev} of these ROC curves for LC and DSS corresponding to range of the SI optimal thresholds (1.8, 1.9. 2.0 and 2.1) were compared to determine the best single SI threshold value that has the maximal AUC value and remains effective for all 3 treatment time points.

Results

The component of at-risk voxels was derived for each of the 102 cervical cancer patients by applying 21 different SI values ranging from 1.0 to 3.0 with 0.1 increments. Figure 1 shows an example histogram of the 102 patients and their total number of at-risk DCE voxels defined by the specific SI threshold value = 2.1 from the DCE-MRI obtained at 2–2.5 weeks into treatment (MRI 2). The histogram demonstrates the number of low DCE at-risk voxels

varied greatly from patient to patient (heterogeneous at-risk components) and among their different treatment outcomes (not shown).

For each MRI at the 3 different time points (before, and at 2–3 or 4–5 weeks into treatment), 21 ROC analyses were derived. The ROC analysis consisted of the total number of low DCE voxels classified by the 21 SI thresholds for the 102 patients and their respective treatment outcome, as shown in Figure 4. The AUC and AUC_{dev} for each of the 21 ROC curves as a function of SI was calculated. The optimal threshold SI values were determined by the maximal AUC and the zero point of the AUC_{dev} from the ROC analysis, before (MRI 1), and at 2–3 weeks (MRI 2) or 4–5 weeks (MRI 3) during treatment, are summarized at Table 2. The optimal SI values are different for LC or DSS and are ranged from 1.8 to 2.1. These optimal SI threshold values varied at different treatment time points, but most significantly differentiate either local control vs. local recurrence (LC) or disease specific death (DSS) vs. survival. To derive a single universal threshold applicable for all 3 treatment time points, 4 possible SI values (1.8, 1.9, 2.0 and 2.1) within the range of optimal SI threshold values for both LC and DSS were considered. New variables with coefficients or weighting factors related to all 3 MRI studies at different time points were generated from logistic regression analysis (28). The respective AUC of these new variables (N=8) was generated using ROC analysis (Table 3). The maximum AUC of these new variables were identified (0.742 for LC and 0.719 in Table 3) and used to determine the universal threshold, which corresponded to SI = 1.9. Therefore, the resultant at-risk voxels defined by SI = 1.9 remained the most effective threshold value for all 3 different time points to significantly differentiate favorable and unfavorable outcome for both LC and DSS (Mann-Whitney U test). The ROC curve classified by SI = 1.9 (Fig. 2) also shows an overall excellent performance compared to most of the other ROC curves at all 3 treatment time points.

Discussion

The daily predicament in the management of cancer patients is to determine, as early as possible, if a specific ongoing cancer therapy for a specific individual patient is indeed effective, because chances for cure for this particular patient becomes much less likely by the time the treatment failure is detected many months or years later. One of the fundamental challenges to provide such personalized cancer care strategy is the lack of non-invasive ways to characterize and quantify each individual tumor's heterogeneity and its heterogeneous response among different patients at different areas within the same tumor and at various treatment time points. The microenvironments among sub-regions within the same tumor have shown heterogeneous biological properties (29) relating to hypoxia (30), tumor responsiveness to treatment (31), and gene expression (32). However, the established clinical prognostic criteria and treatment response assessment for advanced cervical cancer have been based upon morphology-based FIGO staging (24) and tumor volume measured at different treatment times using WHO or RECIST criteria (33). These one- or two-dimensional morphological measurements assume all tumors having ellipsoid shape, and do not account for the irregular tumor border and non-linear shrinkage of the tumor induced by treatment. Functional imaging biomarkers reflecting the underlying pathophysiology, such as tumor perfusion status and heterogeneity, have not been incorporated into the current established clinical prognosticators or strategy (34). Using the current standards, the

treatment-related specific prognosis and ultimate outcomes for individual patients vary widely within the same established stage/prognostic categories. Based upon the tumor morphology information alone, the diagnosis of treatment failure frequently comes too late for meaningful salvage treatment options (35, 36).

For clinical translation to characterize the imaging-based tumor heterogeneity and defined the at-risk tumor voxels, three prerequisites are essential, which are readily available in clinical MR scanners and settings. First, a high spatial resolution 3D morphological image is needed to optimize the localization of the tumor mass, delineate the full extent of tumor involvement, and provide reasonably small tumor voxel size (with adequate signal-to-noise ratio and image quality). Second, a sufficiently high temporal resolution DCE-MRI is needed to achieve high quality DCE perfusion maps that provide adequate functional information for each tumor voxel in order to derive meaningful SI-time curves and the plateau SI. Third, a robust imaging analysis methodology is needed to characterize tumor heterogeneity and display the SI distribution of all the tumor voxel populations for the identification of at-risk tumor voxels with various SI threshold values. Both high spatial resolution 3D anatomic and high temporal resolution functional imaging methodologies were integrated in the current research, in order to more precisely target the at-risk tumor voxels throughout the irregularly shaped tumor and effectively assess the tumor underlying biology/pathophysiology.

The current research applied the DCE perfusion MRI and ROC analysis for clinical validation of the novel optimal SI threshold concept at different time points for the effective classification of at-risk tumor voxels. It was based upon the well-known underlying pathophysiology (perfusion, hypoxia) critically influencing individual tumor's response, and an ideal tumor model for ultimate outcome correlation. Our results support the hypothesis that the quality and quantity of the low-DCE sub-volumes within the heterogeneous tumor volume are significantly correlated with unfavorable treatment outcome. Our patient population, which was the largest patient number with the longest follow-up time (mean: 6.8 years) reported to date, provided a solid database for the clinical endpoint verification because most treatment failures in cervical cancer occur within two years.

One important observation is that the temporal changes of MR perfusion biomarkers including SI can occur in as early as 2 weeks into treatment, which enable the assessment of early response and ultimate outcome prediction. Such information may not be readily appreciated by the temporal changes of tumor volume particularly during early treatment (37). With the clinically validated optimal SI threshold values, at various treatments time points as determined in this research, the evaluation of the early temporal changes related to the high-risk tumor voxels become more realistic and precise for early prediction of treatment failure related to an ongoing treatment. A potential therapeutic window can be provided, particularly early during treatment, to adjust the personalized cancer treatment strategy and profoundly impact the outcomes.

To classify the at-risk tumor voxels within the heterogeneous tumor, previously Yuh et al applied different lower percentile thresholds of the total tumor voxel population (range 2.5 – 20 percentile with 2.5 increments) and correlated this to clinical outcomes (14). The lower

percentile threshold assumes at-risk voxels within a specific tumor having relative poor perfusion compared to the entire tumor voxel population. The authors reported significant correlation with clinical outcome using the lower percentage threshold to classify the at-risk tumor voxels for outcome prediction. They found 5 percentile was the optimal percentile threshold which usually corresponding to approximately SI = 2.05 and the best timing was the 2nd MRI at 2–2.5 weeks into treatment. However, the percentile threshold is a relative parameter pertaining to the total tumor voxel population from a specific tumor and has its limitation. It assumes all tumors must have unfavorable or favorable components (voxels) including those tumors with excellent or very poor perfusion (overestimates or underestimates the at-risk voxels, respectively). Mayr et al. used a fixed absolute SI threshold of 2.1 to classify the functional risk volume for all 3 MRI studies obtained before treatment and at 2–2.5 or 4–4.5 weeks into treatment (38). They found SI = 2.1 was an effective SI threshold for all 3 treatment time points to classify at-risk tumor voxels and predict treatment outcome. However, optimal SI thresholds at different treatment time points were not investigated.

A p-value for each SI threshold can be derived by correlating the number of at-risk tumor voxels defined by such SI threshold with the ultimate outcome. However, the p-values do not provide the comparative effectiveness of among different respective SI values so that the most significant independent SI threshold value, or the optimal SI threshold, could be determined. Therefore ROC analysis is applied to select the optimal SI threshold judged by the maximal area under curve (AUC) of ROC curves for this purpose. Selecting the optimal SI threshold is a pure numerical optimization process, and there is no need to generate p-values for this purpose. In the current study, the optimal absolute SI thresholds at 3 different treatment time points were clinically validated for DCE heterogeneity-based risk assessment. These optimal SI threshold values make the evaluation of the temporal changes of the high-risk tumor components at different treatment time points more realistic in the clinical settings. The identification of the absolute SI threshold of SI = 1.9 as a single effective universal SI threshold for all 3 treatment time points further facilitate the clinical translation of the concept of tumor heterogeneity to quantify at-risk tumor voxels, not just morphologic volume, and thus further improve the prediction of treatment outcomes.

There are limitations of our research at the time the study was conducted and our results have to be reconfirmed with future studies. Despite having the largest patient number and most frequent imaged patient population before and during early treatment, our statistical power was limited by the number of patients studied. In addition, our studies were incremental and performed over a long period of time when there were rapid advancements in MR imaging techniques. The lack of uniformity in imaging technique for all the MR examinations could further limit our statistical analysis, results and conclusions. The variation of the SI was not assessed in the current study and would be valuable for future validation with more uniform imaging acquisition and larger patient population. Furthermore, our studies did not incorporate the higher MR contrast dose, and the most updated imaging techniques with the highest spatial and temporal resolution available today. These factors will likely influence the values of the optimal and universal threshold determined by the current research. Despite these limitations, our preliminary results were able to show encouraging findings that tumor DCE heterogeneity could be readily assessed

in the clinical settings and at-risk tumor voxels could be effectively classified and clinically validated by the ultimate treatment outcomes for early outcome prediction based upon the limited clinical MR studies obtained over the past.

In conclusion, tumor heterogeneity, known to influence treatment outcome, has not been incorporated in current standard cancer management paradigm, including cervical cancer. With the improvement in the temporal and spatial resolutions of clinical scanners, DCE MR provides a unique capability, early during treatment, to characterize perfusion heterogeneity in cervical cancer and classify at-risk low-perfused tumor voxels that likely contribute to treatment failure for risk assessment and early outcome prediction. With the known classic hypoxia/perfusion model and short clinical follow-up time required to verify the long-term clinical outcome in cervical cancer, we are able to apply DCE-MRI and ROC analysis to clinically validate the optimal voxel SI threshold values to classify at-risk voxels which can significantly differentiate treatment outcome before and early during treatment. To facilitate clinical translation, a single universal SI voxel threshold value for all three treatment time points was also identified and clinically validated to further facilitate the clinical translation of the heterogeneity concept in order to effectively classify at-risk tumor components and significantly predict the ultimate treatment outcome. With the ability to analyze the at-risk tumor voxels within the heterogeneous tumor and identify the at-risk portion of the tumor, we were able to predict treatment outcome early during treatment (as early as 2 weeks into treatment), which would be a great contribution to the current management of cancer patients. If our results can be further confirmed with future studies with more consistent imaging technique and larger patient population, this would have profound impact to the management of cancer patients.

Acknowledgments

REFERENCES

1. Dewhirst MW, Oliver R, Tso CY, Gustafson C, Secomb T, Gross JF. Heterogeneity in tumor microvascular response to radiation. Int J Radiat Oncol Biol Phys. 1990; 18:559–568. [PubMed: 2318688]

2. Lyng H, Vorren AO, Sundfor K, et al. Intra- and intertumor heterogeneity in blood perfusion of human cervical cancer before treatment and after radiotherapy. Int J Cancer. 2001; 96:182–190. [PubMed: 11410887]

3. Eifel PJ. Problems with the Clinical Staging of Carcinoma of the Cervix. Semin Radiat Oncol. 1994; 4:1–8. [PubMed: 10717083]

4. Brasch RC, Li KC, Husband JE, et al. In vivo monitoring of tumor angiogenesis with MR imaging. Acad Radiol. 2000; 7:812–823. [PubMed: 11048879]

5. Mayr NA, Hawighorst H, Yuh WT, Essig M, Magnotta VA, Knopp MV. MR microcirculation assessment in cervical cancer: correlations with histomorphological tumor markers and clinical outcome. J Magn Reson Imaging. 1999; 10:267–276. [PubMed: 10508286]

6. Brix G, Schreiber W, Hoffmann U, Guckel F, Hawighorst H, Knopp MV. Methodological approaches to quantitative evaluation of microcirculation in tissues with dynamic magnetic resonance tomography. Radiologe. 1997; 37:470–480. [PubMed: 9340677]

7. Hoskin PJ, Saunders MI, Goodchild K, Powell ME, Taylor NJ, Baddeley H. Dynamic contrast enhanced magnetic resonance scanning as a predictor of response to accelerated radiotherapy for advanced head and neck cancer. Br J Radiol. 1999; 72:1093–1098. [PubMed: 10700827]

8. Neeman M, Dafni H. Structural, functional, and molecular MR imaging of the microvasculature. Annu Rev Biomed Eng. 2003; 5:29–56. [PubMed: 14527310]

9. Link KH, Leder G, Pillasch J, et al. In vitro concentration response studies and in vitro phase II tests as the experimental basis for regional chemotherapeutic protocols. Semin Surg Oncol. 1998; 14:189–201. [PubMed: 9548601]

10. Galmarini CM, Galmarini FC. Multidrug resistance in cancer therapy: role of the microenvironment. Curr Opin Investig Drugs. 2003; 4:1416–1421.

11. Galmarini FC, Galmarini CM, Sarchi MI, Abulafia J, Galmarini D. Heterogeneous distribution of tumor blood supply affects the response to chemotherapy in patients with head and neck cancer. Microcirculation. 2000; 7:405–410. [PubMed: 11142337]

12. Ellingsen C, Natvig I, Gaustad JV, Gulliksrud K, Egeland TA, Rofstad EK. Human cervical carcinoma xenograft models for studies of the physiological microenvironment of tumors. J Cancer Res Clin Oncol. 2009; 135:1177–1184. [PubMed: 19214568]

13. Hockel M, Knoop C, Schlenger K, et al. Intratumoral pO2 predicts survival in advanced cancer of the uterine cervix. Radiother Oncol. 1993; 26:45–50. [PubMed: 8438086]

14. Yuh WT, Mayr NA, Jarjoura D, et al. Predicting control of primary tumor and survival by DCE MRI during early therapy in cervical cancer. Invest Radiol. 2009; 44:343–350. [PubMed: 19661844]

15. Semple SI, Harry VN, Parkin DE, Gilbert FJ. A combined pharmacokinetic and radiologic assessment of dynamic contrast-enhanced magnetic resonance imaging predicts response to chemoradiation in locally advanced cervical cancer. Int J Radiat Oncol Biol Phys. 2009; 75:611–617. [PubMed: 19735887]

16. Loncaster JA, Carrington BM, Sykes JR, et al. Prediction of radiotherapy outcome using dynamic contrast enhanced MRI of carcinoma of the cervix. Int J Radiat Oncol Biol Phys. 2002; 54:759–767. [PubMed: 12377328]

17. Mayr NA, Yuh WT, Magnotta VA, et al. Tumor perfusion studies using fast magnetic resonance imaging technique in advanced cervical cancer: a new noninvasive predictive assay. Int J Radiat Oncol Biol Phys. 1996; 36:623–633. [PubMed: 8948347]

18. Yamashita Y, Baba T, Baba Y, et al. Dynamic contrast-enhanced MR imaging of uterine cervical cancer: pharmacokinetic analysis with histopathologic correlation and its importance in predicting the outcome of radiation therapy. Radiology. 2000; 216:803–809. [PubMed: 10966715]

19. Gong QY, Brunt JN, Romaniuk CS, et al. Contrast enhanced dynamic MRI of cervical carcinoma during radiotherapy: early prediction of tumour regression rate. Br J Radiol. 1999; 72:1177–1184. [PubMed: 10703475]

20. Mayr NA, Yuh WT, Arnholt JC, et al. Pixel analysis of MR perfusion imaging in predicting radiation therapy outcome in cervical cancer. J Magn Reson Imaging. 2000; 12:1027–1033. [PubMed: 11105046]

21. Mayr NA, Yuh WT, Jajoura D, et al. Ultra-early predictive assay for treatment failure using functional magnetic resonance imaging and clinical prognostic parameters in cervical cancer. Cancer. 2010; 116:903–912. [PubMed: 20052727]

22. Hamstra DA, Galban CJ, Meyer CR, et al. Functional diffusion map as an early imaging biomarker for high-grade glioma: correlation with conventional radiologic response and overall survival. J Clin Oncol. 2008; 26:3387–3394. [PubMed: 18541899]

23. Zahra MA, Tan LT, Priest AN, et al. Semiquantitative and quantitative dynamic contrast-enhanced magnetic resonance imaging measurements predict radiation response in cervix cancer. Int J Radiat Oncol Biol Phys. 2009; 74:766–773. [PubMed: 19019563]

24. Pecorelli S. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. Int J Gynaecol Obstet. 2009; 105:103–104. [PubMed: 19367689]

25. Mayr NA, Yuh WT, Zheng J, et al. Tumor size evaluated by pelvic examination compared with 3-D quantitative analysis in the prediction of outcome for cervical cancer. Int J Radiat Oncol Biol Phys. 1997; 39:395–404. [PubMed: 9308943]

26. Mayr NA, Taoka T, Yuh WT, et al. Method and timing of tumor volume measurement for outcome prediction in cervical cancer using magnetic resonance imaging. Int J Radiat Oncol Biol Phys. 2002; 52:14–22. [PubMed: 11777618]

27. Yuh WT. An exciting and challenging role for the advanced contrast MR imaging. J Magn Reson Imaging. 1999; 10:221–222. [PubMed: 10508280]

28. Hilbe, JM.; Joseph, M. Logistic regression models. Chapman & Hall/CRC Press; 2009.

29. Zhao S, Kuge Y, Mochizuki T, et al. Biologic correlates of intratumoral heterogeneity in 18F-FDG distribution with regional expression of glucose transporters and hexokinase-II in experimental tumor. J Nucl Med. 2005; 46:675–682. [PubMed: 15809491]

30. Hockel M, Schlenger K, Aral B, Mitze M, Schaffer U, Vaupel P. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. Cancer Res. 1996; 56:4509–4515. [PubMed: 8813149]

31. Britten RA, Evans AJ, Allalunis-Turner MJ, Franko AJ, Pearcey RG. Intratumoral heterogeneity as a confounding factor in clonogenic assays for tumour radioresponsiveness. Radiother Oncol. 1996; 39:145–153. [PubMed: 8735482]

32. Bachtiary B, Boutros PC, Pintilie M, et al. Gene expression profiling in cervical cancer: an exploration of intratumor heterogeneity. Clin Cancer Res. 2006; 12:5632–5640. [PubMed: 17020965]

33. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009; 45:228–247. [PubMed: 19097774]

34. Kastritis E, Bamias A, Efstathiou E, et al. The outcome of advanced or recurrent non-squamous carcinoma of the uterine cervix after platinum-based combination chemotherapy. Gynecol Oncol. 2005; 99:376–382. [PubMed: 16051322]

35. Eifel PJ, Morris M, Wharton JT, Oswald MJ. The influence of tumor size and morphology on the outcome of patients with FIGO stage IB squamous cell carcinoma of the uterine cervix. Int J Radiat Oncol Biol Phys. 1994; 29:9–16. [PubMed: 8175451]

36. van Nagell JRJ, Roddick JW Jr, Lowin DM. The staging of cervical cancer: inevitable discrepancies between clinical staging and pathologic findings. Am J Obstet Gynecol. 1971; 110:973–978. [PubMed: 5558978]

37. Mayr NA, Wang JZ, Zhang D, et al. Longitudinal changes in tumor perfusion pattern during the radiation therapy course and its clinical impact in cervical cancer. Int J Radiat Oncol Biol Phys. 2010; 77:502–508. [PubMed: 19775824]

38. Mayr NA, Huang Z, Wang JZ, et al. Characterizing tumor heterogeneity with functional imaging and quantifying high-risk tumor volume for early prediction of treatment outcome: cervical cancer as a model. Int J Radiat Oncol Biol Phys. 2012; 83:972–979. [PubMed: 22208967]

Statement of translation relevance

Current cancer treatment strategy has not routinely incorporated tumor heterogeneity known critically influencing cancer treatment outcome. With advancements in clinical functional imaging, each tumor imaging voxel's functional biomarkers can be derived throughout the entire tumor to evaluate tissue heterogeneity. Dynamic Contrast Enhanced (DCE) MRI has demonstrated efficacy in assessing at-risk low DCE perfusion tumor voxels that are associated with treatment failure. However, DCE-MRI has not been fully translated into the clinical setting. Therefore, the establishment of clinically validated DCE threshold metrics before or during early treatment to effectively target and define at-risk tumor voxels would have significant impact on personalized risk assessment and early prediction of treatment failure. This research clinically validated the optimal DCE thresholds at 3 different time points: before treatment and 2 weeks or 4 weeks into treatment, and a single universal threshold for all 3 time points to significantly classify at-risk tissue and predict treatment failure.
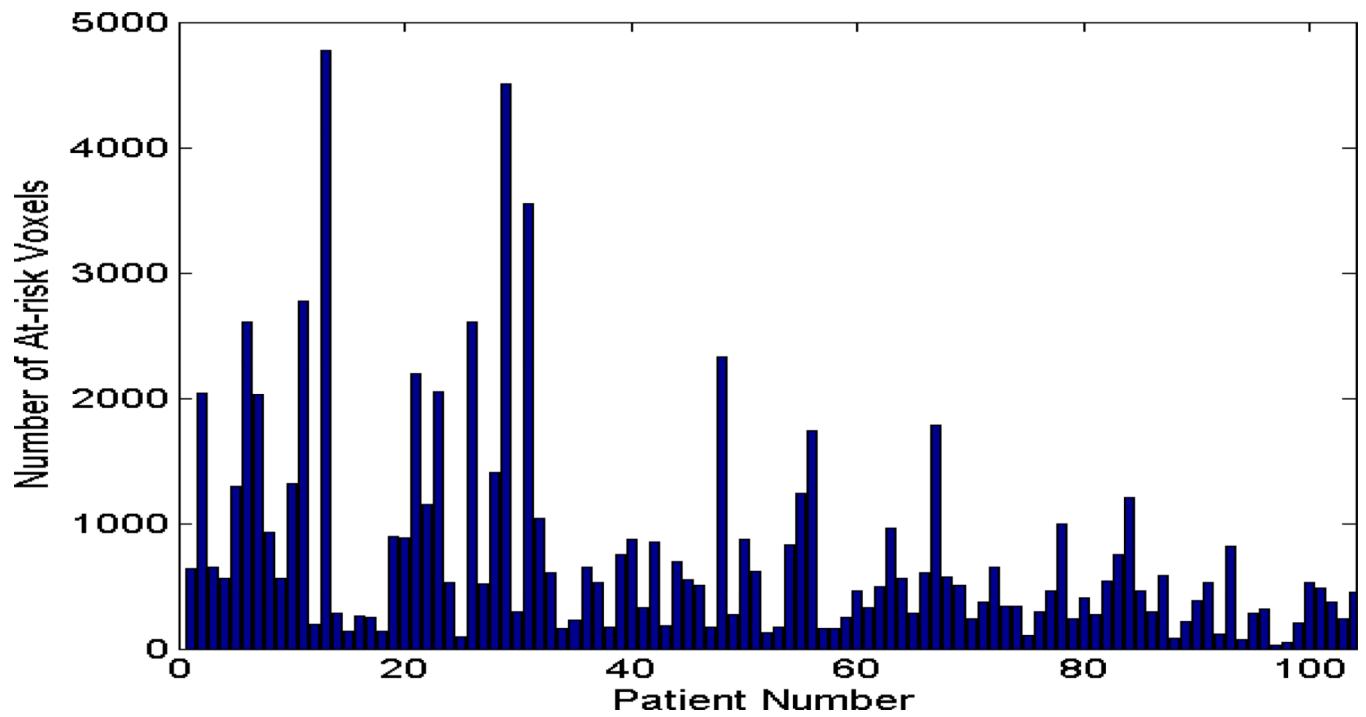
Figure 1. Distribution of at-risk tumor voxels

The distribution of the total number of at-risk tumor voxels of the 102 tumors, which were classified by an example using a specific SI threshold = 2.1, from a DCE MRI study obtained at 2–2.5 weeks into treatment. There were 21 distributions generated by varying the SI values from 1.0 to 3.0 with an increment of 0.1 for ROC analyses.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
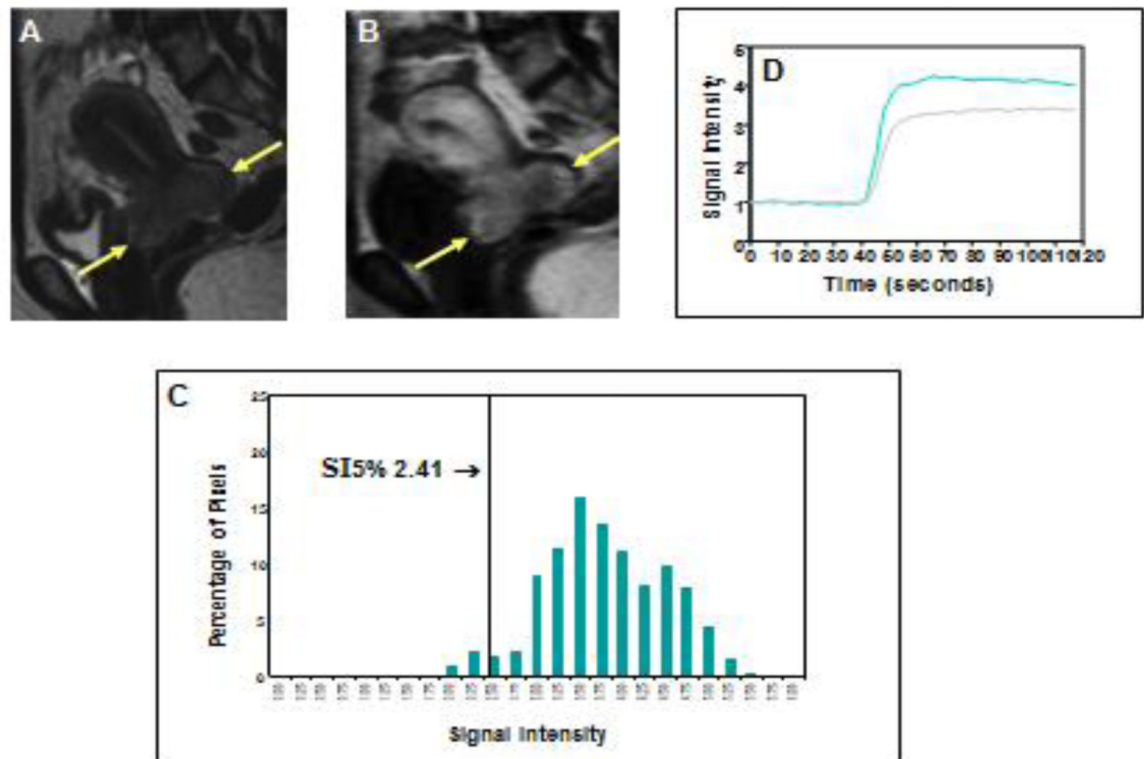
Figure 2. Favorable clinical outcome with high perfusion status early during treatment
Precontrast sagittal T2-weighted image (A) shows a large cervical cancer (arrows) indicating poor prognosis as judged by FIGO criteria. Corresponding DCE MRI at the plateau phase of the DCE imaging (B) shows the large tumor (arrows) with intense and heterogeneous dynamic contrast enhancement, indicating high tumor perfusion during the early part (2 weeks) of the 8-week treatment course. Tumor SI-pixel histogram (C) of the tumor is generated with the SI of the entire tumor pixels by plotting the SI of each tumor pixel along the x-axis and the number of pixels with same SI (frequency) along the y-axis. From this tumor SI histogram, tumor size (area under the DCE curve) can be calculated. The quantity and degree of low DCE subpopulation was quantified by computing SI percentiles. In this case, the 5th percentile (SI5%) (arrow, C) was 2.41, indicating that 5% of the pixels within the heterogeneous tumor fall below an SI of 2.41. This patient had excellent treatment outcome and survived for more than 10 years. Signal intensity-time DCE curves (D) again show abrupt and intense contrast enhancement of the tumor, indicating high tumor perfusion in both MRI 1 (gray-colored curve) and MRI 2 (blue-colored curve), and predicting excellent long term outcome.
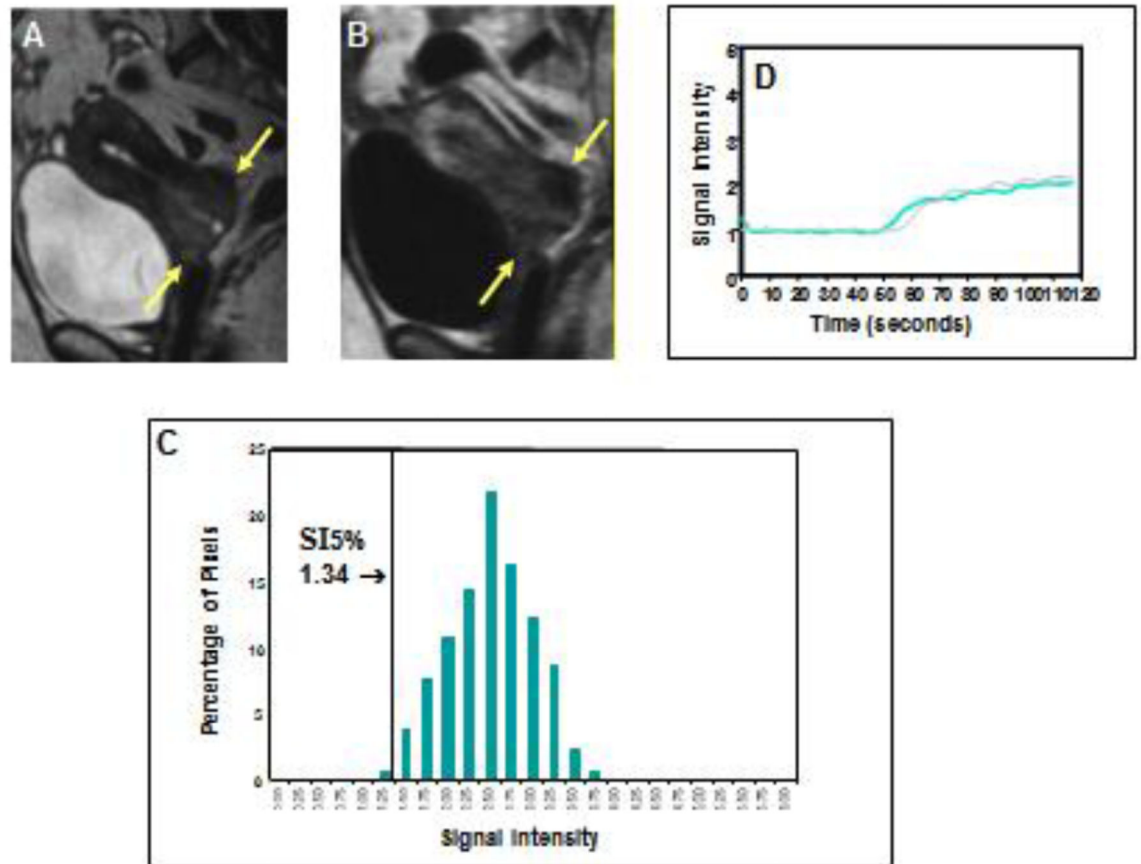
Figure 3. Poor clinical outcome with low perfusion status early during treatment for a small tumor

Precontrast sagittal T2-weighted image (A) showed a much smaller cervical cancer (arrows), as compared with Figure 1, indicating a more favorable prognosis as judged by FIGO criteria. Corresponding DCE MRI at the plateau phase of the DCE imaging (B) showed the small tumor (arrows) with poor contrast enhancement indicating low tumor perfusion during early part (2 weeks) of the treatment course (8 weeks). From the tumor SI-pixel histogram (C), the low-DCE subpopulation was characterized by 5th percentile of SI (SI5%). SI5% was much lower (1.34) than that of Figure 1 (2.41). This patient had a primary tumor recurrence at 2 months after completion of therapy, and died 6 months after completion of therapy. Signal intensity-time DCE curves (D) again show sluggish and low contrast enhancement of the tumor mass indicating low tumor perfusion and poor response to treatment. In contrast to Figure 1, tumor perfusion remains low even 2 to 3 weeks after initiation of therapy (blue-colored curve) and is similar to that of the pretreatment (gray-colored curve).
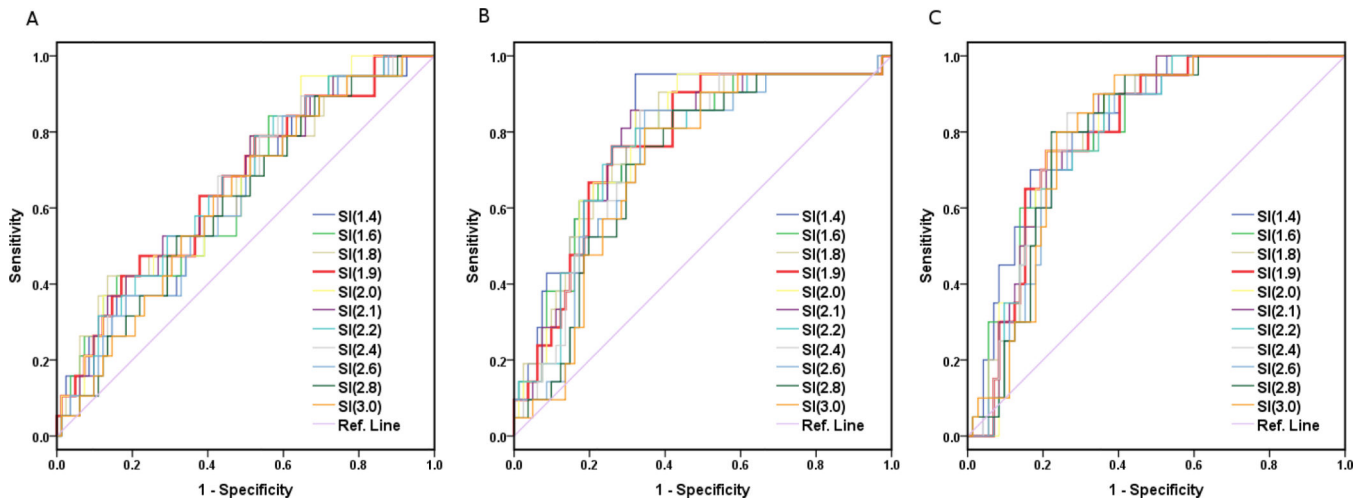
Figure 4. ROC Curves at 3 treatment time points for LC only
ROC graphs only for local control (LC) at (A) pre-treatment, and (B) 2–2.5 weeks and (C) 4–4.5 weeks into treatment. To conserve space, only partial ROC curves are shown. The ROC graph of SI=1.9 (red) is also included to demonstrate its relationship and performance in comparison with other ROC curves. Overall, these ROC curves show a gross trend of gradual increase of AUC from A to B and C with the maximal AUC values of 0.66, 0.78 and 0.80, respectively, and the corresponding optimal SI thresholds values of 2.1, 2.1 and 2.0, respectively. Note that the thickest ROC curve (red) represents the universal SI=1.9, which shows an excellent overall performance (toward the left and top) compared to most of the other ROC curves.

Table 1

Summary of clinical information (102 patients)

Average age	55 (25.0–89.0 years)
FIGO stage	
I	15 (14%)
II	46 (45%)
III	31 (31%)
IV	10 (10%)
Histology	
Squamous cell	84 (82%)
Adenocarcinoma	18 (18%)
Brachytherapy	
Low-dose-rate	98 (96%)
High-dose-rate	4 (4%)
Chemotherapy	48 (46%)

Table 2

Optimal SI threshold value to classify unfavorable tumor voxels at different imaging times

Summary of ROC Analysis for Optimal SI and Universal Threshold

Ultimate treatment outcome	Local tumor control (LC)			Disease-specific survival (DSS)		
MRI Time	**Pre-RT**	**20–25 Gy**	**40–50 Gy**	**Pre-RT**	**20–25 Gy**	**40–50 Gy**
Max AUC (Area under the curve)	0.66	0.78	0.80	0.66	0.71	0.69
Time-specific optimal SI threshold	2.1	2.1	2.0	1.8	2.1	1.9
* p-value (time-specific threshold)	0.02	$8.0 \cdot 10^{-5}$	$3.7 \cdot 10^{-5}$	0.009	0.001	0.006
* p-value (universal threshold of 1.9)	0.025	$1.4 \cdot 10^{-4}$	$4.2 \cdot 10^{-5}$	0.006	0.002	0.006

Optimal SI thresholds to define at-risk voxels, determined by ROC analysis, varied at different treatment time points. Threshold values ranged from 1.8–2.1 (row 5) before treatment (#1), and at 2–2.5 weeks (#2) or 4–4.5 weeks into treatment, respectively. These time-specific optimal SI thresholds correlated significantly with treatment outcomes (row 6). Significant correlation remains when a single universal SI threshold (1.9) is applied to define the unfavorable voxels for all three time points (row 6).

*
Mann-Whitney U test.

Table 3

ROC analysis of the 8 new variables shows maximal

AUC/SI	1.8	1.9	2.0	2.1
LC	0.737	0.742[*]	0.708	0.710
DSS	0.690	0.719[*]	0.659	0.688

[*] AUC at SI=1.9 for both LC and DDS