Original Article

Magnetic Resonance Image Radiomic Reproducibility: The Impact of Preprocessing on Extracted Features from Gross and High-Risk Clinical Tumor Volumes in Cervical Cancer Patients before Brachytherapy

Abstract

Background: Radiomic feature reproducibility assessment is critical in radiomics-based image biomarker discovery. This study aims to evaluate the impact of preprocessing parameters on the reproducibility of magnetic resonance image (MRI) radiomic features extracted from gross tumor volume (GTV) and high-risk clinical tumor volume (HR-CTV) in cervical cancer (CC) patients. **Methods:** This study included 99 patients with pathologically confirmed cervical cancer who underwent an MRI prior to receiving brachytherapy. The GTV and HR-CTV were delineated on T2-weighted MRI and inputted into 3D Slicer for radiomic analysis. Before feature extraction, all images were preprocessed to a combination of several parameters of Laplacian of Gaussian (1 and 2), resampling (0.5 and 1), and bin width (5, 10, 25, and 50). The reproducibility of radiomic features was analyzed using the intra-class correlation coefficient (ICC). **Results:** Almost all shapes and first-order features had ICC values > 0.95. Most second-order texture features were not reproducible (ICC < 0.95) in GTV and HR-CTV. Furthermore, 20% of all neighboring gray-tone difference matrix texture features had ICC > 0.90 in both GTV and HR-CTV. **Conclusion:** The results presented here showed that MRI radiomic features are vulnerable to changes in preprocessing, and this issue must be understood and applied before any clinical decision-making. Features with ICC > 0.90 were considered the most reproducible features. Shape and first-order radiomic features were the most reproducible features in both GTV and HR-CTV. Our results also showed that GTV and HR-CTV radiomic features had similar changes against preprocessing sets.

Keywords: *Brachytherapy, gross tumor volume, high-risk clinical tumor volume, MRI, radiomics, reproducibility*

Submitted: 30-Aug-2022 **Revised:** 09-Nov-2022 **Accepted:** 14-Mar-2023 **Published:** 06-Aug-2024

Mahdi Sadeghi[1,2], Neda Abdalvand[1], Seied Rabi Mahdavi[1,3], Hamid Abdollahi[4,5], Younes Qasempour[4], Fatemeh Mohammadian[6], Mohammad Javad Tahmasebi Birgani[6,7], Khadijeh Hosseini[6], Maryam Hazbavi[6]

[1]Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, [2]Finetech in Medicine Research Center, Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, [3]Radiation Biology Research Center, Iran University of Medical Sciences, Tehran, [4]Student Research Committee, Department of Radiology Technology, Faculty of Allied Medicine, Kerman University of Medical Sciences, Kerman, [6]Department of Radiation Oncology, Golestan Hospital, Ahvaz Jundishapour University of Medical Sciences, [7]Department of Medical Physics, School of Medicine, Ahvaz Jundishapour University of Medical Sciences, Ahvaz, Iran, [5]Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada

Address for correspondence:
Dr. Neda Abdalvand, Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran.
E-mail: abdalvandneda@gmail.com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_57_22

Quick Response Code:

Introduction

Cervical cancer (CC) is one of the most common cancers among women globally, estimated at 500,000 new cases/year.[1] The standard CC stage IB2-IVA treatment is concurrent external beam chemoradiation followed by brachytherapy (BT). Despite significant improvement of survival gained by established treatment, about 20% of patients experienced treatment failure.[2,3] Because of differences between patients' tumor architectures, personalized treatment based on tumor texture might effectively reduce the failure.[4]

Due to its superior soft-tissue resolution, magnetic resonance imaging (MRI) plays an essential role in the management of CC patients at several steps of diagnosis, staging, treatment planning (guidance), response assessment, and follow-up.[3] It is also feasible for personalizing external beam radiotherapy and BT on CC patients.[5-7] In recent decades, image-guided BT (IGBT) benefited from MRI. For example, various guidelines recommended sequential MRI at different times, including before external beam radiation therapy (EBRT) and BT and the treatment planning procedure. It has been recently highlighted that MRI before the first session of BT could be acquired to evaluate remaining tumor extension after EBRT, BT target volume estimation, and uterine canal anatomy determination for preinsertion assessments.[8,9]

Radiomics is a quantitative image analysis used noninvasively to determine tumor

How to cite this article: Sadeghi M, Abdalvand N, Mahdavi SR, Abdollahi H, Qasempour Y, Mohammadian F, *et al.* Magnetic resonance image radiomic reproducibility: The impact of preprocessing on extracted features from gross and high-risk clinical tumor volumes in cervical cancer patients before brachytherapy. J Med Sign Sens 2024;14:23.

heterogeneity by extracting a wide range of features. It is clearly shown that radiomic features can be helpful in tissue characterization, diagnosis, prognosis, response assessment, and prediction, and personalized medicine can be easily adapted through this approach.[10] Radiomic features derived from MRI of CC patients showed excellent performance in radiation therapy decision-making, especially for IGBT individualized-based treatment management.[11-13] Since the microenvironment of tumor and cervix tissue changes significantly during EBRT, in the setting of adaptive IGBT for CC, assessment of remaining gross tumor volume (GTV) at the end of EBRT and before starting BT is an important task.[8,14,15]

In biomarker discovery approaches, for example, developing predictive models with high reliability and accuracy, it is critical to find and use the most reproducible and repeatable markers. For radiomics-based image biomarker discovery, it is clearly shown that radiomic features are vulnerable to image acquisition, processing, reconstruction, segmentation, and analysis changes. This issue was shown for radiomic features in many cancers and almost all imaging modalities.[16-20] MRI radiomic reproducibility analysis studies have concluded that MRI radiomic features vary over image acquisition, segmentation, and preprocessing.[21-23] Preprocessing is a critical step in radiomic analysis. It has been suggested that some preprocessing approaches such as gray-level discretization, resampling, and applying some filter-based image processing such as Laplacian of Gaussian (LoG) have significant impacts on images and consequently on radiomic features.[23,24]

The present study aims to assess the reproducibility of radiomic features derived from manually segmented cervix tumors in patients who underwent T2-weighted MRI before BT. Then, the influence of different feature extraction parameters is evaluated, and the optimum setting is introduced.

Materials and Methods

Patients

This retrospective study was performed at the BT ward of the Radiotherapy Department of Golestan Hospital in Ahvaz.

Inclusion and exclusion criteria

Patients with histopathologically proven CC, stage IB-IVA, were enrolled in this research. All patients received whole-pelvic EBRT with a 45–52.2 Gy dose before BT. All subjects underwent MRI to evaluate tumor response to EBRT before the initialization of BT. We included only patients who had managed to receive definitive chemoradiation, and patients in adjuvant or neoadjuvant chemoradiation settings (i.e., metastatic or recurrent patients or those who had done surgery before chemoradiation) were excluded from the study. Patients with pre-BT tumor size <1 cm^3 or no visible tumors on T2-weighted MRI were excluded from the study. Patients with cardiac pacemaker were not eligible for MRI acquisition and were then excluded from the study.

Imaging

All MRI scans were acquired with a 1.5 Tesla MAGNETOM (Siemens, Erlangen, Germany) MRI scanner. Axial T2-weighted fast spin-echo MR scans were acquired. The information on image acquisition is shown in Table 1. Based on some reports, to better distinguish tumor volume, normal cervix, and vagina, aqueous gel was injected into the vagina.[25,26] No intravenous contrast agent was used. The images were acquired before the first session of BT without any applicator insertion to eliminate any effect of GYN (Gynecology) applicators on extracted features.

Segmentation

GTV and high-risk clinical tumor volume (HR-CTV) based on T2-weighted MRI images were manually delineated based on the American Brachytherapy Society guidelines and The Groupe Europeen de Curietherapie European Society for Radiotherapy and Oncology recommendations for MR-based target volume delineation by a 10-year experienced radiation oncologist with subspecialization at GYN-BT.[14,27] All regions in the cervix that had higher signal (dark) than normal cervix (gray) were included as GTV. GTV plus normal cervix was segmented as HR-CTV. Finally, three-dimensional volumes of GTV and HR-CTV were created. Furthermore, no gel is included in the segmentation. For 3D slicer, you can refer to: https://www.slicer.org/ In brief, 3D Slicer, is a free, open source software package for visualization and image analysis. Since it has a number of contributors, around the world, you can refer to its website. Based on the Wikipedia, it is developed by Harvard University, National Institutes of Health was used for delineations. Figure 1 shows the segmentation of a typical patient.

Image processing and feature extraction

The radiomic module of 3D Slicer software was used to feature extraction. 3D Slicer is an open-source and integrated platform for medical image computing.[28] Each radiomic feature has a specific value that can be

Table 1: Specifications of T2-weighted magnetic resonance image acquisition parameters

Parameter	Value
Slice thickness (mm)	4.5
Slice spacing (mm)	1.17
Time of echo (ms)	80
Time of repetition (ms)	3590
Matrix size	256×256
Number of excitations	1

All distances are based on mm and all times are in ms

obtained by applying its own mathematical analysis to medical images. Radiomic features and their equations are available at https://pyradiomics.readthedocs.io/en/latest/features.html. To find the most reproducible set of extraction parameters, we changed resampled voxel size, LoG kernel size, and bin width in various combinations with each other.

Then, we extracted features from 7 different preprocessing parameters that were obtained by combining these preprocessing parameters, as described in Table 2. These sets aimed to study the effect of preprocessing parameters separately and/or in combination. At set number 1, bin width changed (5, 10, 25, and 50) with fixed resampling size of 0.5 mm and LoG of 1. The experiments were conducted in four sets. In the first set, the bin width was changed to 5, 10, 25, and 50, while keeping the resampling size fixed at 0.5 mm and using a Laplacian of Gaussian (LoG) filter with a sigma value of 1. In the second set, the bin width was changed to 5, 10, 25, and 50, with a resampling size of 0.5 mm and a LoG filter with a sigma value of 2. In the third set, the bin width was again changed to 5, 10, 25, and 50, but with a fixed resampling size of 1 mm and a LoG filter with a sigma value of 1. Finally, in the fourth set, the bin width was changed to 5, 10, 25, and 50, with a resampling size of 1 mm and a LoG filter with a sigma value of 2.

Set numbers 5, 6, and 7 are more complex than other sets, and besides changing bin width (5, 10, 25, and 50), they have different resampling sizes (0.5 and 1) with or without different LoG (1 and 2).

Statistical analyses

To assess reproducibility, we applied the intra-class correlation coefficient (ICC), a principal reproducibility measure used in many previous radiomic studies.[29-31] R package "IRR" version 0.84.1 was used for ICC analysis. ICC can be calculated from the following equation:

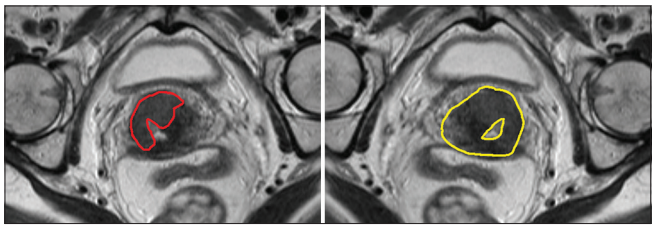

Figure 1: Segmentation approach on T2W images. Red and yellow segments show GTV and HR-CTV, respectively. GTV – Gross tumor volume; HR-CTV – High-risk clinical tumor volume

$$ICC = \frac{MS_R - MS_W}{MS_R + (K - 1)\,MS_W}$$

where MSR and MSW are mean squares for rows and residual sources of the variance, respectively, and k is the number of observers involved.

In this study, we obtained ICC numbers in all combined preprocessing states, and also, we divided them into the following categories: (1) ICC <50%, (2) 50% < ICC <80%, (3) 80% < ICC <90%, (4) 90% < ICC <95%, and (5) ICC >95%. Features with ICC >95% were introduced as the most reproducible categories.

To show the results, we used GraphPad Prism 8.0 software, GraphPad Software, Inc. San Diego, California, US and Excel (Office, 2016) for better data presentations in terms of heat maps and box plots.

Results

In this study, to better figure out the results, we categorize them into the following sections. First, patient data will be presented in terms of number, age, FIGO (International Federation of Gynecology and Obstetrics) stage, pathology, and EBRT dose. Second, ICC results will be described by using heatmaps for each feature set. We also show some ICC results in terms of percentage and mean ± Standard Deviation (SD) of ICC in different combinations of preprocessing parameters.

Patients

In this study, a total of 130 patients were initially enrolled. However, after a thorough assessment, five patients were excluded due to the presence of metastases, eight patients had undergone a supracervical hysterectomy, 13 patients did not undergo magnetic resonance (MR) imaging before brachytherapy (BT), and five patients were excluded because their Gross Tumor Volume (GTV) volume was less than or equal to 1 cm3. This left a total of 99 patients who were included in the final analysis. The characteristics of these patients are presented in detail in Table 3.

Radiomic features

We extracted a total of 107 T2-weighted MRI-based radiomic features quantifying (1) first-order statistics based on pixel gray-level histograms, 18 features; (2) shape metrics, 14 features; and (3) statistical features derived from texture matrices including gray-level co-occurrence

Table 2: Preprocessing sets

	1	2	3	4	5	6	7
Resampling	0.5	0.5	1	1	0.5	1	0.5,1
LoG	1	2	1	2	1.2	1.2	1.2
Bin width	5, 10, 25, 50	5, 10, 25, 50	5, 10, 25, 50	5, 10, 25, 50	5, 10, 25, 50	5, 10, 25, 50	5, 10, 25, 50

LoG – Laplacian of Gaussian

matrix (GLCM), gray-level size zone matrix (GLSZM), gray-level dependence matrix (GLDM), gray-level run-length matrix (GLRLM), and neighboring gray-tone difference matrix (NGTDM), 75 features. The features are detailed in Supplementary Table 1.

Intra-class correlation coefficient

The mean ± Std of ICC in different preprocessing states is shown in Table 4, and our results on different ICC categories for all feature/preprocessing sets are shown in Table 5. The results were presented based on the percentage in Table 5. Vertical values (1–5) represent five categories of ICC ranges which 5 is the most reproducible range and 1 is the worth one, and horizontal values (1–7) represent seven sets of preprocessing parameters used in this study.

Table 3: Patient demographic and tumor characteristics of patients

Demographic Characteristic	Number
Total number	99
Age range (median)	28–81 (53)
FIGO stage	
IB	1
IIA	16
IIB	57
IIIA	8
IIIB	12
IVA	5
Pathology	
SCC	60
Adenocarcinoma	30
Clear cell	9
EBRT dose range (median)	45–52.2 (47.92)

SCC – Squamous cell carcinoma; EBRT – External beam radiation therapy; FIGO – International Federation of Gynecology and Obstetrics

Our ICC result for GLDM features is shown in Figure 2. Except for dependence nonuniformity, Large dependence high gray-level emphasis, low gray-level emphasis, and small dependence low gray-level emphasis were found as radiomic textures with higher ICC values (range: 0.63–0.99) for both GTV and HR-CTV in all preprocessing sets. Most features have a wide range of ICC variations from 0.14 to 0.93.

The ICC results for GLRLM features, as shown in Figure 3, for both volumes and all preprocessing sets, including gray-level nonuniformity, run length nonuniformity, and short-run low gray-level emphasis, were found as features with the highest ICC (range: 0.68–0.96).

For GLSZM features, we showed the ICC results in Figure 4. As was depicted, gray-level nonuniformity and zone entropy were those features with the highest ICC (range: 0.63–0.92) for both volumes and all preprocessing sets.

For NGTDM features [Figure 5], only coarseness was found as a highly reproducible feature (ICC > 0.93) for both volumes and all preprocessing sets (range: 0.93–1).

Figure 6 also shows ICC percentages of 7 preprocessing sets for GTV and HR-CTV. For all features, preprocessing numbers 5, 6, and 7 belong to the more stable ones.

Our ICC results for both GTV and HR-CTV were shown as heat maps in Figures 2-5 and 7-9 based on the feature sets for all preprocessing sets.

In Figure 7, the ICC values for shape features are shown. As expected, almost all features, except the feature of sphericity for HR-CTV in preprocessing set number 4, were found as most reproducible (ICC >95%).

The results for first-order features are depicted in Figure 8. Two features, including entropy and uniformity, had ICC

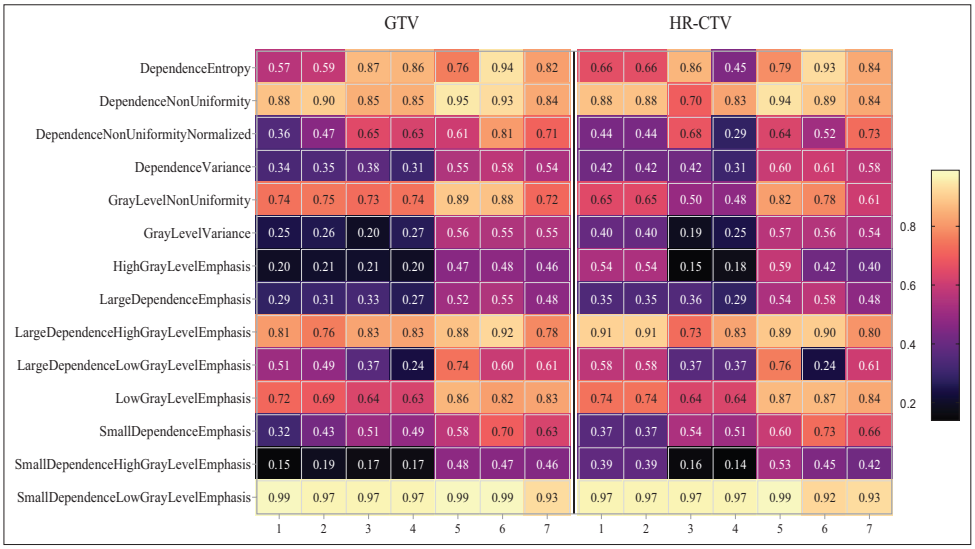

Figure 2: GLDM radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to GLDM feature names. GLDM – Gray-level dependence matrix

Table 4: Mean±standard deviation of intra-class correlation coefficient in different combinations of preprocessing parameters

Feature group	Volume of interest	Mean±SD of ICC in different preprocessing states						
		1	2	3	4	5	6	7
Shape	GTV	1	1	1	1	1	1	1
	HR-CTV	1	1	1	0.99±0.02	1	1	1
1st order	GTV	0.92±0.23	0.92±0.22	0.92±0.23	0.92±0.23	0.95±0.16	0.94±0.16	0.94±0.16
	HR-CTV	0.92±0.21	0.92±0.21	0.91±0.23	0.91±0.22	0.95±0.15	0.92±0.21	0.94±0.16
GLCM	GTV	0.41±0.30	0.43±0.28	0.45±0.32	0.45±0.30	0.62±0.21	0.65±0.20	0.58±0.18
	HR-CTV	0.49±0.25	0.49±0.25	0.42±0.32	0.42±0.31	0.63±0.19	0.56±0.20	0.58±0.18
GLDM	GTV	0.51±0.28	0.53±0.26	0.55±0.28	0.53±0.29	0.70±0.18	0.73±0.19	0.67±0.16
	HR-CTV	0.59±0.21	0.59±0.21	0.52±0.26	0.47±0.26	0.72±0.16	0.67±0.22	0.66±0.17
GLRLM	GTV	0.42±0.23	0.42±0.23	0.42±0.23	0.40±0.23	0.63±0.17	0.63±0.16	0.58±0.15
	HR-CTV	0.52±0.18	0.52±0.18	0.39±0.20	0.41±0.21	0.66±0.14	0.61±0.15	0.57±0.16
GLSZM	GTV	0.37±0.18	0.39±0.18	0.41±0.21	0.42±0.20	0.64±0.14	0.65±0.19	0.54±0.17
	HR-CTV	0.48±0.16	0.48±0.16	0.43±0.19	0.43±0.19	0.65±0.12	0.64±0.14	0.56±0.15
NGTDM	GTV	0.51±0.32	0.52±0.32	0.48±0.30	0.54±0.30	0.74±0.19	0.75±0.18	0.62±0.22
	HR-CTV	0.55±0.27	0.55±0.27	0.48±0.31	0.52±0.29	0.74±0.19	0.75±0.17	0.60±0.24

ICC – Intra-class correlation coefficient; GLCM – Gray-level co-occurrence matrix; GLSZM – Gray-level size zone matrix; GLDM – Gray-level dependence matrix; GLRLM – Gray-level run-length matrix; NGTDM – Neighboring gray-tone difference matrix; GTV – Gross tumor volume; HR-CTV – High-risk clinical tumor volume; SD – Standard deviation

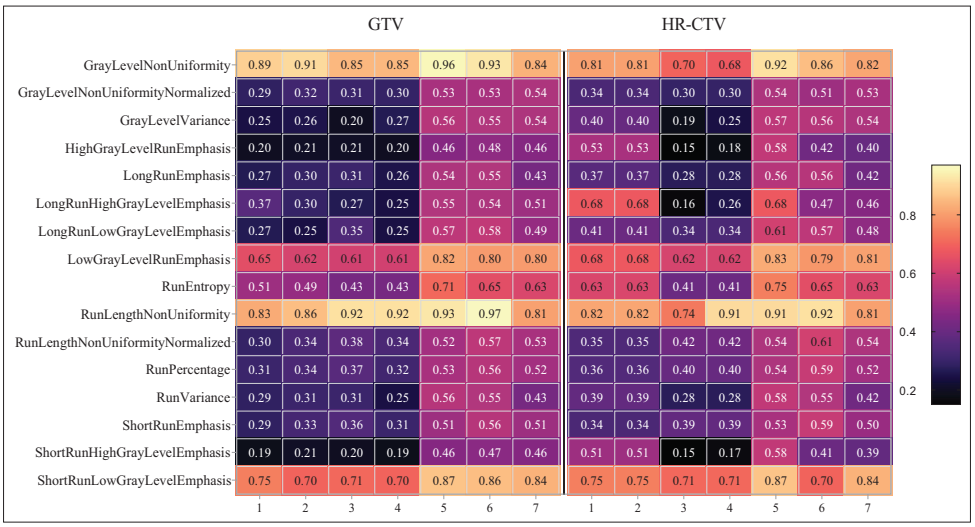

Figure 3: GLRLM radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to GLRLM feature names. GLRLM – Gray-level run-length matrix

below 0.54 (range: 0.24–0.53) for all preprocessing sets. The energy feature also had a range of ICC from 0.78 to 1. All remaining features had ICC = 1.

For GLCM texture features, in all preprocessing sets, as shown in Figure 9, correlation, Idn, Idmn, Imc1, Imc2, and Maximal Correlation Coefficient (MCC) were found as features with higher reproducibility (range: 0.6–1, most of them >0.9) than other GLCM features. The range of ICC for the remaining features was found as 0.1–0.7.

Discussion

Preprocessing is a critical step in radiomic analysis. It is clearly shown that some preprocessing algorithms can reduce noise, improve image quality, and harmonize/normalize images.[32,33] Furthermore, it is identified that radiomic features are vulnerable to applying preprocessing algorithms. This issue is studied for almost all imaging modalities, and other ongoing evaluations are in progress, but there is still no comprehensive conclusion.[34,35] Some studies used statistical analyses such as the concordance correlation coefficient, *t*-test, or Spearman correlation. Still, there is no established reference for the gold standard test.[36,37] Hence, we used the ICC metric accepted for reproducibility and repeatability in different studies of T2-weighted MRI.[5,35,38] For gray-level discretization, Duron *et al*.[24] showed that MRI texture radiomic features significantly change by bin size and bin number.[39] On the impact of resampling on radiomic features, Park *et al*. showed the effect of pixel

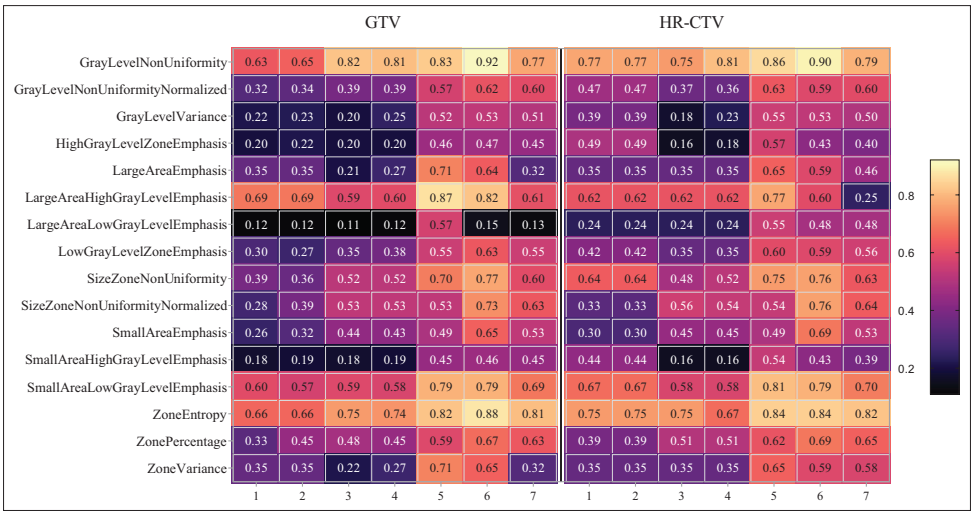

Figure 4: GLSZM radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to GLSZM feature names. GLSZM – Gray-level size zone matrix

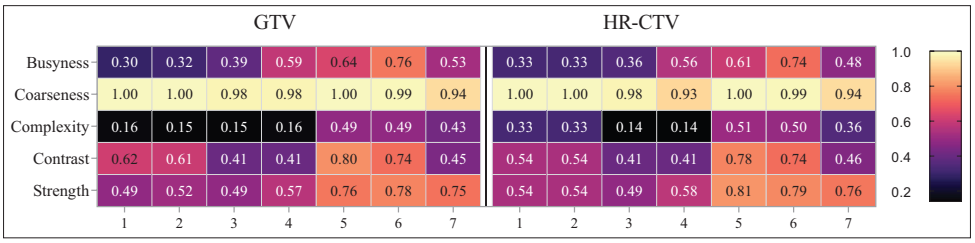

Figure 5: NGTDM radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to NGTDM feature names. NGTDM – Neighboring gray-tone difference matrix

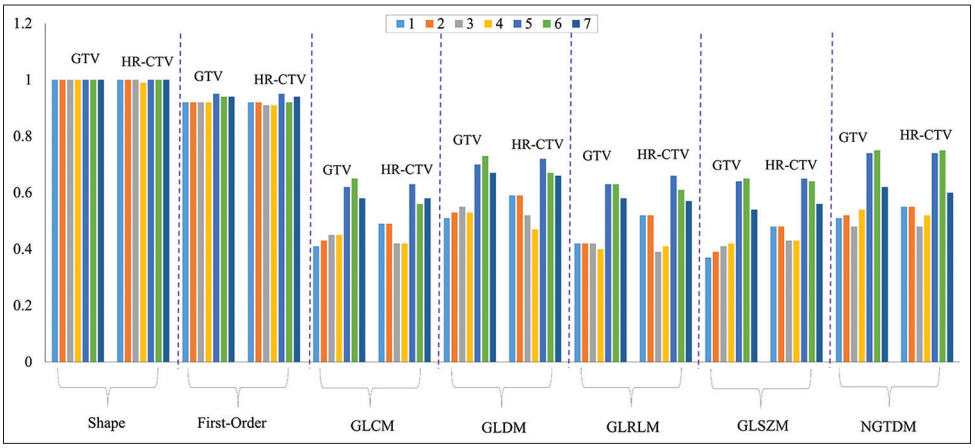

Figure 6: ICC percentage in all radiomic feature sets. The vertical axis is ICC values and the horizontal axis is feature groups. ICC – Intra-class correlation coefficient

size resampling and interpolation on MRI radiomic features in patients with CC.[23] They observed that many radiomic features were robust, while others had several changes. However, preprocessing is vital in radiomic analysis because of image normalization and noise reduction.

In the present inquiry, we assessed and compared the consequences of applying different preprocessing parameters, including bin width, LoG filter, and voxel size resampling on radiomic features extracted from MR images of two target volumes that are globally established for BT treatment planning in CC patients, including GTV and HR-CTV.

To our knowledge, there is no published study on HR-CTV radiomic features, and most papers on the reproducibility of GTV radiomic features rely on images with a BT applicator in place. Although radiomics on pre-BT images can be beneficial for predictive modeling in BT procedures, there is a lack of evaluation on this time point imaging and volumes. In this era, we focused on using T2-weighted MR images

Table 5: Intra-class correlation coefficient percentage in all features and preprocessing sets

		Different combinations of preprocessing															
Feature set-volume	**ICC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Feature set-volume**	**ICC**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Shape									First order								
GTV	1	0	0	0	0	0	0	0	GTV	1	11	11	11	11	6	6	0
	2	0	0	0	0	0	0	0		2	0	0	0	0	6	6	11
	3	0	0	0	0	0	0	0		3	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0		4	0	0	0	0	0	0	6
	5	100	100	100	100	100	100	100		5	89	89	89	89	89	89	89
HR-CTV	1	0	0	0	0	0	0	0	HR-CTV	1	11	11	11	11	0	11	6
	2	0	0	0	0	0	0	0		2	0	0	6	0	11	0	6
	3	0	0	0	0	0	0	0		3	0	0	0	6	0	0	6
	4	0	0	0	7	0	0	0		4	0	0	0	0	0	0	0
	5	100	100	100	93	100	100	100		5	89	89	83	83	89	89	82
GLDM									GLRLM								
GTV	1	50	36	43	14	14	21	43	GTV	1	69	75	75	75	13	13	25
	2	36	36	29	50	29	50	36		2	19	13	13	13	63	75	44
	3	7	21	21	21	29	21	7		3	13	6	6	6	13	13	25
	4	0	0	0	0	0	7	0		4	0	0	6	6	6	6	0
	5	7	7	7	14	14	0	14		5	0	6	0	0	6	6	0
HR-CTV	1	43	43	21	50	0	21	21	HR-CTV	1	50	50	50	63	75	19	38
	2	36	36	57	29	64	43	36		2	38	38	50	31	13	69	38
	3	7	7	14	14	21	21	29		3	13	13	0	0	6	6	25
	4	0	0	0	0	7	14	7		4	0	0	0	0	6	6	0
	5	14	14	7	7	7	0	0		5	0	0	0	6	6	0	0
GLSZM									NGTDM								
GTV	1	63	63	50	50	13	13	25	GTV	1	40	20	60	20	20	20	40
	2	38	38	44	44	56	69	69		2	40	60	20	60	40	60	40
	3	0	0	6	6	25	13	6		3	0	0	0	0	20	0	0
	4	0	0	0	0	0	6	0		4	0	0	0	0	0	0	20
	5	0	0	0	0	0	0	0		5	20	20	20	20	20	20	0
HR-CTV	1	63	63	44	44	17	19	31	HR-CTV	1	40	40	60	20	0	0	60
	2	38	38	56	56	75	69	56		2	40	40	20	60	60	80	20
	3	0	0	0	6	19	13	13		3	0	0	0	0	20	0	0
	4	0	0	0	0	0	0	0		4	0	0	0	20	0	0	20
	5	0	0	0	0	0	0	0		5	20	20	20	0	20	20	0
GLCM																	
GTV	1	63	63	63	79	25	13	33									
	2	17	17	17	8	50	58	46									
	3	4	4	4	4	4	4	21									
	4	4	4	0	4	0	4	0									
	5	13	13	17	17	21	21	0									
HR-CTV	1	58	58	46	58	4	29	29									
	2	21	21	29	17	71	54	46									
	3	4	4	8	13	13	4	25									
	4	13	13	4	0	4	8	0									
	5	4	4	21	13	17	4	0									

ICC values from 1 to 5 are scales for feature reproducibility (1 as the lowest and 5 as the most reproducible features). The values under set numbers 1–7 show the percent of features that are located in 1–5 ICC categories. ICC – Intra-class correlation coefficient; GLCM – Gray-level co-occurrence matrix; GLSZM – Gray-level size zone matrix; GLDM – Gray-level dependence matrix; GLRLM – Gray-level run-length matrix; NGTDM – Neighboring gray-tone difference matrix; GTV – Gross tumor volume; HR-CTV – High-risk clinical tumor volume

without BT applicators in place at a specific time point before BT initialization. We evaluated the reproducibility of HR-CTV and GTV radiomic features of CC patients in different arrangements of preprocessing parameters.

This study applied a combination of preprocessing parameters standard in image processing and radiomic analysis studies. Based on our findings, shape and first-order features, except for entropy and

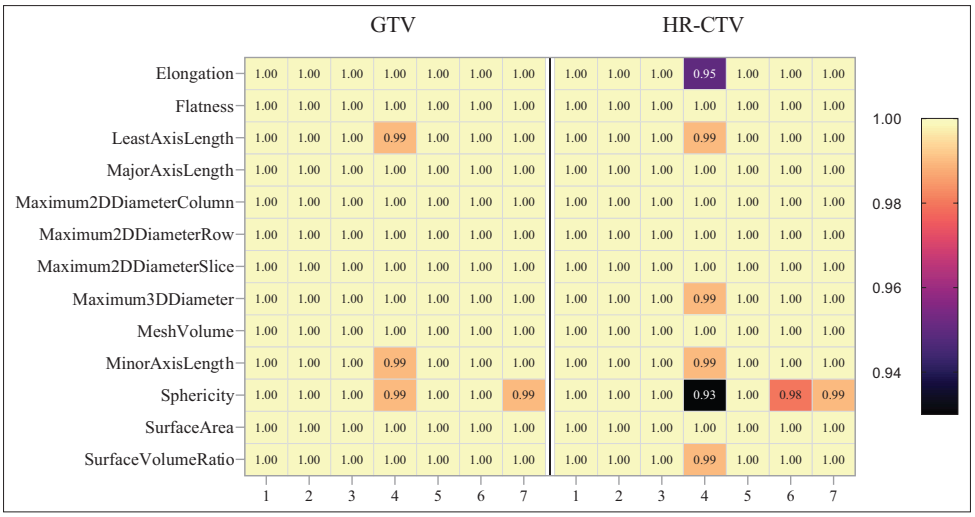

Figure 7: Shape radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to shape feature names

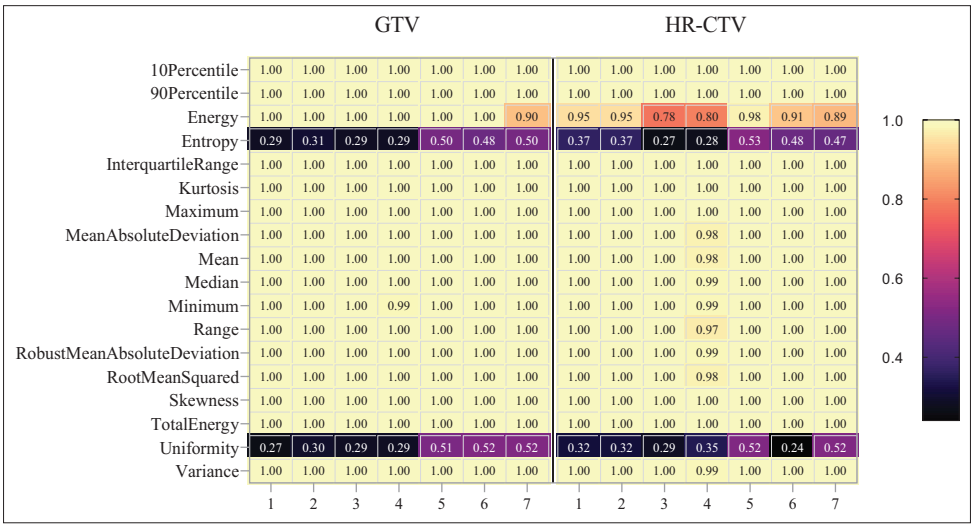

Figure 8: First-order radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to 1st order feature names

uniformity (0.27<ICC<0.53), are the most robust features in T2-weighted MRI after EBRT of CC patients which had the highest reproducibility over changes in all preprocessing settings. No specific trends were observed between ICC of both volumes for entropy, uniformity, and other shape and first-order features. In some preprocessing sets, ICC of GTV is higher than HR-CTV, and for other sets, ICC of HR-CTV is more. Since post-EBRT radiomic features extracted from MRI could be used for BT response analysis in radiomic predictive modeling, it is critical to find the most robust radiomic features.[40,41]

Furthermore, studies have indicated that radiotherapy-induced radiomic feature changes are feasible biomarkers for assessing tumor response or recurrence, normal tissue toxicity, and differentiation between tumor and normal tissue responses.[10,42] In this era, robust features over changing issues such as preprocessing would be worthful. This study proposes shape and first-order features as the most robust post-EBRT radiomic features.

On the other hand, most intensity-based and GLCM, most neighborhood gray-tone difference (NGTD), most GLSZM, and most GLRLM features had a range of mean ICC from 0.41 to 0.75. Parameter set numbers 5, 6, and 7, which belong to the most complex combination of parameters to each other, have higher reproducibility (0.62<ICC<0.75). The preprocessing number 6 has the most ICC value for all features, and it occurred when resampling is one and other two parameters varied. One can conclude that features are more robust with a resampling size of 1 mm.

Based on the results obtained from this cohort, it was obtained that there are no differences between the reproducibility of GTV and HR-CTV radiomic features. HR-CTV and GTV have different tissue inhomogeneity structures and amounts; GTV is more heterogeneous than HR-CTV. Results of our

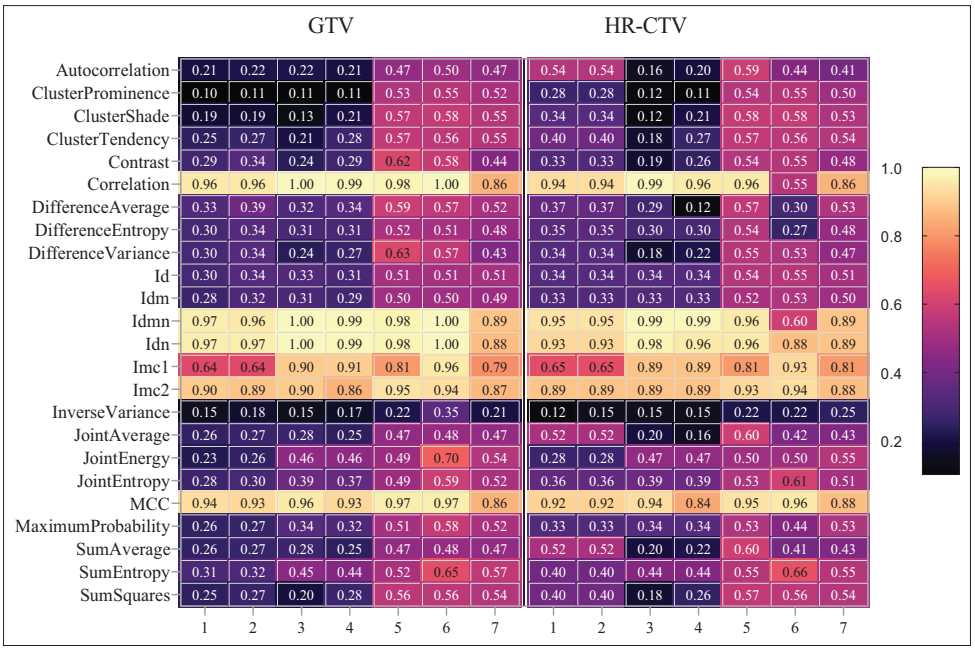

Figure 9: GLCM radiomic feature reproducibility. In the horizontal axis, 1–7 refer to different arrangements of preprocessing parameters and vertical names related to GLCM feature names. GLCM – Gray-level co-occurrence matrix

ICC computations propose that different preprocessing parameter sets could have similar consequences on the robustness of radiomic features extracted from different tissues with various levels of heterogeneity. Our study assessed two BT volumes, including GTV and HR-CTV. We aimed to understand how microscopic disease's radiomic features in HR-CTV respond to preprocessing settings compared to GTV. Although we did not perform statistical analysis, Based on visual observations such as heat maps and Figure 6, it appears that there are no noticeable variations in the reproducibility of GTV and HR CTV features. Although statistical analysis was not conducted, the results suggest that these features exhibit similar patterns. However, the ICC values have some differences in the two volumes in some preprocessing sets. For example, in sets, number 1 and 2 ICCs related to GTV are lower than HR-CTV, and set numbers 3 and 4 are reversed.

One can find that changing the LoG filter, bin width, or resampling size cannot produce more stable features. In all different arrangements of these parameters, all features have a similar power of reproducibility. This finding could be helpful for other studies in the future for preprocessing parameter selection.

Finally, this study has some limitations; first, the sample size is small, and future studies are needed to confirm our results; second, the effect of segmentation on radiomic features was not assessed as it might be the primary bias in radiomic studies.

Conclusion

The results presented here showed that MRI radiomic features are vulnerable to changes in preprocessing, and this issue must be understood and applied before any clinical decision-making. Our results also showed that both GTV and HR-CTV radiomic features had similar changes against preprocessing sets.

Ethical approval

All procedures performed in studies involving human participants were under the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent

Due to the retrospective nature of this study, there was no need for informed consent.

Consent for publication

Consent for publication was obtained for every person's data included in the study.

Data availability

The data and implementation materials.

Financial support and sponsorship

This research was supported by grant no. 16525 from the School of Medicine, Iran University of Medical Sciences.

Conflicts of interest

There are no conflicts of interest.

References

1. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, *et al.* Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis. Lancet Glob Health 2020;8:e191-203.

2. Chemoradiotherapy for Cervical Cancer Meta-Analysis Collaboration. Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: A systematic review and meta-analysis of individual patient data from 18 randomized trials. J Clin Oncol 2008;26:5802-12.
3. Naga Ch P, Gurram L, Chopra S, Mahantshetty U. The management of locally advanced cervical cancer. Curr Opin Oncol 2018;30:323-9.
4. Volm M, Efferth T. Prediction of cancer drug resistance and implications for personalized medicine. Front Oncol 2015;5:282.
5. Dimopoulos JC, Schard G, Berger D, Lang S, Goldner G, Helbich T, *et al.* Systematic evaluation of MRI findings in different stages of treatment of cervical cancer: Potential of MRI on delineation of target, pathoanatomic structures, and organs at risk. Int J Radiat Oncol Biol Phys 2006;64:1380-8.
6. Wachter-Gerstner N, Wachter S, Reinstadler E, Fellner C, Knocke TH, Pötter R. The impact of sectional imaging on dose escalation in endocavitary HDR-brachytherapy of cervical cancer: Results of a prospective comparative trial. Radiother Oncol 2003;68:51-9.
7. Schernberg A, Kumar T, Achkar S, Espenel S, Bockel S, Majer M, *et al.* Incorporating magnetic resonance imaging (MRI) based radiation therapy response prediction into clinical practice for locally advanced cervical cancer patients. Semin Radiat Oncol 2020;30:291-9.
8. Dimopoulos JC, Petrow P, Tanderup K, Petric P, Berger D, Kirisits C, *et al.* Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (IV): Basic principles and parameters for MR imaging within the frame of image based adaptive cervix cancer brachytherapy. Radiother Oncol 2012;103:113-22.
9. Pötter R, Federico M, Sturdza A, Fotina I, Hegazy N, Schmid M, *et al.* Value of magnetic resonance imaging without or with applicator in place for target definition in cervix cancer brachytherapy. Int J Radiat Oncol Biol Phys 2016;94:588-97.
10. Amini M, Nazari M, Shiri I, Hajianfar G, Deevband MR, Abdollahi H, *et al*. Multi-level multi-modality (PET and CT) fusion radiomics: Prognostic modeling for non-small cell lung carcinoma. Phys Med Biol 2021;66:205017.
11. Hua W, Xiao T, Jiang X, Liu Z, Wang M, Zheng H, *et al*. Lymph-vascular space invasion prediction in cervical cancer: Exploring radiomics and deep learning multilevel features of tumor and peritumor tissue on multiparametric MRI. Biomed Signal Process Control 2020;58:101869.
12. Bowen SR, Yuh WT, Hippe DS, Wu W, Partridge SC, Elias S, *et al.* Tumor radiomic heterogeneity: Multiparametric functional imaging to characterize variability and predict response following cervical cancer radiation therapy. J Magn Reson Imaging 2018;47:1388-96.
13. Lucia F, Visvikis D, Desseroit MC, Miranda O, Malhaire JP, Robin P, *et al.* Prediction of outcome using pretreatment (18) F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2018;45:768-86.
14. Haie-Meder C, Pötter R, Van Limbergen E, Briot E, De Brabandere M, Dimopoulos J, *et al.* Recommendations from Gynaecological (GYN) GEC-ESTRO Working Group (I): Concepts and terms in 3D image based 3D treatment planning in cervix cancer brachytherapy with emphasis on MRI assessment of GTV and CTV. Radiother Oncol 2005;74:235-45.
15. Pötter R, Haie-Meder C, Van Limbergen E, Barillot I, De Brabandere M, Dimopoulos J, *et al.* Recommendations from gynaecological (GYN) GEC ESTRO working group (II): Concepts and terms in 3D image-based treatment planning in cervix cancer brachytherapy-3D dose volume parameters and aspects of 3D image-based anatomy, radiation physics, radiobiology. Radiother Oncol 2006;78:67-77.
16. Hu P, Wang J, Zhong H, Zhou Z, Shen L, Hu W, *et al.* Reproducibility with repeat CT in radiomics study for rectal cancer. Oncotarget 2016;7:71440-6.
17. Liu Y, Zhang Y, Cheng R, Liu S, Qu F, Yin X, *et al.* Radiomics analysis of apparent diffusion coefficient in cervical cancer: A preliminary study on histological grade evaluation. J Magn Reson Imaging 2019;49:280-90.
18. Jethanandani A, Lin TA, Volpe S, Elhalawani H, Mohamed AS, Yang P, *et al.* Exploring applications of radiomics in magnetic resonance imaging of head and neck cancer: A systematic review. Front Oncol 2018;8:131.
19. Shiri I, Hajianfar G, Sohrabi A, Abdollahi H, *P* Shayesteh S, Geramifar P, *et al.* Repeatability of radiomic features in magnetic resonance imaging of glioblastoma: Test-retest and image registration analyses. Med Phys 2020;47:4265-80.
20. Qasempour Y, Mohammadi A, Rezaei M, Pouryazadanpanah P, Ziaddini F, Borbori A, *et al.* Radiographic texture reproducibility: The impact of different materials, their arrangement, and focal spot size. J Med Signals Sens 2020;10:275-85.
21. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, *et al.* Robust Radiomics feature quantification using semiautomatic volumetric segmentation. PLoS One 2014;9:e102107.
22. Traverso A, Kazmierski M, Welch ML, Weiss J, Fiset S, Foltz WD, *et al.* Sensitivity of radiomic features to inter-observer variability and image pre-processing in Apparent Diffusion Coefficient (ADC) maps of cervix cancer patients. Radiother Oncol 2020;143:88-94.
23. Park SH, Lim H, Bae BK, Hahm MH, Chong GO, Jeong SY, *et al.* Robustness of magnetic resonance radiomic features to pixel size resampling and interpolation in patients with cervical cancer. Cancer Imaging 2021;21:19.
24. Duron L, Balvay D, Vande Perre S, Bouchouicha A, Savatovsky J, Sadik JC, *et al.* Gray-level discretization impacts reproducible MRI radiomics texture features. PLoS One 2019;14:e0213459.
25. Engelaere C, Poncelet E, Durot C, Dohan A, Rousset P, Hoeffel C. Pelvic MRI: Is endovaginal or rectal filling needed? Korean J Radiol 2018;19:397-409.
26. Brown MA, Mattrey RF, Stamato S, Sirlin CB. MRI of the female pelvis using vaginal gel. AJR Am J Roentgenol 2005;185:1221-7.
27. Viswanathan AN, Beriwal S, De Los Santos JF, Demanes DJ, Gaffney D, Hansen J, *et al.* American Brachytherapy Society consensus guidelines for locally advanced carcinoma of the cervix. Part II: High-dose-rate brachytherapy. Brachytherapy 2012;11:47-52.
28. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 2012;30:1323-41.
29. Xue C, Yuan J, Lo GG, Chang AT, Poon DM, Wong OL, *et al.* Radiomics feature reliability assessed by intraclass correlation coefficient: A systematic review. Quant Imaging Med Surg 2021;11:4431-60.
30. Abdollahi H, Chin E, Clark HD, Hyde DE, Thomas S, Wu J, *et al*. Radiomics-guided radiation therapy: opportunities and challenges. Phys Med Biol. 2022;13(67).
31. Fiset S, Welch ML, Weiss J, Pintilie M, Conway JL, Milosevic M, *et al.* Repeatability and reproducibility of MRI-based radiomic features in cervical cancer. Radiother Oncol 2019;135:107-14.

32. Krajnc D, Papp L, Nakuz TS, Magometschnigg HF, Grahovac M, Spielvogel CP, *et al.* Breast tumor characterization using [(18) F] FDG-PET/CT imaging combined with data preprocessing and radiomics. Cancers (Basel) 2021;13:1249.
33. Rastegar S, Beigi J, Saeedi E, Shiri I, Qasempour Y, Rezaei M, *et al.* Radiographic image radiomics feature reproducibility: A preliminary study on the impact of field size. J Med Imaging Radiat Sci 2020;51:128-36.
34. Saeedi E, Dezhkam A, Beigi J, Rastegar S, Yousefi Z, Mehdipour LA, *et al.* Radiomic feature robustness and reproducibility in quantitative bone radiography: A study on radiologic parameter changes. J Clin Densitom 2019;22:203-13.
35. Scalco E, Belfatto A, Mastropietro A, Rancati T, Avuzzi B, Messina A, *et al.* T2w-MRI signal normalization affects radiomics features reproducibility. Med Phys 2020;47:1680-91.
36. Lecler A, Duron L, Balvay D, Savatovsky J, Bergès O, Zmuda M, *et al.* Combining multiple magnetic resonance imaging sequences provides independent reproducible radiomics features. Sci Rep 2019;9:2068.
37. Peng X, Yang S, Zhou L, Mei Y, Shi L, Zhang R, *et al.* Repeatability and reproducibility of computed tomography radiomics for pulmonary nodules: A multicenter phantom study. Invest Radiol 2022;57:242-53.
38. Abdollahi H, Tanha K, Mofid B, Razzaghdoust A, Saadipoor A, Khalafi L, *et al.* MRI radiomic analysis of IMRT-induced bladder wall changes in prostate cancer patients: A relationship with radiation dose and toxicity. J Med Imaging Radiat Sci 2019;50:252-60.
39. Abdollahi H, Mofid B, Shiri I, Razzaghdoust A, Saadipoor A, Mahdavi A, *et al.* Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer. Radiol Med 2019;124:555-67.
40. Fave X, Zhang L, Yang J, Mackin D, Balter P, Gomez D, *et al.* Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. Sci Rep 2017;7:588.
41. Nardone V, Reginelli A, Grassi R, Boldrini L, Vacca G, D'Ippolito E, *et al.* Delta radiomics: A systematic review. Radiol Med 2021;126:1571-83.
42. Cusumano D, Boldrini L, Yadav P, Casà C, Lee SL, Romano A, *et al.* Delta radiomics analysis for local control prediction in pancreatic cancer patients treated using magnetic resonance guided radiotherapy. Diagnostics (Basel) 2021;11:72.

Supplementary Table 1: Radiomic features
Shape
Elongation
Flatness
Least axis length
Major axis length
Maximum 2D diameter column
Maximum 2D diameter row
Maximum 2D diameter slice
Maximum 3D diameter
Mesh volume
Minor axis length
Sphericity
Surface area
Surface volume ratio
Voxel volume
GLDM
Dependence entropy
Dependence nonuniformity
Dependence nonuniformity normalized
Dependence variance
Gray-level nonuniformity
Gray-level variance
High gray-level emphasis
Large dependence emphasis
Large dependence high gray-level emphasis
Large dependence low gray-level emphasis
Low gray-level emphasis
Small dependence emphasis
Small dependence high gray-level emphasis
Small dependence low gray-level emphasis
NGTDM
Busyness
Coarseness
Complexity
Contrast
Strength
First order
10 percentiles
90 percentiles
Energy
Entropy
Interquartile range
Kurtosis
Maximum
Mean absolute deviation
Mean
Median
Minimum
Range
Robust mean absolute deviation
Root mean squared
Skewness
Total energy
Uniformity

Contd...

Supplementary Table 1: Contd...
First order
Variance
GLRLM
Gray-level nonuniformity
Gray-level nonuniformity normalized
Gray-level variance
High gray-level run emphasis
Long run emphasis
Long run high gray-level emphasis
Long run low gray-level emphasis
Low gray-level run emphasis
Run entropy
Run length nonuniformity
Run length nonuniformity normalized
Run percentage
Run variance
Short run emphasis
Short run high gray-level emphasis
Short run low gray-level emphasis
GLCM
Autocorrelation
Cluster prominence
Cluster shade
Cluster tendency
Contrast
Correlation
Difference average
Difference entropy
Difference variance
Id
Idm
Idmn
Idn
Imc1
Imc2
Inverse variance
Joint average
Joint energy
Joint entropy
MCC
Maximum probability
Sum average
Sum entropy
Sum squares
GLSZM
Gray-level nonuniformity
Gray-level nonuniformity normalized
Gray-level variance
High gray-level zone emphasis
Large area emphasis
Large area high gray-level emphasis
Large area low gray-level emphasis
Low gray-level zone emphasis
Size zone nonuniformity

Contd...

Supplementary Table 1: Contd...

GLSZM
Size zone nonuniformity normalized
Small area emphasis
Small area high gray-level emphasis
Small area low gray-level emphasis
Zone entropy
Zone percentage
Zone variance

GLCM – Gray-level co-occurrence matrix; GLSZM – Gray-level size zone matrix; GLDM – Gray-level dependence matrix; GLRLM – Gray-level run-length matrix; NGTDM – Neighboring gray-tone difference matrix; 2D – Two-dimensional; 3D – Three-dimensional; MCC – Maximal Correlation Coefficient