Published in final edited form as:

J Magn Reson Imaging. 2018 May ; 47(5): 1388–1396. doi:10.1002/jmri.25874.

Tumor Radiomic Heterogeneity: Multiparametric Functional Imaging to Characterize Variability and Predict Response Following Cervical Cancer Radiation Therapy

Stephen R. Bowen, PhD^{1,2}, William T. C. Yuh, MD², Daniel S. Hippe, MS², Wei Wu, MD³, Savannah C. Partridge, PhD², Saba Elias, MS⁴, Guang Jia, PhD⁵, Zhibin Huang, PhD⁶, George A. Sandison, PhD¹, Dennis Nelson, PhD⁷, Michael V. Knopp, MD⁴, Simon S. Lo, MD¹, Paul E. Kinahan, PhD², and Nina A. Mayr, MD¹

¹University of Washington School of Medicine, Department of Radiation Oncology, 1959 NE Pacific St, Seattle, WA 98195, USA

²University of Washington School of Medicine, Department of Radiology, 1959 NE Pacific St, Seattle, WA 98195, USA

³Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology, Department of Radiology, Wuhan, Hubei, CN

⁴The Ohio State University, Department of Radiology, Columbus, OH

⁵Louisiana State University, Department of Physics, Baton Rouge, LA

⁶East Carolina University, Department of Radiation Oncology, Greenville, NC

⁷MIM Software, Inc., Cleveland, OH

Abstract

Background—Robust approaches to quantify tumor heterogeneity are needed to provide early decision support for precise individualized therapy.

Purpose—To conduct a technical exploration of longitudinal changes in tumor heterogeneity patterns on DCE MRI, DWI and FDG PET/CT, and their association to radiation therapy (RT) response in cervical cancer.

Study Type—Prospective observational study with longitudinal MRI and PET/CT pre-RT, early-RT (2 weeks), and mid-RT (5 weeks).

Corresponding author contact information: Stephen R. Bowen, Ph.D., Assistant Professor, Departments of Radiation Oncology and Radiology, University of Washington Medical Center, 1959 NE Pacific St, Box 356043, Seattle, WA 98195, srbowen@uw.edu.

Disclosure of Potential Conflicts of Interest

Stephen R. Bowen was supported by a Research Scholar Award from the Radiological Society of North American and an NIH/NCI grant for work performed as part of the current study. William T.C. Yuh, Daniel S. Hippe, Savannah C. Partridge, Michael V. Knopp, and Nina A. Mayr were all supported by an NIH/NCI grant for work performed as part of the current study. Dennis Nelson has a commercial interest as President of MIM Software, Inc. for work performed outside of the current study. Paul Kinahan has a commercial interest as co-founder of PET/X, LLC for work performed outside of the current study. Saba Elias, Guang Jia, Zhibin Huang, Norman J. Beauchamp, George A. Sandison, and Simon S. Lo declare no conflicts of interest.

Population—Twenty-one FIGO IB₂-IVA cervical cancer patients receiving definitive external beam RT and brachytherapy.

Field Strength/Sequence—1.5 Tesla, pre-contrast axial T1-weighted, axial and sagittal T2-weighted, sagittal DWI (multi-b values), sagittal DCE MRI (<10 sec temporal resolution), post-contrast axial T1-weighted.

Assessment—Response assessment one month after completion of treatment by board-certified radiation oncologist from manually delineated tumor volume changes.

Statistical Tests—Intensity histogram (IH) quantiles (DCE SI_{10%} and DWI ADC_{10%}, FDG-PET SUV_{max}) and distribution moments (mean, variance, skewness, kurtosis) were extracted. Differences in IH features between timepoints and modalities were evaluated by Skillings-Mack tests with Holm's correction. Area under receiver-operating characteristic curve (AUC) and Mann-Whitney testing was performed to discriminate treatment response using IH features.

Results—Tumor IH means and quantiles varied significantly during RT (SUV_{mean}: ↓28–47%, SUV_{max}: ↓30–59%, SI_{mean}: ↑8–30%, SI_{10%}: ↑8–19%, ADC_{mean}: ↑16%, $p < 0.02$ for each). Among IH heterogeneity features, FDG-PET SUV_{CoV} (↓16–30%, $p = 0.011$) and DW-MRI ADC_{skewness} decreased ($p = 0.001$). FDG-PET SUV_{CoV} was higher than DCE-MRI SI_{CoV} and DW-MRI ADC_{CoV} at baseline ($p < 0.001$) and 2 weeks ($p = 0.010$). FDG-PET SUV_{kurtosis} was lower than DCE-MRI SI_{kurtosis} and DW-MRI ADC_{kurtosis} at baseline ($p = 0.001$). Some IH features appeared to associate with favorable tumor response, including large early RT changes in DW-MRI ADC_{skewness} (AUC=0.86).

Data Conclusion—Preliminary findings show tumor heterogeneity was variable between patients, modalities, and time points. Radiomic assessment of changing tumor heterogeneity has potential to personalize treatment and power outcome prediction.

Keywords

MRI; PET; DCE; DWI; tumor heterogeneity; radiomics

Introduction

Although heterogeneity of tumor cells [1] and their microenvironment [2–4] greatly impact patient-specific response to cancer therapy, including cervical cancer therapy [5, 6], strategies for objectively quantifying and associating tumor heterogeneity to outcome have yet to be standardized and validated for clinical implementation. Evaluation of regional tumor microscopic heterogeneity can be achieved by invasive biopsy, but cannot evaluate the entire tumor bed, and is prone to sampling error. Furthermore, it is difficult to perform sequential biopsy of specific tumor regions and assess temporal changes due to morphologic changes of the tumor as it shrinks during treatment [7]. A robust and non-invasive approach to quantify tumor heterogeneity, prior to or early during treatment, may have a profound impact on the management of individual cancer patients by providing early prediction of treatment outcome and the opportunity to tailor therapy as part of the precision medicine paradigm [8].

Radiomics is an expanding field of clinical research that extracts high dimensional sets of imaging features for association to clinical diagnosis, prognosis, and therapeutic interventions. Unlike tissue biomarkers which invasively sample a small tumor region at the micron spatial scale, radiomic biomarkers non-invasively interrogate the entire tumor at the millimeter scale over multiple time points. Radiomics features of tumor heterogeneity show promise as predictive correlates of clinical outcomes in a variety of cancer patients [9] and have also been linked to oncogenic expression [10]. The majority of published investigations on tumor radiomics have focused on a single imaging modality, predominantly CT, MRI and to a lesser extent PET.

In cervical cancer patients, changes in FDG PET standardized uptake value (SUV) metrics post-therapy predict for recurrent disease and clinical outcome [11, 12], but these changes after treatment completion cannot be utilized to directly adapt or alter the initial definitive treatment course in order to improve outcome. Pre-treatment FDG PET uptake gradients are also associated with treatment response and recurrence risk, but correlation of these uptake gradients with simpler volumetric imaging measures suggests limited complementary information [13]. High dimensional sets of textural features on FDG PET change dynamically with therapy [14]. Further validation of their clinical utility for patient treatment outcome prediction will depend on judicious feature selection [15]. Diffusion-weighted (DW) MRI has proven useful in the management of cervical cancer [16], with features of apparent diffusion coefficient (ADC) that detect early signs of treatment response [17]. It remains unclear whether combinations of MRI and PET metrics of tumor heterogeneity can demonstrate improved prediction of treatment response and clinical outcome [18], particularly when evaluating temporal changes of multimodal imaging features across treatment time points.

The purpose of this study was to conduct a technical exploration of longitudinal changes in tumor heterogeneity patterns on DCE MRI, DWI and FDG PET/CT, as well as preliminary association with response to radiation therapy (RT) therapy in cervical cancer patients. This extends prior work on DCE MRI features of at-risk tumor regions that demonstrated high prognostic value [19, 20].

Methods and materials

Patient and Treatment Characteristics

Twenty-one patients with cervical cancer of variable stage (FIGO IB₂-IVA) were prospectively enrolled onto an observational study (NCT01992861) following approval by institutional review board at the University of Washington and after obtaining each patient's informed consent. Three serial MRI and three serial PET examinations were performed before and during treatment. Pre-treatment clinical evaluations included routine initial clinical work-up following FIGO guidelines [21]. Patients were treated with standard external beam RT, concurrent weekly cisplatin-based chemotherapy and standard-of-care brachytherapy. External beam RT dose prescriptions ranged from 45.0–50.4 Gy in conventional 1.8 Gy per fraction, including whole pelvis and cone down radiation fields. High dose rate ¹⁹²Ir brachytherapy ranged from 30.0–39.6 Gy biologically equivalent dose in 2 Gy according to the linear-quadratic model of cell survival for cervical cancer ($\alpha/\beta=10$).

DCE MRI, DW MRI and FDG PET/CT acquisition

Patients underwent serial MRI and PET examinations at three time points: prior to external beam RT, early during external beam RT (at 20–25 Gy of pelvic RT), and midway during RT (at 40–50 Gy). MRI examinations were performed on 1.5 Tesla scanners (SIGNA series, GE Healthcare, Waukesha, WI) and included pre-contrast axial T1-weighted, axial and sagittal T2-weighted, and sagittal DWI sequences, followed by sagittal DCE MRI and post-contrast axial T1-weighted sequences. Further details of the DCE MRI imaging protocol can be found in a prior study [22].

The images used for quantitative analysis in the research study were the matched sagittal T2-weighted, DWI, and DCE MRI sequences, described in more detail below. Pre-contrast T2-weighted imaging was performed using a 2D fast spin echo sequence with TR 300–5000 ms, TE 100–120 ms, field-of-view (FOV) of 30–40 cm, matrix 256×256, slice thickness 3–5 mm, 3–4 averages. DWI was performed using a 2D single-shot echo planar imaging sequence with TR 6000–8000 ms, TE 80–110 ms, FOV of 30–40 cm, matrix 128 × 128, slice thickness 5–8 mm, and b values of 0, 100, 600, and 1000 s/mm². DCE MRI was performed using a high temporal resolution 3D fast gradient echo sequence with TR 5–7 ms, TE 1–3 ms, flip angle 15–30°, FOV of 30–40 cm, matrix 256 × 256, slice thickness 5–8 mm, temporal resolution of <10 sec, and total scan time of 10 minutes. An intravenous bolus of 1 mmol/kg commercially available gadolinium chelated MR contrast agent was administered by power injector after the fifth DCE phase, at an injection rate equal to or greater than 5 mL/s.

FDG PET/CT examinations (Discovery STE, GE Healthcare, Waukesha, WI) were conducted at time points that coincided with MRI examinations prior to and during radiation therapy. The exam consisted of a helical CT for attenuation correction, reconstructed to 2 mm slice thickness for improved MR-CT registration, followed by a whole body static PET acquisition spanning 6–7 axial fields of view and 5 minutes per couch position. PET imaging commenced 60 minutes post weight-based FDG injection (nominal) with the most caudal axial FOV so as to minimize FDG bladder filling-induced artifact in the pelvis. Images were reconstructed using ordered subset expectation maximization (OSEM) and a post-reconstruction 3-dimensional Gaussian filter was applied.

Intensity Histogram Feature Extraction

DCE MRI SI ratio, DW MRI ADC, and FDG PET SUV images generated on the respective scanners were co-registered to the reference sagittal T2-weighted MR image at each time point using a rigid normalized mutual information algorithm (MIM 6.5, MIM Software Inc., Cleveland, OH) applied between MR-MR and MR-CT (PET). Image processing of DCE MRI to generate SI ratio maps has been described previously [22]. In brief, the first-pass DCE plateau phase was identified on the time-intensity curve. The plateau phase and pre-contrast phase of the DCE time series were fused and adjustments in alignment were applied to account for inter-frame tumor motion. From this fusion, the SI ratio of the plateau phase and pre-contrast phase was computed at every voxel which produced the final SI ratio images. Tumor and uterus volumes were delineated on the reference MRI (sagittal T2-weighted), with the latter structure assisting in final anatomic alignment of all images due to

bladder and rectal filling-induced changes in uterus position. The reference tumor contour at each time point was rigidly copied to the DCE MR SI ratio, DWI ADC, and FDG PET image sets and converted to their respective native digital grids. Voxels on DWI ADC maps that suffered from significant distortion artifacts at air-tissue interfaces were excluded from the analysis through automatic thresholding, while voxels on FDG PET impacted by spill-in of bladder signal were similarly removed. Voxel value distributions of each image modality were extracted from the tumor contour on the native grid without digital resampling, which preserved covariance in image voxel neighborhoods at the cost of contour boundary variability. The resultant process introduced a small contour transformation rather than a large image value transformation. Due to these variations in contour boundaries between imaging modalities, particularly in small residual tumors at the 5-week mid-RT time point, we focused our investigation on intensity histogram features. Other radiomic features of image texture can be sensitive to ROI size and ROI boundary shape [23].

The following set of intensity histogram (IH) features were extracted from the tumor ROI at each time point for each image modality (DCE SI ratio, DWI ADC, FDG PET SUV):

1. mean: 1st moment of intensity distribution
2. coefficient of variation: standard deviation divided by the mean, representing normalized 2nd moment of intensity distribution
3. skewness: normalized 3rd moment of intensity distribution
4. kurtosis: normalized 4th moment of intensity distribution
5. quantiles: DCE SI ratio 10th percentile, DWI ADC 10th percentile, FDG PET SUV maximum of the respective intensity distributions.

The moments of the intensity distribution describe the shape of the intensity histogram as statistical measures of heterogeneity. The quantiles represent the most commonly reported values associated with specific tumor voxels/regions (low perfusion, low diffusion, high glucose metabolism) that are presumed to be more resistant to therapy [22, 24].

Statistical analysis

For each image modality, differences in IH feature distributions between time points (pre-treatment, 2-week early-treatment, and 5-week mid-treatment) were evaluated using the non-parametric Skillings-Mack test [25]. This statistical test is analogous to the Wilcoxon signed-rank test for paired data but allows for three or more groups to be compared. At each time point, differences in IH feature distributions between modalities (FDG PET, DCE MRI, DW MRI) were also evaluated using the Skillings-Mack test. Two-tailed *p* values were adjusted for multiple comparisons using Holm's method, which is a more powerful alternative to the Bonferroni correction [26]. The multiple comparison adjustment was applied separately to the group of tests comparing time points and the group of tests comparing modalities.

An exploratory analysis of individual IH features and changes in features during therapy were compared between patients with and without tumor response assessed at 1-month after the completion of brachytherapy, using the Mann-Whitney test. Tumor response was defined

as <10% residual volume from baseline. The area under the receiver operating characteristic (ROC) curve (AUC) was used to summarize the strength of association between each individual IH feature and response. Multivariate regression analysis was not performed due to the small sample size. As this part of the analysis was exploratory and hypothesis-generating in nature, the corresponding p-values were not adjusted for multiple comparisons. All statistical analyses were conducted in OriginLab Pro 9.1 (OriginLab Corporation, Northampton, MA) or the R statistical computing language (version 3.1.1; R Foundation for Statistical Computing, Vienna, Austria).

Results

Images and image intensity histograms are shown for a representative patient in Figures 1 and 2, respectively, at each time point (pre-treatment, 2-week early-treatment, 5-week mid-treatment) and for each modality (FDG PET, DCE MRI, DW MRI). Figure 1 shows the time-dependent tumor (magenta contour) decrease in SUV on FDG PET (top row), increase in DCE signal enhancement on MRI (middle row), and variable ADC on MRI (bottom row). Figure 2 displays the corresponding intensity histograms of the images of the same patient, whose characteristic tumor heterogeneity included small increases in FDG PET SUV skewness (right-skewed) along with large decreases in DCE MR skewness and DW MR ADC skewness (left-skewed) with therapy. Large tumor volume regression was observed on the 5-week mid-treatment images and histograms.

Characteristic time-dependent trends were observed for the heterogeneity properties of DCE, ADC and FDG PET SUV during treatment. Mean and quantile IH features evolved significantly during RT, and the changes during therapy were not concordant among the functional imaging parameters. SUV_{mean} and SUV_{max} both decreased during treatment by 28–47% and by 30–59%, respectively (adjusted $p < 0.001$), relative to baseline values. Conversely, the DCE parameters SI_{mean} and $SI_{10\%}$ increased by 8–30% and 8–19%, respectively (adjusted $p < 0.015$), and ADC_{mean} increased by 16% (adjusted $p < 0.001$), relative to baseline values. Among the IH features of heterogeneity, greatest changes were observed in FDG PET SUV and ADC. The SUV coefficient of variation decreased intra-treatment by 16–30% (adjusted $p = 0.011$), and ADC skewness decreased by 0.62–0.95 (adjusted $p = 0.001$). SUV kurtosis showed only small longitudinal changes during treatment, whereas DCE (adjusted $p = 0.21$) and ADC kurtosis (adjusted $p = 0.075$) numerically decreased from the pre-therapy to the 2-week mid-treatment imaging, though these trends did not reach statistical significance. Figure 3 summarizes the time-dependent changes of MRI and PET image features during radiation therapy for the patient cohort.

Figure 4 illustrates the imaging modality-dependent differences in intensity histogram features at each time point prior to and during treatment. The coefficient of variation of FDG PET SUV was significantly higher than the coefficient of variation of DCE MRI SI or DW MRI ADC prior to treatment (adjusted $p < 0.001$) and at 2 weeks during treatment (adjusted $p = 0.01$). The kurtosis of FDG PET SUV was lower prior to treatment than kurtosis of DCE MRI SI and DW MRI ADC (adjusted $p = 0.001$).

A hypothesis-generating analysis was conducted, correlating 1-month post treatment response with individual IH features (Table 1). Out of the 21 subjects, one without sufficient follow up to determine response was excluded from this analysis. Six features were found to be significantly associated with response ($p < 0.05$ without adjustment for multiple comparisons), with greater decreases in DW-MRI ADC_{skewness} (becoming more left-skewed) between the baseline and 2 week scan having strongest observed association with favorable treatment response (AUC = 0.86).

Discussion

Although biological heterogeneity of cervical cancer plays an important role in patient-specific tumor sensitivity to therapy [5], strategies for objectively quantifying and associating tumor heterogeneity to outcome have yet to be standardized and validated for clinical implementation. Functional and molecular imaging voxels represent millimeter-scale resolution, and thereby are incapable of resolving the microscopic heterogeneity observable on tumor biopsy specimens. On the other hand, invasive biopsy-based evaluation does not encompass the entire tumor bed nor can it perform reliable longitudinal assessment of tumor heterogeneity during therapy [7, 27]. Invasive biopsy assessment of tumor heterogeneity is also less feasible for deeply seated tumors and those in close proximity to vital structures, such as major nerves and blood vessels. Despite its limitations, image-voxel derived heterogeneity as part of radiomics analysis is closely related to biologic heterogeneity when using PET [28, 29], CT [10, 30], or MRI [31, 32]. Such radiomic heterogeneity provides a non-invasive technique to obtain essential information of “macro-heterogeneity” over the entire tumor burden of disease, and can quantify temporal changes in tumor heterogeneity during cytotoxic cancer therapy. Macroscopic tumor heterogeneity derived from image voxel intensity distributions would also facilitate clinical translation of targeted tumor treatment, such as surgical resection or radiation therapy that carry the same inherent millimeter level of delivery precision. Radiomic biomarkers of tumor heterogeneity may eventually complement existing genomic and proteomic biomarkers to form a unique patient profile that informs personalized care strategies.

This exploratory technical investigation established a generalized framework for extracting multimodal imaging features of tumor heterogeneity, and demonstrated two key observations: tumor heterogeneity varies between image modality for the same patient, and tumor heterogeneity varies between time points during cytotoxic therapy for the same image modality. Statistical differences in certain heterogeneity metrics between PET and MRI techniques suggest that they may provide complementary information for response assessment and outcome prediction of cervical cancer patients. The following longitudinal changes in image features of tumor heterogeneity across the patient cohort were observed between baseline and mid-therapy time points: FDG PET SUV decrease, DCE MRI SI (plateau enhancement ratio) increase, DW MRI ADC increase, FDG PET SUV coefficient of variation decrease, and DW MRI ADC skewness decrease. Our observation of DW MRI ADC increase during treatment is consistent with previous reports that show a general trend of improved diffusion in cervical cancer with radiation [17] that relate to increased water motility, increased microvessel density [33], and reduced tumor cellularity during treatment. Similarly, the observed DCE MRI SI increase suggests improvements in blood perfusion

with therapy [19, 20, 22, 34–37]. Lastly, the reduction in mean and maximum FDG PET SUV during therapy is in keeping with observations of declining tumor glycolytic metabolism [24] that relate to glucose transporter expression [28].

This represents the first study to report on DCE MRI, DWI, and FDG PET imaging features of cervical cancer heterogeneity in the same patient cohort and their changes during therapy. We found that FDG-PET intensity histogram (IH) feature changes during therapy are different than analogous DCE-MRI and DW-MRI IH features changes. We did not observe correlation of pre-therapy FDG-PET SUV with DW-MRI ADC as previously reported [38], nor were the time-dependent changes in parameters concordant between PET and MRI. Decreases in FDG-PET SUV coefficient of variation suggest that radiation therapy may lower intra-tumor metabolic heterogeneity. Decreases in DW-MRI ADC skewness are characteristic of large tumor region shifts to higher levels of diffusion coupled with smaller tumor regions of persistently low diffusion. Beyond population trends, high variability in IH features between individual patients was observed through large inter-quartile and overall ranges. Such inter-tumor variation and unique temporal evolution of imaging heterogeneity may provide critical insight into individual patient response to therapy.

A strength of the proposed framework is its design around intensity histogram distribution feature changes instead of image voxel changes during therapy, making it substantially less sensitive to stochastic uncertainties in image formation and therefore well-suited to multimodality patient data. It has been recently reported that these image formation uncertainties can have a detrimental effect on quantitation of certain radiomic features [23] and necessitate comprehensive standardization [39]. Methods to analyze multimodal image data on a grid of common voxel size can compromise quantitative accuracy of tumor radiomic heterogeneity, since MRI reconstruction grids consist of high in-plane resolution but coarse slice thickness in most 2D acquisitions, whereas PET reconstructions grids consist of nearly isotropic coarse resolution from 3D acquisitions. The intensity histogram feature extraction in this study was performed on the native digital reconstruction grid of each MRI and PET image to preserve the voxel intensity distributions. Furthermore, longitudinal response assessment of intensity histogram features does not require deformable image registration if independent contours are generated at each time point. This is an important consideration in disease sites with substantial morphological tumor changes in response to therapy, highlighted here when applied to cervical cancer. Lastly, distortion artifacts that present on DW-MRI ADC maps late in cervical cancer therapy due to variable air-tissue rectal interfaces challenge the accuracy of voxel-based response assessment and even textural feature response assessment but are mitigated by intensity histogram-based response assessment through careful exclusion of distorted tumor voxels.

As part of a hypothesis-generating sub-analysis, we identified several candidate IH features which may correlate with 1-month post-treatment response. In particular, greater decreases in DW-MRI ADC skewness at 2 weeks, relative to baseline, appeared to be associated with an increased likelihood of tumor response and had the largest estimated AUC among the IH features. This supports the argument that left-skewed distribution shifts to higher tumor diffusion result in favorably treatment response, which may be biologically tied to a higher proportion of tumor volume with declining cellularity as a result of sensitivity to radiation.

However, this sub-analysis is exploratory, without adjustment for the number of comparisons, and the associated findings need to be replicated in a larger study before they can be utilized.

The general framework to extract multimodal tumor heterogeneity features can identify candidate imaging biomarkers for application to outcome prediction during and following therapeutic intervention. Such predictions can inform on selection, dosing, and sequencing of novel combination therapies guided by functional imaging. This will promote personalized therapeutic approaches in patients that are adapted to radiomic profiles of tumor heterogeneity. The tools developed under this framework can be integrated into clinical workflows through standardized reporting of radiomic features, analogous to genomic and proteomic features. Causal relationships between these features can be characterized on a patient-specific basis [40], leading to future omics analyses of imaging and tissue biomarkers of tumor heterogeneity.

The investigation has several limitations. The sample size is small and we do not report long-term follow-up for outcome prediction. A larger sample size would provide the necessary statistical power to examine higher dimensional sets of radiomics features beyond image intensity, including those derived from image texture and shape. A larger sample would also power the construction and validation of a multivariate model for clinical outcome prediction. Given these limitations, our preliminary study was focused on characterization of select image features and identifying candidate imaging biomarkers of therapeutic response. Other limitations include anatomical variations between imaging studies, during the same imaging study (e.g. between different pulse sequences, frames of the DCE time-series) or at different treatment time points. These were caused by differences in bladder filling, rectal distention or imaging position. Strict protocol standardization, patient fasting instructions, image co-registration, and image intensity histogram feature extraction methodology minimized the impact of this variability. Despite the small sample size and other inherent limitations, we were able to detect substantial changes in IH features over time and differences between modalities while accounting for multiple comparisons. Further investigation of these candidate imaging biomarkers of tumor heterogeneity is warranted in a larger patient population to train and validate multivariate models for outcome prediction early during therapy.

An exploratory technical investigation of tumor radiomic heterogeneity analysis of multimodal DCE MRI, DW MRI and FDG PET was developed to identify candidate imaging biomarkers of treatment response. Image intensity histogram features of tumor heterogeneity varied between patients, between modalities, and between time points during radiation therapy. The characteristic heterogeneity patterns shown with different imaging features and their temporal changes at different treatment time points offer great potential for incorporation into precision radiation oncology. These preliminary results will require further confirmation within a larger patient population and validation of tumor radiomic heterogeneity for clinical outcome prediction.

Acknowledgments

This investigation was supported by NIH/NCI R01CA155454 and a Research Scholar award (RSCH1405) from the Radiological Society of North America.

References

1. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochimica et biophysica acta*. 2010; 1805:105–17. [PubMed: 19931353]
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144:646–74. [PubMed: 21376230]
3. Bachtary B, Boutros PC, Pintilie M. Gene expression profiling in cervical cancer: an exploration of intratumor heterogeneity. *Clin Cancer Res*. 2006; 12:5632–40. [PubMed: 17020965]
4. 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; 135:1177–84. [PubMed: 19214568]
5. Grigsby PW, Zigelboim I, Powell MA, Mutch DG, Schwarz JK. In vitro chemoresponse to cisplatin and outcomes in cervical cancer. *Gynecologic oncology*. 2013; 130:188–91. [PubMed: 23583416]
6. Simpson-Herren L, Noker PE, Wagoner SD. Variability of tumor response to chemotherapy. II. Contribution of tumor heterogeneity. *Cancer Chemother Pharmacol*. 1988; 22:131–6. [PubMed: 3409443]
7. Mayr NA, Yuh WTC, Taoka T, Wang JZ, Wu DH, Montebello JF, et al. Serial therapy-induced changes in tumor shape in cervical cancer and their impact on assessing tumor volume and treatment response. *AJR American journal of roentgenology*. 2006; 187:65–72. [PubMed: 16794157]
8. Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. *The New England journal of medicine*. 2012; 366:489–91. [PubMed: 22256780]
9. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012; 48:441–6. [PubMed: 22257792]
10. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*. 2014; 5:4006.
11. Schwarz JK, Siegel BA, Dehdashti F, Grigsby PW. Metabolic response on post-therapy FDG-PET predicts patterns of failure after radiotherapy for cervical cancer. *International journal of radiation oncology, biology, physics*. 2012; 83:185–90.
12. Pallardy A, Bodet-Milin C, Oudoux A, Champion L, Bourbouloux E, Sagan C, et al. Clinical and survival impact of FDG PET in patients with suspicion of recurrent cervical carcinoma. *European journal of nuclear medicine and molecular imaging*. 2010; 37:1270–8. [PubMed: 20306032]
13. Kidd EAGP. Intratumoral metabolic heterogeneity of cervical cancer. *Clin Cancer Res*. 2008; 15:5236–41.
14. Yang F, Thomas MA, Dehdashti F, Grigsby PW. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *European journal of nuclear medicine and molecular imaging*. 2013; 40:716–27. [PubMed: 23340594]
15. Brooks FJ, Grigsby PW. FDG uptake heterogeneity in FIGO IIb cervical carcinoma does not predict pelvic lymph node involvement. *Radiat Oncol*. 2013; 8:294. [PubMed: 24365202]
16. Chen J, Zhang Y, Liang B, Yang Z. The utility of diffusion-weighted MR imaging in cervical cancer. *European journal of radiology*. 2010; 74:e101–6. [PubMed: 19442466]
17. Harry VN, Semple SI, Gilbert FJ, Parkin DE. Diffusion-weighted magnetic resonance imaging in the early detection of response to chemoradiation in cervical cancer. *Gynecologic oncology*. 2008; 111:213–20. [PubMed: 18774597]
18. Vandecasteele K, Delrue L, Lambert B, Makar A, Lambein K, Denys H, et al. Value of magnetic resonance and (1)(8)FDG PET-CT in predicting tumor response and resectability of primary

- locally advanced cervical cancer after treatment with intensity-modulated arc therapy: a prospective pathology-matched study. *International journal of gynecological cancer: official journal of the International Gynecological Cancer Society*. 2012; 22:630–7. [PubMed: 22237382]
19. Huang Z, Mayr NA, Lo SS, Grecula JC, Wang JZ, Jia G, et al. Characterizing at-Risk Voxels by Using Perfusion Magnetic Resonance Imaging for Cervical Cancer during Radiotherapy. *Journal of cancer science & therapy*. 2012; 4:254–9. [PubMed: 23638244]
 20. Huang Z, Yuh KA, Lo SS, Grecula JC, Sammet S, Sammet CL, et al. Validation of optimal DCE-MRI perfusion threshold to classify at-risk tumor imaging voxels in heterogeneous cervical cancer for outcome prediction. *Magnetic resonance imaging*. 2014; 32:1198–205. [PubMed: 25179141]
 21. Narayan K, Lin MY. Staging for cervix cancer: Role of radiology, surgery and clinical assessment. *Best practice & research Clinical obstetrics & gynaecology*. 2015; 29:833–44. [PubMed: 25898789]
 22. Yuh WT, Mayr NA, Jarjoura D, Wu D, Grecula JC, Lo SS, et al. Predicting control of primary tumor and survival by DCE MRI during early therapy in cervical cancer. *Investigative radiology*. 2009; 44:343–50. [PubMed: 19661844]
 23. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. *J Med Imaging (Bellingham)*. 2015; 2:041002. [PubMed: 26251842]
 24. Kidd EA, Thomas M, Siegel BA, Dehdashti F, Grigsby PW. Changes in cervical cancer FDG uptake during chemoradiation and association with response. *International journal of radiation oncology, biology, physics*. 2013; 85:116–22.
 25. Skillings JH, Mack GA. On the Use of a Friedman-Type Statistic in Balanced and Un-Balanced Block-Designs. *Technometrics*. 1981; 23:171–7.
 26. Goeman JJ, Solari A. Multiple hypothesis testing in genomics. *Statistics in medicine*. 2014; 33:1946–78. [PubMed: 24399688]
 27. Mayr NA, Wang JZ, Lo SS, Zhang D, Grecula JC, Lu L, et al. Translating response during therapy into ultimate treatment outcome: a personalized 4-dimensional MRI tumor volumetric regression approach in cervical cancer. *International journal of radiation oncology, biology, physics*. 2010; 76:719–27.
 28. Zhao S, Kuge Y, Mochizuki T. 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; 46:675–82. [PubMed: 15809491]
 29. van Baardwijk A, Bosmans G, van Suylen RJ, van Kroonenburgh M, Hochstenbag M, Geskes G, et al. Correlation of intra-tumour heterogeneity on 18F-FDG PET with pathologic features in non-small cell lung cancer: a feasibility study. *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology*. 2008; 87:55–8. [PubMed: 18328584]
 30. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Scientific reports*. 2015; 5:11044. [PubMed: 26251068]
 31. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ breast cancer*. 2016;2.
 32. Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, et al. Deciphering Genomic Underpinnings of Quantitative MRI-based Radiomic Phenotypes of Invasive Breast Carcinoma. *Scientific reports*. 2015; 5:17787. [PubMed: 26639025]
 33. Liu Y, Ye Z, Sun H, Bai R. Grading of uterine cervical cancer by using the ADC difference value and its correlation with microvascular density and vascular endothelial growth factor. *European radiology*. 2013; 23:757–65. [PubMed: 22955960]
 34. Mannelli L, Patterson AJ, Zahra M, Priest AN, Graves MJ, Lomas DJ, et al. Evaluation of nonenhancing tumor fraction assessed by dynamic contrast-enhanced MRI subtraction as a predictor of decrease in tumor volume in response to chemoradiotherapy in advanced cervical cancer. *AJR American journal of roentgenology*. 2010; 195:524–7. [PubMed: 20651213]
 35. Zahra MA, Tan LT, Priest AN, Graves MJ, Arends M, Crawford RA, et al. Semiquantitative and quantitative dynamic contrast-enhanced magnetic resonance imaging measurements predict

radiation response in cervix cancer. International journal of radiation oncology, biology, physics. 2009; 74:766–73.

36. Mayr NA, Huang Z, Wang JZ, Lo SS, Fan JM, Grecula JC, 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. International journal of radiation oncology, biology, physics. 2012; 83:972–9.
37. Lyng H, Vorren AO, Sundfor K, Taksdal I, Lien HH, Kaalhus O, et al. Intra- and intertumor heterogeneity in blood perfusion of human cervical cancer before treatment and after radiotherapy. Int J Cancer. 2001; 96:182–90. [PubMed: 11410887]
38. Olsen JR, Esthappan J, DeWees T, Narra VR, Dehdashti F, Siegel BA, et al. Tumor volume and subvolume concordance between FDG-PET/CT and diffusion-weighted MRI for squamous cell carcinoma of the cervix. Journal of magnetic resonance imaging: JMRI. 2013; 37:431–4. [PubMed: 23023940]
39. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. Scientific reports. 2015; 5:11075. [PubMed: 26242464]
40. Panth KM, Leijenaar RT, Carvalho S, Lieuwes NG, Yaromina A, Dubois L, et al. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology. 2015; 116:462–6. [PubMed: 26163091]

Figure 1.

MRI-PET fusion, DCE MR plateau (post-contrast), and DW MR ADC images for a representative patient at three different time points: prior to radiation therapy, early during RT (2 weeks), and midway during RT (5 weeks). Gross tumor was delineated at each time point (magenta contour), from which voxel distribution histogram features were extracted. Tumor voxels suffering from bladder artifact on FDG PET or distortion artifacts on DW MRI were excluded by threshold from calculation of SUV and ADC histogram features, respectively. Note the variable heterogeneity in image intensity between modalities (FDG PET, DCE MRI, DW MRI) and variable changes during therapy.

Figure 2.

Intensity histograms of FDG-PET SUV, DCE-MRI SI ratio, and DW-MRI ADC at time points prior to, 2 weeks during, and 5 weeks during external beam radiation therapy for the same patient as in Figure 1, from which quantitative features of heterogeneity can be extracted. Note the variable histogram shapes and distributions between modalities and across time points. This patient is characterized by increasing FDG-PET SUV skewness, decreasing DCE-MRI SI ratio skewness, and decreasing DW-MRI ADC skewness with therapy.

Figure 3.

Statistical box-whisker plots of tumor radiomic heterogeneity features of all patients compared between time points for each modality. (A) mean, (B) quantiles, (C) coefficient of variation, (D) skewness, (E) kurtosis of FDG PET standardized uptake value (SUV), DCE MRI plateau signal intensity ratio (SI), and DW MRI apparent diffusion coefficient (ADC) tumor voxel distributions. Mean (open marker), median (line), inter-quartile range (box) and range (whisker) across the patient cohort are shown for each imaging modality. Distributions of radiomic features were compared between imaging time points for each modality using the Skillings-Mack non-parametric test. P values were adjusted for multiple comparisons: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 4.

Statistical box-whisker plots of tumor radiomic heterogeneity features of all patients compared between modalities at each time point. (A) coefficient of variation, (B) skewness, (C) kurtosis of FDG PET standardized uptake value (SUV), DCE MRI plateau signal intensity ratio (SI), and DW MRI apparent diffusion coefficient (ADC) tumor voxel distributions. Mean (open marker), median (line), inter-quartile range (box) and range (whisker) across the patient cohort are shown for each time point. Distributions of radiomic features at the same time point were compared between modalities using the Skillings-Mack non-parametric test. P values were adjusted for multiple comparisons: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Candidate intensity histogram (IH) features of non-responders to radiotherapy as defined by 1-month post-treatment tumor volume regression ($< 10\%$ residual volume from baseline). Imaging biomarkers are sorted by decreasing AUC. Only those with $p < 0.05$ (not adjusted for multiple comparisons) are shown out of 75 IH features.

Table 1

IH Features	Responder [*]		P-value [†]	AUC
	Yes (N=13)	No (N=7)		
2wk ADC _{skewness} [‡]	-0.9 ± 0.7	-0.2 ± 0.3	0.010	0.86
Pre FDG SUV _{mean}	8.1 ± 2.4	5.7 ± 1.4	0.024	0.81
Pre FDG SUV _{max}	20.3 ± 5.4	14.2 ± 3.4	0.024	0.81
5wk ADC _{mean} [‡]	1.2 ± 0.2	1.4 ± 0.1	0.028	0.81
% 5wk FDG SUV _{mean}	-54.9 ± 18.4	-35.0 ± 16.7	0.037	0.79
5wk FDG SUV _{skewness}	0.6 ± 0.6	-0.2 ± 1.0	0.037	0.79

AUC = area under the receiver operating characteristic curve; SD = standard deviation;

^{*} Values are mean $\pm$ SD;

[†] Mann-Whitney test, no adjustment for multiple comparisons;

[‡] Excluding one responder and one non-responder with missing DCE measurements at the baseline or 2 week scan.