
Advanced image characterisation using radiomics on multi-parametric magnetic resonance images of cervical cancer patients across multiple time-points

Author:

Brijesh PATEL
9568975

Supervisors:

Dr Alan McWILLIAM
Dr Corinne JOHNSON-HART

School of Physics and Astronomy
and
The Christie NHS Foundation Trust

January 2019

This project was performed in collaboration with Harry HOWELL.

Abstract

In this project, a method was developed to establish how to extract a set of radiomics features from a sample of 8 patients with cervical cancer provided by the Christie NHS Foundation radiotherapy group. Dixon and T2-weighted magnetic resonance images (MRI) were obtained for up to 4 different visits across 2 time-points – at the start and end – for each of these patients. After a clinician had drawn on contours (delineations) for each image, a set of 105 features were obtained with their corresponding intra-class correlation coefficients (ICCs) then calculated for the primary tumour values, tumour nodes and rectal wall tissue. These ICCs were utilised to determine the repeatability and in order to produce a short list of stable features. These will then be explored further next semester in order to evaluate their clinical significance.

1. Introduction

MRI-guided radiation therapy (radiotherapy) is a common means to treat early-stage cervical cancer, where malignant tumours form in cervix tissues. Cervical cancer is the third-most common cancer amongst women worldwide and the nineteenth-most common in the UK, where its 1-year and 10-year survival rates are 82.8% and 63.0% respectively as of 2016. [1] Nevertheless, although the latter statistic has improved by 17% since the 1980s, this rate is lower than similar forms of cancer such as prostate (84%) and breast (78%), despite having the highest percentage of preventable cases (99.8%). [1] Factors such as late diagnosis, stage of cancer and age all contribute to the success of treatment. [1] The objective of this project was to use advanced image characterisation using the Python package, PyRadiomics, in order to obtain a more numerical assessment of the progression of cervical cancer. This report seeks to detail the first part of the project, which was to develop code and a workflow in order to extract quantitative features and establish the set of the most stable ones. To find this list of stable features, analysis was completed for images taken at the start and ends of a visit, and then repeated across several different visits for 8 patients.

2. Theory

2.1 Basics of radiation therapy

MRI can be used to produce detailed images which can enable clinicians to track the progress of a cancerous region. For this project, images were acquired through T2 Spin Turbo Echo and Dixon sequences. Once the analysis has been conducted, clinicians can use radiotherapy, whereby the beam placements, which provide the ionising radiation dosage that kills the cancer cells, can be adjusted accordingly. [2] Computed tomography (CT) scans also provide a non-invasive method whereby many X-ray measurements are taken around one axis of rotation in order to form a high-resolution 3D image of the organs and tissue inside a patient. The 3D image volume consists of a sequence of tomographic cross-sections ('virtual' 2D slices) across an area, or in this case, the womb and the cervix. However, MRI is non-ionising and generally provides greater resolution compared to CT for soft tissue (due to soft tissue being rich in water molecules, in which the protons provide the signals in MRI) but sometimes both types of scans are used to provide the best insight into a tumour region. [3]

In external beam radiotherapy, a high dosage of ionising radiation (e.g. X-rays, of energies 50 keV–20 MeV) is produced by accelerating electrons close to the speed of light in a linear accelerator (linac), which then hits a tungsten target (**Figure 1**). The Compton Effect is the most significant interaction here to produce the X-rays, where the incident photons transfer energy to valence electrons, ionising the atoms in the target. [4] The linear accelerator is rotated around the cancerous region in order to minimise injury to the surrounding healthy tissue but still maximising damage to the tumour. The ionising radiation damages the DNA in the cancerous tissue preventing them from growing and dividing, leading to necrosis (cell death). Internal beam therapy, known as brachytherapy, may be used too where the radiation source is placed next to the cancerous tissue which increases irradiation to a more localised area, minimising damage to surrounding healthy tissue. [5]

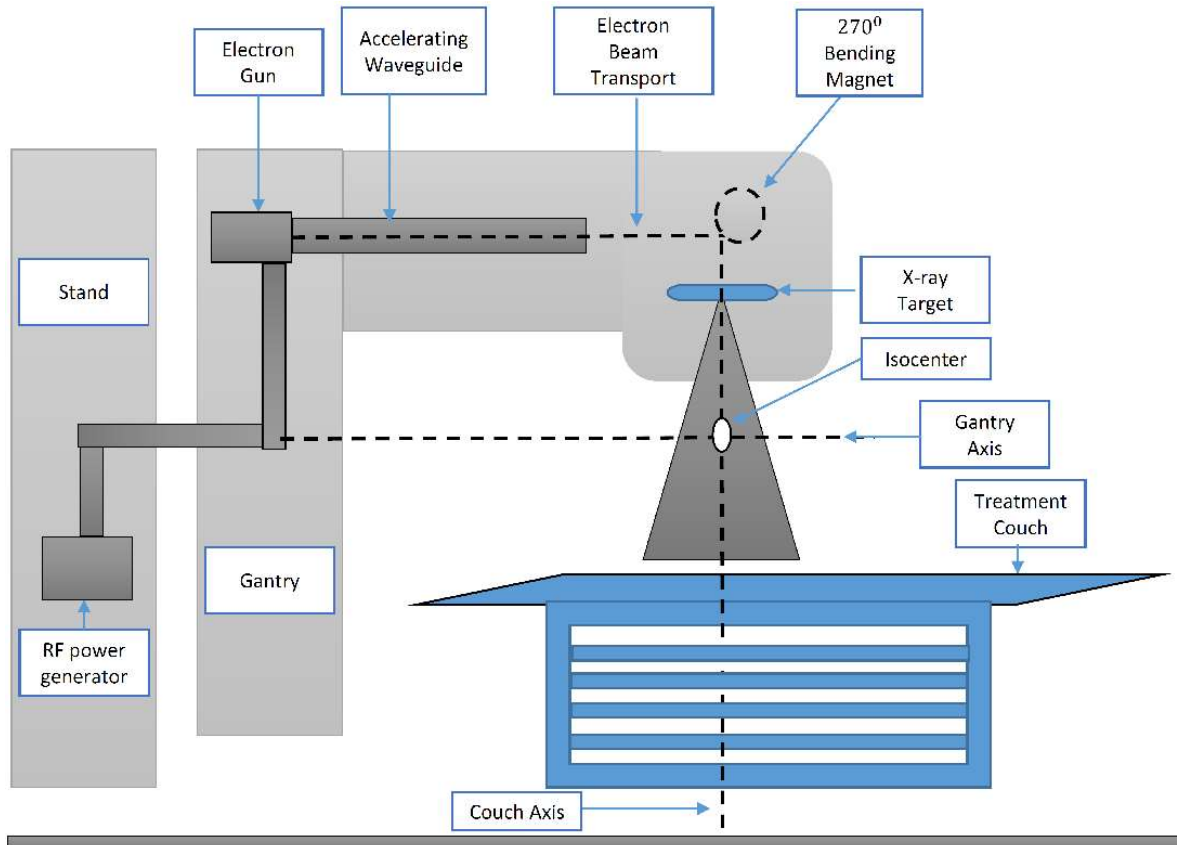


Figure 1: a diagram of a radiotherapy apparatus, which uses a linear accelerator (linac). Thermionic electrons from the electron gun are accelerated to high speeds via a high-frequency AC voltage through the consecutive electrodes in the wave guide. The electrons leave and then collide with a heavy metal target. The X-rays produced match the shape of the tumour upon leaving the gantry, which is rotated around the patient. [6]

2.2 Treatment of cervical cancer

External beam radiotherapy (EBR) is used for early stages and for locally advanced cancer (i.e. cancer that has spread). [7] CT scans may be performed before treatment to help design the treatment specifically for an individual through calculations of the dose needed by using the electron density information of different tissues. [8] EBR may be used in conjunction with chemotherapy as this makes the cervical cancer cells more sensitive to the radiotherapy treatment; this procedure is then followed by brachytherapy if needed. During brachytherapy, a radioactive source is placed next to the cervix, though the precise method is dependent on whether the woman has undergone hysterectomy or not (surgery where the uterus is removed); a low dose rate (LDR) or pulsed dose rate (PDR) is given for up to 24 hours during the irradiation. At the end of the surgery, radiotherapy is sometimes repeated if there is a high risk the cancer may return. [7]

2.3 T2 MRI scans

MRI scans use nuclear magnetic resonance (NMR) where the magnetic properties of atomic nuclei enable randomly orientated water protons within tissue to align via the application of a strong magnetic field. Initially, these spins will not provide a signal since the overall magnetisation will be cancelled out by the random orientations. However, when the alignment of the protons becomes perturbed by externally applied radio frequencies (RFs), the protons

align parallel or anti-parallel to the applied field causing a net magnetisation [9]. The nuclei then return to their default positions, releasing weak radiation at the Larmor frequency, $\omega(\mathbf{r})$, given by,

$$\omega(\mathbf{r}) = -\gamma(H_0 + \mathbf{G} \cdot \mathbf{r}), \quad (1)$$

where γ is the gyromagnetic ratio, H_0 is the effective equilibrium magnetisation and \mathbf{G} is the calibrated magnetic field gradient applied [10] given by,

$$\mathbf{G} = \left(\frac{\partial H_0}{\partial x}, \frac{\partial H_0}{\partial y}, \frac{\partial H_0}{\partial z} \right), \quad (2)$$

where each component is the strength of the magnetic field in the x , y and z directions through the magnetic material.

The emitted signals correspond to two relaxation times; T1 measures the characteristic relaxation time taken for excited protons to return to equilibrium by realigning with the magnetic field (in the longitudinal plane), and depends on the enthalpy of the system (ratio between the number of nuclei with parallel and anti-parallel spins) [11]. T2 measures the decay time for excited protons to lose coherence with each other amongst the nuclei spinning perpendicular to the external magnetic field (in the transverse plane).

As the tissue is not homogenous, this can lead to loss of phase coherence quickly since the nuclei will not all precess at the same rate; a spin echo needs to be used to ensure the contrast in the images for the tissue is discernible. In order to improve the contrast in the image, the spin-echo acquisition method used involves the creation of two consecutive RF pulses. [12] The first RF pulse creates a free induction decay (FID), where a RF pulse at the Larmor frequency will cause a short-lived, transient oscillation in the nuclei. The transverse component of the magnetization will precess around its Larmor frequency, which will then induce an oscillating potential difference in the detection coil around the sample that can subsequently be used to produce an NMR spectrum. [13]

The second RF pulse is the spin echo (SE), with its process shown in **Figure 2**. When the first pulse, the FID, is applied at 90° , the spins will rotate into the transverse plane. The spin groups may precess at different rates and lose phase coherence in accordance with the T2 relaxation time. The SE flips the spins by 180° , enabling the spin groups which were initially precessing faster to become in sync with the groups precessing more slowly, improving their phase coherence. In this case, the echo time (TE) is defined to be twice the time interval between the peaks of the FID and SE intensities; if a spin echo were not used, then a short TE would not allow tissues to be differentiated easily causing poor contrast in the images produced. [13]

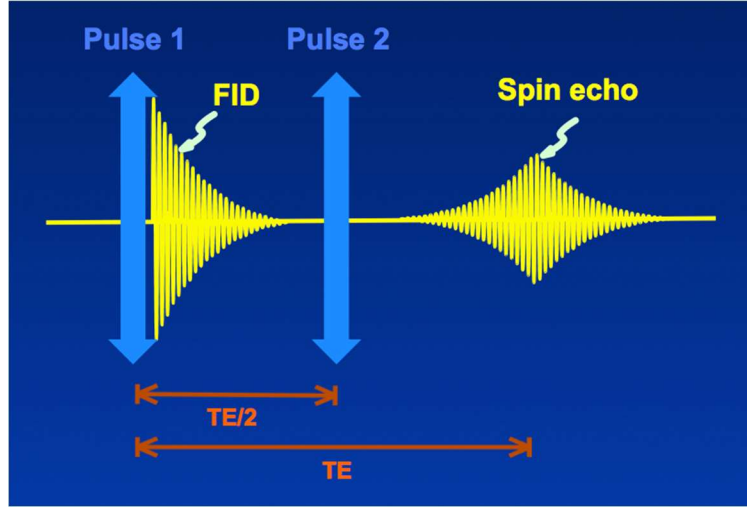


Figure 2: a diagram to show a spin echo is generated. The first pulse creates an FID and the second creates an SE. The TE value is measured between the peaks of both pulses. [13]

T2-weighted scans have longer echo times and repetition time (RT) values – the time between the RF pulses applied – than T1-weighted scans. This produces brighter scans where the fat and water is highlighted, with hard structures such as bone being dark. [14]

The Bloch equations describe how the nuclear magnetisation inside the scanner evolves over time and a permutation for the xy -plane can be given by,

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{H} - \frac{M_{xy}}{T_2} - \frac{M_z - M_0}{T_1}, \quad (3)$$

where \mathbf{M} is the macroscopic spin magnetisation of the polarised water protons, \mathbf{H} is the effective magnetic field, M_0 the equilibrium magnetisation, M_{xy} is the magnetisation in the xy -plane and M_z is the magnetisation in the z -plane. The first term is the precession term, the second is the relaxation term and the third is equilibrium term. The relaxation values vary depending on the interaction of water with larger molecules in the tissues. A Fourier transform on the spatial frequency distribution, $S(k)$, gives the spin density, $\rho(\mathbf{r})$, as given by,

$$\rho(\mathbf{r}) = \frac{1}{2\pi} \int_{k_1}^{k_2} S(k) e^{i\mathbf{k}(t) \cdot \mathbf{r}} d\mathbf{k}, \quad (4)$$

where $\mathbf{k}(t)$ is the spatial frequency vector in a unit of \mathbf{k} -space, $d\mathbf{k}$. [10] The \mathbf{k} -space is an array of values which represent the spatial frequencies in the image, with \mathbf{k}_1 and \mathbf{k}_2 representing the lower and upper bounds of the frequencies used. [15]

The applied gradient field can be used to provide an increasing magnetic field across the area, which increases the frequency and spin-flip energy of emitted signal. This emitted signal contains frequencies for two areas of proton concentration which can be separated via a Fourier transform using (4), enabling the proton frequency to be mapped in two areas. A rotating gradient, \mathbf{G} , allows proton densities to be visualised a 2D map, as proton signals are highly sensitive to different types of tissue (when spin echo is used). [16] The Fourier transforms then

convert these frequencies to locations on an image plane, with intensities determined by levels of greyness. MRI scans allow the tumour to be visualised in the axial, sagittal and coronal planes as shown in **Figure 3**, where white regions are tumours due to the increased density and abnormal cell concentration. The image reconstruction using a Fourier transform is shown in **Figure 4**.

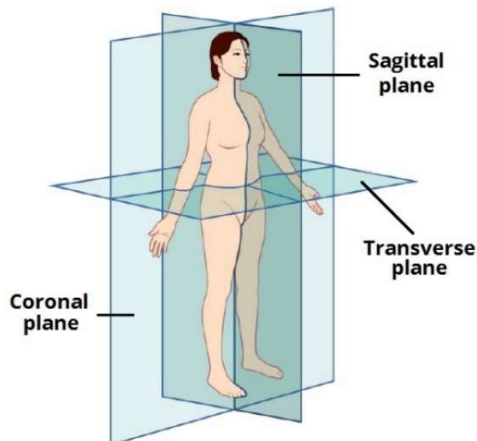


Figure 3: an image to show the anatomical planes: sagittal, coronal and transverse (axial). The coronal and sagittal planes are longitudinal planes which are perpendicular to the axial plane. [17]

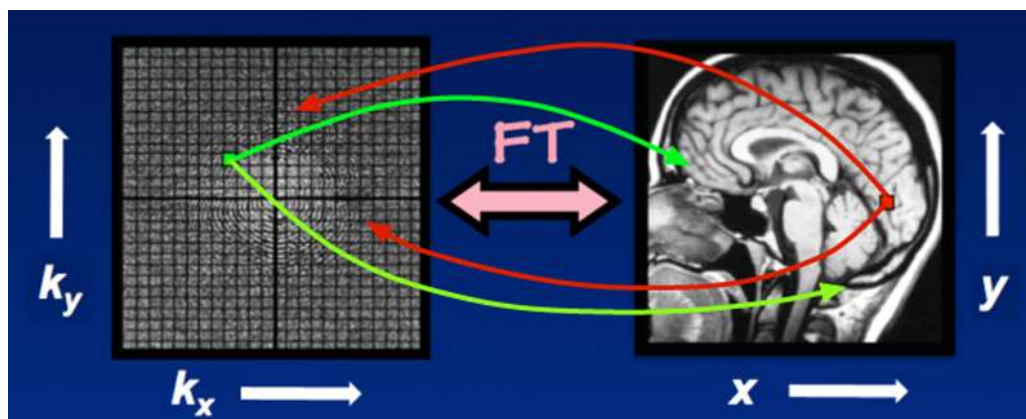


Figure 4: an example of the image reconstruction method used for an MR scan. The k -space signal in the first image - where the white area is the presence of protons - can undergo a Fourier transform into a 2D image as shown. Each point in the k -space corresponds to a spatial frequency. [15]

2.4 Dixon method

Images acquired through the Dixon method enable fat-only and water-only images to be created, where the resonance frequencies of fat and water are used. The Dixon technique is based on the idea of chemical shift in order to acquire fat suppression and one of its main advantages is that it is less affected by artefacts compared to other methods of fat suppression. [18] Through a spin-echo method, the first images are acquired when the TE of protons of fat and water are completely in phase and the second image is obtained when the TE of water and fat protons is completely out of phase. By adding the intensities of the pixel values of the images together, the water protons from the image can be identified (where the fat signal is suppressed) whereas subtracting the out of phase image from the in phase images gives an image with all the fat protons (where the water signal is suppressed) instead. [19]

2.5 Radiomics

Radiomics is a developing field of study where a large number of quantitative features are extracted from medical images; a clinician will draw delineations visually and characteristics (such as heterogeneity) of the tumour can be obtained by running radiomics algorithms. This could potentially be used in the future as biomarkers to lead to greater understanding of the tumour and help with predictions on how the disease is progressing. [20] **Figure 5** displays the process needed in order to use radiomics for analysis of image data.

The features can be grouped into two types of features: “semantic” and “agnostic”. The former describes the regions of interest (i.e. the cervix and surrounding womb) whilst the latter describes the heterogeneity of the tumour through quantitative descriptors. Agnostic features are statistical outputs based on the mathematical information contained and can be grouped into first-order, second-order and higher-order statistics:

- First-order: histogram-based features including mean, median, minimum, maximum, uniformity of the intensities in the image plus skewness and kurtosis in the histogram of intensity values. [21]
- Second-order: texture-based features which give relationships between voxels of similar contrast values. [21]
- Higher-order: filter grid-based features where patterns are imposed onto images and the number of grid elements containing voxels which have a specific value is xfound. [21]

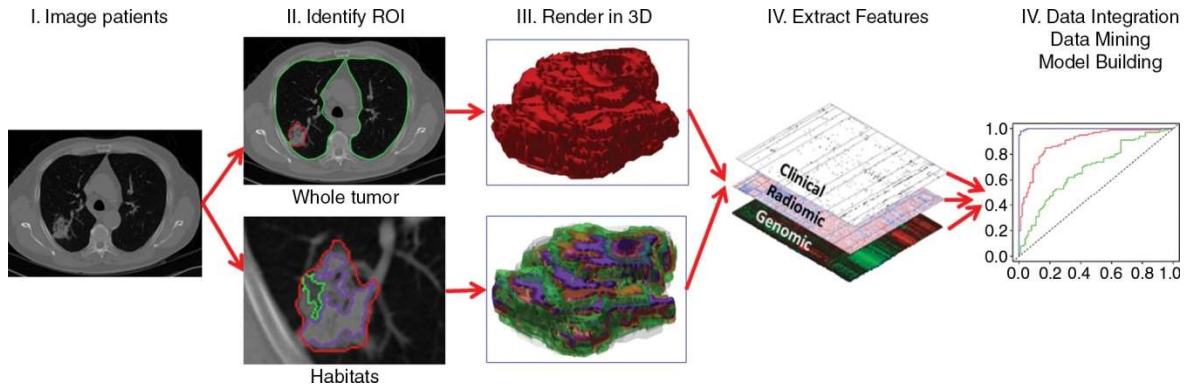


Figure 5: a flowchart to show the step-by-step process in analysing the data obtained through imaging the patient, to drawing delineations, to producing a 3D image, to feature extraction and finally, using this data to build models. [21]

Stable features can be identified by determining which features have the strongest correlations within a group (e.g. a given delineation). The repeatability is tested for by measuring the variability in values of the extracted features from images acquired under the same conditions such as same patient, same scanner and same image parameters. The reproducibility is tested for by the variation in values between images acquired through different conditions including different scanners, kernels, slice thicknesses etc. For this project, the repeatability was tested for by finding the intraclass correlation coefficients (ICCs) of radiomics features within a visit for a patient (for a given delineation) and the stability was measured by finding a count of

the ICCs using the mean values for each feature between patients (for a given delineation) within certain standard deviation limits and ranges. [20]

3. Method

3.1 Initial Steps

The list of stable features were found using MR scans of T2-weighted and Dixon images for 1 patient who had 4 visits (each with a scan at the start and end per visit) and for 7 patients who had 3 visits each. Radiotherapy was performed between the start and endpoints in a visit and the apparatus used was the Siemens MAGNETOM Aera MRI scanner [22], with its mechanisms are shown in **Figure 6**. The T2-weighted and Dixon sequences used resulted in two separate structure files being created (one for each type of sequence); this is because each sequence would have factors affecting the position of the image so one structure file cannot simply be superimposed on both sets of images. Here, the clinician used the Embrace II protocol [23] to follow guidelines in drawing delineations onto the MR images, an example of which is given by **Figure 7**. The WorldMatch medical image viewing software, developed by Marcel Van Herk [24], was subsequently set up to load in these scans (acquired from RayStation [25]). Lua was then utilised to produce scripts in order to manipulate these scans using the Advanced Visual Systems feature and to, for instance, burn on binary masks containing delineations onto another image [24]. A flowchart (**Figure 8**) was developed to show the process utilised in order to generate, extract and analyse the radiomics features.

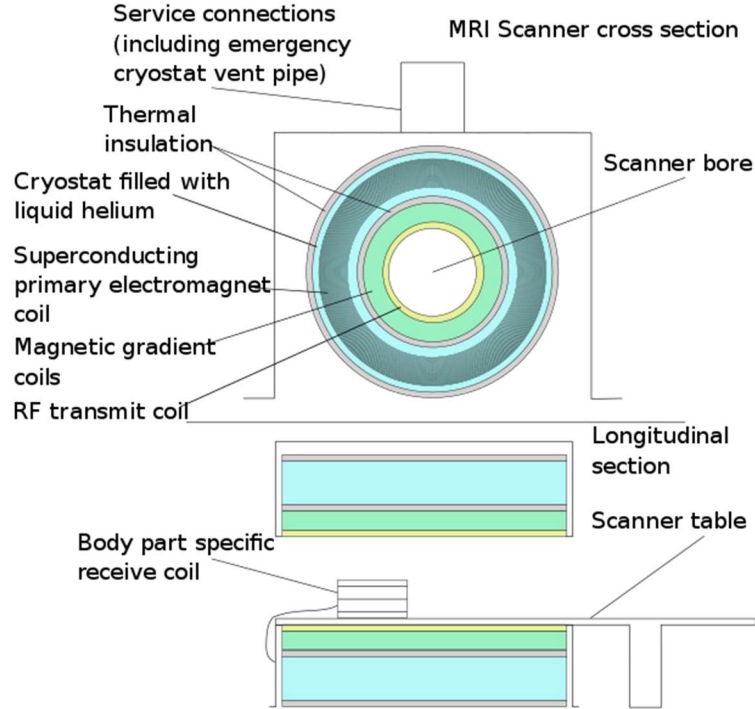


Figure 6: a diagram to show the parts of a superconducting MRI scanner. A ferrous alloy is cooled by liquid Helium to around 4 K, where it becomes a superconductor exhibiting very low resistance when an electric current is applied. A stable, magnetic field of strength 1.5 T is produced by the primary coil; shim coils correct for changes in the homogeneity of the magnetic field; the gradient coils distort the primary field so spatial information can be found through the excitation of the protons. [26]

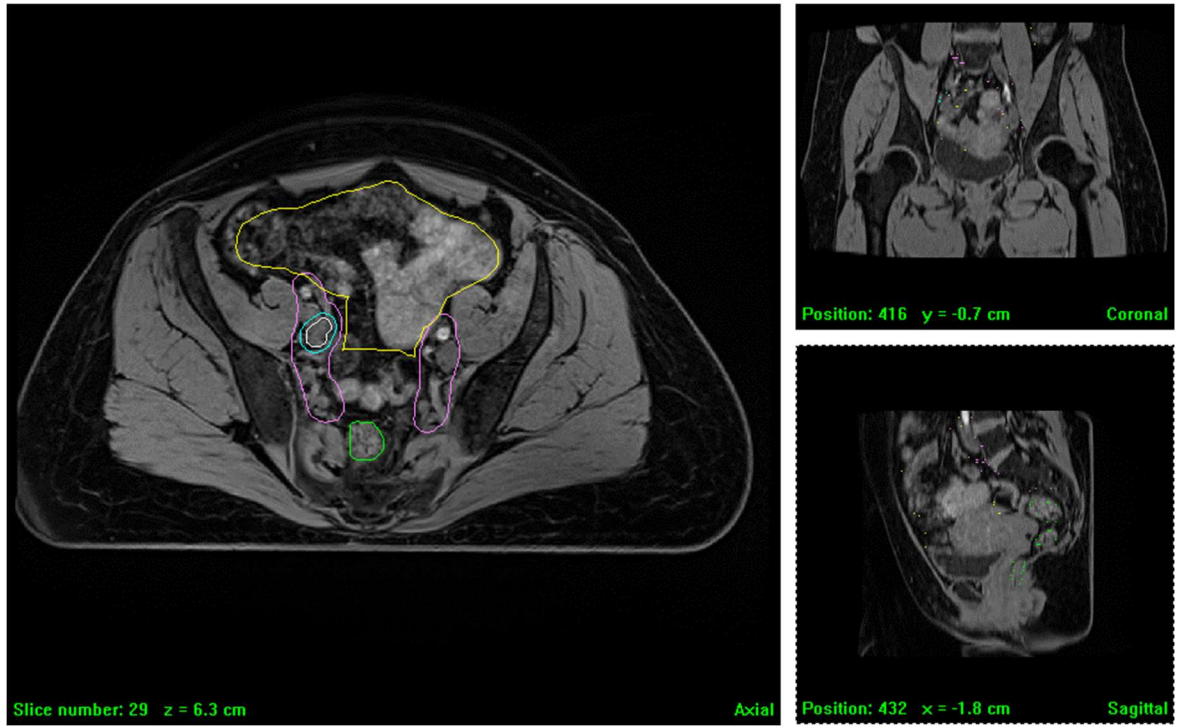


Figure 7: MR scans for the cervix loaded in for a patient at the start of their first visit. A single slice is shown in the axial (left), coronal (top-right) and sagittal planes (bottom-right). The following contours have been shown: GTV-N1 (white), CTV-N1 (cyan), CTV-E (bright pink), bowel (yellow) and rectum boundary (green).

3.2 Types of tumour regions or regions of interests (ROIs)

Delineations were drawn for the gross tumour volume (GTV) and clinical target volume (CTV), with a planning target volume (PTV) used during the radiotherapy process. Their ICRU 50/62/83 definitions have been quoted below:

- GTV: the GTV is the region in which there is “gross palpable or visible extent and location of malignant growth”. [27] No GTV can be defined if the tumour is removed before radiotherapy is performed.
 - GTV-T is the primary tumour volume for the GTV [27]
 - GTV-N refers to the lymph nodes for the GTV [27]
- CTV: the CTV is the “tissue volume that contains a GTV and subclinical microscopic disease which has to be eliminated.” [27]
 - CTV-HR includes “the GTV, the whole cervix and presumed extra-cervical tumour spread”, thereby representing the “macroscopic load” [28]
 - CTV-IR includes the region where there is “significant microscopic tumour load” [28]
 - CTV-LR. Low risk includes region where there is “potential microscopic tumour spread” [29]
 - CTV-E corresponds to nodes in which there is assumed microscopic disease where the cancerous cells may have spread to (EMBRACE II, 9.3.4 [23])

- PTV: the Planning Target Volume (PTV) is defined to be the “CTV plus a fixed or variable margin” [27] and is used when performing radiotherapy. PTV accounts for patient movement and setup errors in the beams.

Flow Chart for Radiomics Extraction

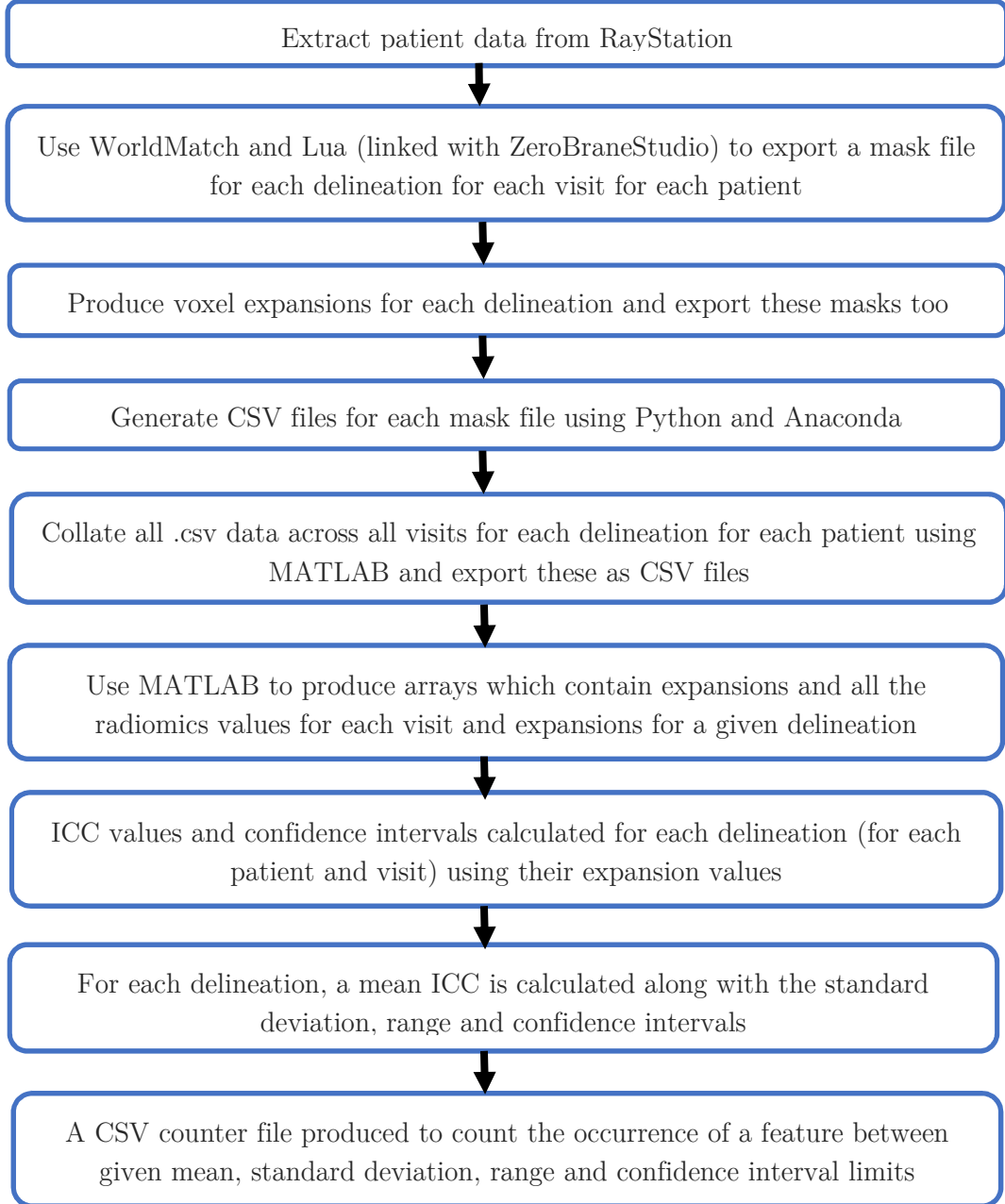


Figure 8: a flowchart to show the process used to extract and interpret radiomics features obtained from a set of MR images.

3.3 WorldMatch and Lua: NIFTI Generation

Once the images had been transferred from the RayStation database to local computers, Lua code was developed in order to automate the process of loading in the DICOM image format (.DCM) slices, followed by loading in the two structure files in turn: one for the T2-weighted scan and the other for the Dixon scan. The Dixon structure file contained delineations on GTV-N, GTV-N1, GTV-N2, GTV-N3, GTV-N4, CTV-N, CTV-N1, CTV-N2, CTV-N3, CTV-N4, CTV-E and Rectum-Boundary; the T2 contained GTV-T, CTV-HR, CTV-IR and CTV-LR. Hence, a mask file was created for each delineation in the NIFTI (.nii) file format, whereby data from all the slices of the MRI scans was combined. **Algorithm 1** (Appendix) displays pseudo-code detailing the process in which these NIFTI files can be generated from Lua code using ZeroBraneStudio [30] linked with WorldMatch.

The Rectum-Boundary followed a slightly different method of NIFTI creation. The rectum contains gas which moves between the start and endpoints of a visit, causing the radiomics values to change - sometimes to 0 if no gas was present since no signal was provided by the protons during the scan. By developing code which would exclude the rectum gas, the effect of this on the rectum radiomic outputs was minimised: by expanding the rectal wall (which is elastic) isotopically and then subtracting the original rectum mask, this mask was loaded into PyRadiomics. Here, the original contour for the Rectum-Boundary was first expanded by 0.26 before it was removed from the original mask. This factor of 0.25 was used since it was found that the median rectal wall thickness was 2.6 mm (amongst 19 control subjects); [31] the rectum boundary was used as a way of comparing normal tissue to the other tissue types.

3.4 Python: Radiomics features extraction

Python code was developed to perform the radiomics extraction, which were output as .csv files. PyRadiomics, extracted 105 features in the categories of First Order, Shape-based, Gray Level Cooccurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTD) and Gray Level Dependence Matrix (GLDM). The algorithm outlining the function of the Python code is displayed in Algorithm 2 (Appendix).

3.5 MATLAB: CSV file collation

MATLAB [32] code developed to loop through all the visit directories for each patient, producing MATLAB tables where for a particular delineation, data for all the visits were collected into columns, storing values for each of the 105 features. This has been outlined in **Algorithm 3** (Appendix), and an example of a table has been provided in **Table 1**, with ‘expansions’ that will be explained in section 3.6.

	Feature Name	Visit 1 End	Visit 1 Start	Visit 2 End etc...
M1 (Contraction)	original_shape_Flatness	0.776339	0.783755	0.682818
	original_shape_LeastAxis	51.07532	50.29826	44.89094
	⋮	⋮	⋮	⋮
P0 (Original)	original_shape_Flatness	0.705733	0.82144	0.741248
	original_shape_LeastAxis	47.74353	45.68017	42.48585
	⋮	⋮	⋮	⋮
P1 (Expansion)	original_shape_Flatness	0.740193	0.744006	0.699005
	original_shape_LeastAxis	37.78459	35.04062	33.05076
	⋮	⋮	⋮	⋮

Table 1: a table to show the feature values which are collated for the CTV_HR delineation for a given patient. The contractions and expansions will be explained alongside the method used for ICC calculations in **Section 3.6**.

3.6 MATLAB: ICC calculations

An ICC value was calculated for each delineation for each patient for each visit. This chosen method has been used frequently across various radiomics studies in order to assess the repeatability – where stable features show little variation across multiple images in the same patient – and reproducibility – where stable features show little variation across different clinics with different protocols, machinery etc. [33]

The ICC can be measured in several different ways, according to McGraw and Wong [33], as shown in **Table 2**, depending on the data set used. It was determined that the “two-way mixed effects, absolute agreement, single rater” – ICC(2,1) – model was the most accurate for the current data set contours:

- Two-mixed effects: only the selected rater is of interest, whereby the rater is defined as the clinician who draws the contours. It is only the reliability of the contours drawn by this clinician that is of interest, and hence, the results of the ICC calculations cannot be generalised to other clinicians in this case. The two-way mixed effects model is appropriate since intrarater reliability is being tested for multiple scores (i.e. multiple visits) from the same rater (i.e. clinician).
- Absolute agreement: the PyRadiomics package ensures that the values of the radiomics features remains the same each time the package runs its calculations on the same set of images.
- Single rater: the contours drawn are from a single clinician.

In order to generate the ICCs, expansion of 1, 2 and 3 voxels along with a contraction of 1 voxel (where 1 voxel was measured to be 1 mm³) were applied to all delineations (except the Rectum-Boundary). These were given the suffixes P1, P2, P3 and M1 respectively, with P0 being the original contour. The aforementioned expansions were found because in order to calculate the ICC, multiple measurements are required by the same rater (clinician), which can be replicated by the voxel expansions since they mimic the movement of gas, as well as the slight differences clinicians would make when drawing contours onto the same image.

Contractions of greater than 1 voxel were not supported by WorldMatch, potentially due to limitations caused by the size of the cervix area. The M1 contraction for the Rectum-Boundary involved first contracting the original mask by 1 voxel and then subtracting the original, and similarly for the expansions. An example of a contraction of 1 voxel and an expansion of 3 voxels are shown for the CTV-E delineation in **Figure 9**.

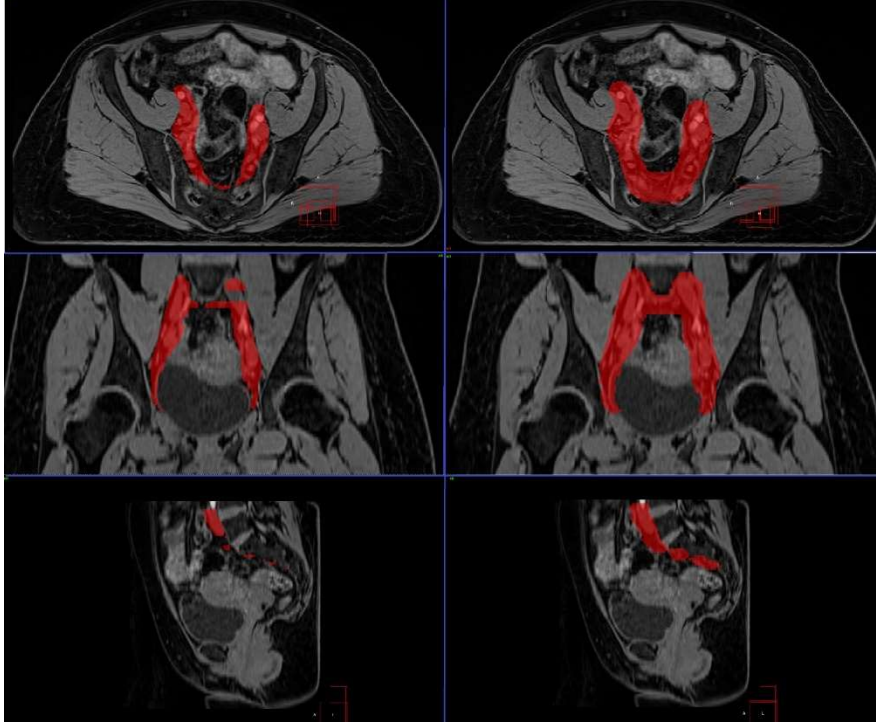


Figure 9: MR scans for the cervix for the CTV-E delineation at the end of a visit for a given patient. The left images are where a contraction of 1 voxel were applied; the right images show an expansion of 3 voxels. The top images are in the axial plane, the middle two are in the coronal plane and the bottom two are in the sagittal plane. [24]

When expansions/contractions are found, the software applies a kernel convolution where values of surrounding voxels determine how the central pixel is modified. As a result, larger voxel expansions resulted in greater smoothing of the rectum boundary compared to smaller voxel expansions, so the error associated with the larger boundary is reduced.

Algorithm 4 (Appendix) displays how the expansions/contractions were used to generate ICC values. A MATLAB function created by Arash Salarian [34] – **Algorithm 5** (Appendix) – enabled the calculation of the ICC(A,1), also called the Case 2: A-1 scenario in this code, which used the formula

$$\text{ICC} = \frac{\text{MS}_R - \text{MS}_E}{\text{MS}_R + (k - 1)\text{MS}_E + \frac{k}{n}(\text{MS}_C - \text{MS}_E)}, \quad (5)$$

where MS_R is the mean square of a feature value between start and endpoints of a visit, MS_E is the variance, MS_C is the mean square for the voxel expansions, n is the number of contour expressions and k is the number of measurements made per visit (i.e. start and end). The

nomenclature used by A. Salarian corresponded to that used by McGraw and Wong and the function also generated the confidence intervals for the ICC values [33].

A mean ICC was then calculated for each delineation, for each patient across all the visits. The standard deviations were calculated too, along with the range between the maximum and minimum values, in order to measure the consistency of the ICC values. In the data available, one of the uncertainties in each mean ICC value, as defined by the 95% confidence intervals added in quadrature, was found to be around 2 orders of magnitude smaller than the sample standard deviations across the visits. Hence, the effect of the confidence intervals on each ICC was neglected and the error on the ICC was defined to be just the standard deviation.

Equivalent ICC Forms Between Shrout and Fleiss (1979) and McGraw and Wong (1996)		
McGraw and Wong (1996) Convention ^a	Shrout and Fleiss (1979) Convention ^b	Formulas for Calculating ICC ^c
One-way random effects, absolute agreement, single rater/measurement	ICC (1,1)	$\frac{MS_R - MS_W}{MS_R + (k+1)MS_W}$
Two-way random effects, consistency, single rater/measurement	–	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E}$
Two-way random effects, absolute agreement, single rater/measurement	ICC (2,1)	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$
Two-way mixed effects, consistency, single rater/measurement	ICC (3,1)	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E}$
Two-way mixed effects, absolute agreement, single rater/measurement	–	$\frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E)}$
One-way random effects, absolute agreement, multiple raters/measurements	ICC (1,k)	$\frac{MS_R - MS_W}{MS_R}$
Two-way random effects, consistency, multiple raters/measurements	–	$\frac{MS_R - MS_E}{MS_R}$
Two-way random effects, absolute agreement, multiple raters/measurements	ICC (2,k)	$\frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}}$
Two-way mixed effects, consistency, multiple raters/measurements	ICC (3,k)	$\frac{MS_R - MS_E}{MS_R}$
Two-way mixed effects, absolute agreement, multiple raters/measurements	–	$\frac{MS_R - MS_E}{MS_R + \frac{MS_C - MS_E}{n}}$

Table 2: table shows a summary of the main types of ICC that can be utilised. ICC(2,1) was found to be most appropriate for the patient data set available [33].

3.7 Python: Produce counters of each ICC for given conditions

In order to assess stability, code was created to import in all the features and their corresponding ICCs/errors for a given delineation. The code would determine if the feature value for an ICC had excellent, good, moderate or weak reliability by checking if it fell within certain bounds, also accounting for the standard deviation and statistical range. Excellent features were identified with the mean in the region $ICC_{\text{mean}} > 0.9$, standard deviation in the region $ICC_{\text{SD}} < 0.2$ and the range in the region $ICC_{\text{range}} < 0.2$. The code would weigh each occurrence (i.e. an excellent value was given a weight of 4 points, good was 3 points, moderate was 2 points and a poor value was given 1 point), and produce a counter for each feature. This was repeated for the other delineations, with CSV files being exported. In addition, additional CSV files were produced containing normalised values, where each count was divided by the total number of appearances of a feature in a given delineation. The counts for CTV N1-N4 and GTV N1-N4 were added together to produce column values for CTV-N and GTV-N respectively. **Algorithm 6** (Appendix) shows the procedure.

4. Results

The normalised table produced from the previous section was then sorted from largest to smallest count for each delineation, with the most stable features (with the highest counts) shown in Tables 3, 4, 5 and 6, in order of largest to smallest count. An arbitrary range of the top 10 features were identified but due to several values for the features being identical, the length of the lists range from 9 to 11 features. The remaining results are shown in the Appendix, which have been colour coded using Microsoft Excel [35], with the GTV-N ordered.

Normal tissue

RECTUM BOUNDARY
original_shape_SurfaceVolumeRatio
original_firstorder_Energy
original_firstorder_TotalEnergy
original_glszm_ZonePercentage
original_gldm_SmallDependenceEmphasis
original_gldm_DependenceNonUniformityNormalized
original_gldm_SmallDependenceLowGrayLevelEmphasis
original_glrlm_RunPercentage
original_glrlm_ShortRunEmphasis
original_glrlm_RunLengthNonUniformityNormalized
original_gldm_SmallDependenceHighGrayLevelEmphasis

Table 3: a list of the most stable features identified for the rectum boundary

Lymph nodes

GTV_N	CTV_N
original_shape_SurfaceArea	original_shape_SurfaceArea
original_shape_Volume	original_shape_Volume
original_glrlm_GrayLevelNonUniformity	original_glrlm_GrayLevelNonUniformity
original_shape_MinorAxis	original_shape_MinorAxis
original_shape_Maximum2DDiameterSlice	original_shape_Maximum2DDiameterSlice
original_glrlm_RunLengthNonUniformity	original_shape_Maximum3DDiameter
original_gldm_DependenceNonUniformity	original_firstorder_Energy
original_shape_Maximum3DDiameter	original_firstorder_TotalEnergy

Table 4: a list of the most stable features for the lymph nodes

Primary tumours

GTV_T	CTV_HR
original_glrlm_RunLengthNonUniformity	original_shape_Volume
original_shape_Volume	original_ngtdm_Coarseness
original_firstorder_Energy	original_glrlm_GrayLevelNonUniformity
original_firstorder_TotalEnergy	original_glrlm_RunLengthNonUniformity
original_gldm_DependenceNonUniformity	original_shape_LeastAxis
original_shape_SurfaceVolumeRatio	original_firstorder_Energy
original_ngtdm_Coarseness	original_firstorder_TotalEnergy
original_shape_SurfaceArea	original_gldm_DependenceNonUniformity
original_glszm_SizeZoneNonUniformity	original_shape_SurfaceVolumeRatio
	original_shape_SurfaceArea

CTV_IR	CTV_LR
original_shape_Volume	original_shape_Volume
original_glrlm_GrayLevelNonUniformity	original_glrlm_GrayLevelNonUniformity
original_firstorder_Energy	original_firstorder_Energy
original_firstorder_TotalEnergy	original_firstorder_TotalEnergy
original_gldm_GrayLevelNonUniformity	original_gldm_GrayLevelNonUniformity
original_ngtdm_Coarseness	original_ngtdm_Coarseness
original_glrlm_RunLengthNonUniformity	original_glrlm_RunLengthNonUniformity
original_shape_LeastAxis	original_gldm_DependenceNonUniformity
original_gldm_DependenceNonUniformity	original_shape_SurfaceVolumeRatio
	original_shape_SurfaceArea
	original_glszm_GrayLevelNonUniformity

Table 5: a list of the most stable features for the primary tumour contours

Elective tissue

CTV_E
original_shape_SurfaceVolumeRatio
original_glrlm_RunLengthNonUniformity
original_gldm_DependenceNonUniformity
original_ngtdm_Coarseness
original_glrlm_GrayLevelNonUniformity
original_shape_Volume
original_firstorder_Energy
original_firstorder_TotalEnergy
original_gldm_GrayLevelNonUniformity

Table 6: a list of the most stable features identified for the elective tissue region

5. Discussion

All 4 groups of features produced a range across 5 categories: Shape, GLRLM, GLDM, NGTDM and First Order which have the following properties [36] :

- Shape: descriptors of the 3D size and shape of ROI.
- GLRLM: this matrix calculates the length of number of pixels which are adjacent to each other and have the same grey level value in a ROI.
- GLDM: measures how frequently pairs with certain values with given spatial relationships occur.
- NGTDM: measures the difference between the average grey value and the neighbouring grey values.
- First Order: describes how the intensities of the voxels within an image area are distributed.

A brief description of the most common features identified has been given in **Table 7** (on page 18), with all the features defined by the official PyRadiomics documentation. [36] The primary tumour region shows 5 features which are stable and similar across the 4 delineations and there does not seem to be a single category the majority of features fall in. The lymph

nodes, interestingly, have the same top 5 features, with the features being dominated by the shape and GLRLM categories.

6. Direction for Semester 2

The most stable features were identified in this part of the project. In the following semester, these features will be investigated further to see how they change for each patient between visits. These features will also be examined in a larger cohort to see if the progression of the feature values can be correlated with patient outcome. An objective is to find which features can be used as future clinical biomarkers, by using some form of hypothesis testing. In addition, the clinical relevance of the stable features will be investigated too beyond the mathematical descriptions given.

The code developed so far will need to be optimised in order to handle a larger amount of data. It is likely that the bottleneck in the overall process – the code library which runs the radiomics calculations – cannot be optimised further as the speed of calculations is determined by the characteristics of the imported PyRadiomics library. Hence, other aspects of the code will be improved instead – e.g. the parts which require manual input will be automated, which can be achieved by reducing the time complexity of the algorithms (e.g. by reducing the number of loops) and by increasing the number of conditional statements to perform checks.

7. Conclusions

The first half of the project as outlined in this report enabled the creation of a workflow in order to progress from MRI scans (with delineations) to the extraction of radiomics features from this data. ICC values were calculated for each visit (between the start and time points), followed by a mean ICC count for each patient (between all the visits). A counter spreadsheet containing a weighted count of ICCs for each feature and delineation was produced thereafter. As the was to produce a list of stable radiomics features for the normal and cancerous tissues (which were divided into the primary tumour, nodes and elective tissue areas), a short list ranging from 9 to 11 features was identified. This was accomplished by checking if the mean ICC values were sufficiently high, with small standard deviations. By using this list of features, they can be explored further through hypothesis testing in the following semester in order to evaluate their clinical significance to see if they can be used as potential biomarkers for the progression of the disease.

Contour	Feature	Description
Normal tissue	original_shape_SurfaceVolumeRatio	The surface volume to ratio indicates how spherical the shape of the tumour is, and is dependent on the ROI. [36]
	original_glszm_ZonePercentage	A measure for the texture coarseness of zones (number of voxels with same intensities) where a high value indicates a large proportion of the ROI has smaller zones (a fine texture). [36]
Primary Tumours	original_shape_Volume	The mesh volume of the ROI, obtained by forming a triangular mesh inside the ROI, with the sum of all the volumes of the tetrahedrons calculated. [36]
	original_firstorder_Energy	This value is a measure of the energy, which is determined by the magnitude of all the voxel values within the image. A larger value indicates more dark grey areas. [36]
	original_firstorder_TotalEnergy	Total energy is the value of energy but scaled by the volume of the voxel. [36]
	original_ngtdm_Coarseness	Coarseness of the feature which measures the rate of change in the space between voxel values, where 0 is completely homogenous. [36]
	original_gldm_DependenceNonUniformity	Measures similarity of dependence across the image, where a lower value means greater homogeneity. [36]
Lymph Nodes	original_shape_SurfaceArea	A triangular mesh is formed within the ROI, with the sum of all the surface areas found. [36]
	original_glrlm_GrayLevelNonUniformity	Grey level nonuniformity measures how the grey-level intensity values in the image vary with a lower value meaning greater homogeneity. [36]
	original_shape_MinorAxis	The minor axis uses the voxel co-ordinates to determine the length of second-largest (i.e. minor) axis of the ellipsoid around the ROI. [36]
	original_shape_Maximum2DDiameterSlice	Maximum 2D diameter is the largest distance provided by a pair of Euclidean co-ordinates between surface mesh vertices in the axial plane. [36]
Elective Tissue	original_glrlm_RunLengthNonUniformity	Similarity of run lengths is measured across the image; lower values indicate more homogeneity. [36]
	original_gldm_GrayLevelNonUniformity	Measures how similar the grey-level intensities are in the image, with a greater value indicating greater similarity. [36]

Table 7: a table to show each tissue type and some of their most features and mathematical definitions

References

- [1] Cancer Research UK. (2019). *Cervical cancer statistics*. [online] Available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/cervical-cancer> [Accessed 5 Jan. 2019].
- [2] Imaging Technology News. (2019). *MRI-guided Radiation Therapy*. [online] Available at: <https://www.itnonline.com/article/mri-guided-radiation-therapy-0> [Accessed 5 Jan. 2019].
- [3] Herman, G. (2009). *Fundamentals of Computerized Tomography*. London: Springer London.
- [4] Gazda, M. (2019). *Principles of radiation therapy*. [ebook] pp.10-12. Available at: <https://pdfs.semanticscholar.org/f6c6/e44bd495775ba1433c6c42ff68f71b7f823f.pdf> [Accessed 5 Jan. 2019].
- [5] The Science Behind Radiation Therapy. (2019). [ebook] American Cancer Society. Available at: <https://www.cancer.org/content/dam/CRC/PDF/Public/6151.00.pdf> [Accessed 5 Jan. 2019].
- [6] Ahmed, S., Ahmed, S., Ahmed, F. and Hussain, A. (2016). SU-F-T-266: Dynalogs Based Evaluation of Different Dose Rate IMRT Using DVH and Gamma Index. *Medical Physics*, [online] 43(6Part16), pp.3523-3524. Available at: https://www.researchgate.net/publication/297806150_Dynamic_Log_Files_Analysis_For_Different_Dose_Rate_IMRT_Using_DVH_and_Gamma_Index.
- [7] Jo's Cervical Cancer Trust. (2019). *Treatments for cervical cancer*. [online] Available at: <https://www.jostrust.org.uk/about-cervical-cancer/cervical-cancer/treatments> [Accessed 5 Jan. 2019].
- [8] Knöös, T. (2019). *Dose Planning and Dose Delivery in Radiation Therapy*. [ebook] Lund University. Available at: https://inis.iaea.org/collection/NCLCollectionStore/_Public/22/077/22077139.pdf [Accessed 5 Jan. 2019].
- [9] Fil.ion.ucl.ac.uk. (2019). *MRI Physics*. [online] Available at: https://www.fil.ion.ucl.ac.uk/mfd_archive/2006/Presentations/MRI%20Physics.ppt [Accessed 5 Jan. 2019].
- [10] Mathematics and physics of emerging biomedical imaging. (1996). Washington, D.C.: National Academy Press, p.Ch 4.
- [11] Köchli, V. and Marincek, B. (2013). *How does MRI work?: An Introduction to the Physics and Function of Magnetic Resonance Imaging*. Springer, p.26.
- [12] Questions and Answers in MRI. (2019). *Image contrast*. [online] Available at: <http://mriquestions.com/image-contrast-trte.html> [Accessed 5 Jan. 2019].
- [13] Chemistry LibreTexts. (2019). 9. *What is the Free Induction Decay?*. [online] Available at: [https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_\(Analytical_Chemistry\)/Analytical_Sciences_Digital_Library/JASDL/Courseware/Quantitative_NMR/02_Basic_NMR_Theory/09._What_is_the_Free_Induction_Decay%3F](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Supplemental_Modules_(Analytical_Chemistry)/Analytical_Sciences_Digital_Library/JASDL/Courseware/Quantitative_NMR/02_Basic_NMR_Theory/09._What_is_the_Free_Induction_Decay%3F) [Accessed 5 Jan. 2019].
- [14] Casemed.case.edu. (2019). *MRI Basics*. [online] Available at: <http://casemed.case.edu/clerkships/neurology/web%20neurorad/mri%20basics.htm> [Accessed 5 Jan. 2019].
- [15] Questions and Answers in MRI. (2019). *k-space*. [online] Available at: <http://mriquestions.com/what-is-k-space.html> [Accessed 5 Jan. 2019].
- [16] Hyperphysics.phy-astr.gsu.edu. (2019). *Magnetic Resonance Imaging*. [online] Available at: <http://hyperphysics.phy-astr.gsu.edu/hbase/Nuclear/mri.html> [Accessed 5 Jan. 2019].
- [17] TeachMeAnatomy. (2019). *Anatomical Planes*. [online] Available at: <https://teachmeanatomy.info/the-basics/anatomical-terminology/planes/>. [Accessed 5 Jan. 2019].
- [18] Gaillard, F. (2019). *Dixon method | Radiology Reference Article | Radiopaedia.org*. [online] Radiopaedia.org. Available at: <https://radiopaedia.org/articles/dixon-method?lang=gb> [Accessed 5 Jan. 2019].
- [19] [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/26583362>.

- [20] Mackin, D. (2017). *Uncertainty, Robustness, and Harmonization in Radiomics Studies of Lung Cancer*.
- [21] Gillies, R., Kinahan, P. and Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, [online] 278(2), pp.563-577. Available at: <https://pubs.rsna.org/doi/full/10.1148/radiol.2015151169>.
- [22] Healthcare.siemens.co.uk. (2019). *MAGNETOM Aera*. [online] Available at: <https://www.healthcare.siemens.co.uk/magnetic-resonance-imaging/0-35-to-1-5t-mri-scanner/magnetom-aera> [Accessed 5 Jan. 2019].
- [23] Tanderup, K. et al. (2019). *EMBRACE-II*. 1st ed. [ebook] Available at: <https://www.embracestudy.dk/UserUpload/PublicDocuments/EMBRACE%20II%20Protocol.pdf> [Accessed 5 Jan. 2019].
- [24] van Herk, Marcel., *WorldMatch*. Version 8.17b. Netherlands Cancer Institute.
- [25] RaySearch Laboratories. *RayStation 7*. Version 6.99.0.33. Available at: <https://www.raysearchlabs.com/>
- [26] Wikimedia (2019). *MRI Schematic Labelled*. [image] Available at: https://upload.wikimedia.org/wikipedia/en/6/62/Mri_scanner_schematic_labelled.svg [Accessed 5 Jan. 2019].
- [27] jouhar (2019). *Icru 50,62,83 volume deliniation*. [online] Slideshare.net. Available at: <https://www.slideshare.net/althafjouhar/icru-506283-volume-deliniation> [Accessed 5 Jan. 2019].
- [28] Amos3.aapm.org. (2019). *Defining targets for brachytherapy*. [online] Available at: <http://amos3.aapm.org/abstracts/pdf/97-25798-352470-109852-834306709.pdf> [Accessed 5 Jan. 2019].
- [29] Brachyacademy.com. (2013). [online] Available at: https://www.brachyacademy.com/wp-content/uploads/2013/10/03_GYN_GEC_ESTRO_November2013.pdf [Accessed 5 Jan. 2019].
- [30] Paul Kulchenko. *ZeroBraneStudio*. Version 1.80. Available at: <https://www.raysearchlabs.com/>
- [31] Rasmussen, S. and Riis, P. (1985). Rectal Wall Thickness Measured by Ultrasound in Chronic Inflammatory Diseases of the Colon. *Scandinavian Journal of Gastroenterology*, [online] 20(1), pp.109-114. Available at: <https://www.ncbi.nlm.nih.gov/pubmed/3887546>.
- [32] MathWorks. *MATLAB*. Version r2015a. Available at: <https://www.mathworks.com/>.
- [33] Koo, T. and Li, M. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), pp.155-163.
- [34] Mathworks.com. (2019). *Intraclass Correlation Coefficient (ICC) - File Exchange - MATLAB Central*. [online] Available at: <https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc> [Accessed 5 Jan. 2019].
- [35] Microsoft Office., *Excel*. Version 365. Microsoft.
- [36] Pyradiomics.readthedocs.io. (2019). *Radiomic Features — pyradiomics 2.1.2.post4+g4b18ce2 documentation*. [online] Available at: <https://pyradiomics.readthedocs.io/en/latest/features.html#>. [Accessed 5 Jan. 2019].
- [37] Herman, G. (2009). *Fundamentals of Computerized Tomography*. London: Springer London.

Risk Assessment

This project was completed in the experiment will be carried out in both the Christie NHS Foundation Trust hospital and the School of Physics and Astronomy Schuster building. All work was carried out via the use of personal laptops in this project.

Hazard Identification:

- Use of portable electrical equipment.
- Prolonged use of display screen equipment, computer keyboard and mouse.

Risks involved:

- Increased risk of repetitive strain injury (RSI) and/or eye strain from prolonged laptop use.
- Laptops and portable electrical equipment such as chargers and cables cause a trip hazard.

Actions to be taken:

- Ensure that any electrical equipment used at the Christie hospital or in the Schuster building have PASS labels (which have not been expired) before use.
- Take regular breaks to relieve eye strain and sit upright in stable, back-supporting chairs where possible.

Appendix

Algorithm 1 for WorldMatch/Lua: creation of NIFTI files

Initialisation

Initialise variables with empty values for each delineation:

 Select which delineations are required

 Select directory containing all the patients

Functions

FileSize Function:

 Counts number of bits in the file

 Return FileSize

ScanDirectory Function:

 Extract list of full directories in a given FilePath

 Return FilePath

Main Code

Loop over the entire directory containing DCMs, Masks and Structures:

 Identify Dixon structure files (RS1) by checking with FileSize (Dixon has larger file size) . File size limits may need to be varied as the structure file sizes may vary

 Load in DCM images into WorldMatch

 Load in delineations from selected mask

 Export DCM slices and a mask combined as a single NIFTI file

 Repeat for T2 structure file

Algorithm 2 for Python: extraction of radiomics features and CSV generation

Initialisation

Import the libraries: radiomics, csv, os, time

masterDirectory = Select directory containing subdirectories

delineationList = Select delineations for which features are needed

Functions

DirectoryExtract Function:

 Extracts list of subdirectories into an array

 Return list of subdirectories

RadiomicsCalculator Function:

 Loads in specified structure file (T2 or Dixon)

 Loads in given NIFTI file

 Write CSV file in current patient directory

FindFile Function:

 Search in substring in a list of strings:

 Return string name

Main Code

Loop over the entire directory containing patients:

 patientDirectory = DirectoryExtract[index_Loop_1]

 Loop over each patient:

 paramPath = DirectoryExtract[index_Loop_2]

 t2Image = FindFile(t2_fileName)

```

DixonImage = FindFile(Dixon_fileName)
if image from list is 'true':
    RadiomicsCalculator extracts features using correct image file

```

Algorithm 3 for MATLAB: collation of CSV files

Initialisation

Select delineations to collate the .csv data from
patientDirectory = Select a patient directory

Main Code

Loop over all the patient visit names:
 visitDirectory = a selected directory
 Within visitDirectory:
 Create list of .csv files
 Create cell arrays and store in MATLAB file

Algorithm 4 for MATLAB: ICCs and errors calculations

Initialisation

Load MATLAB file with arrays
Initiate variables for Feature Name, Mean ICC, Standard Deviation, Range, CI Errors

Main Code

Read cell array
Create table with data in each of the above columns
Use the ICC function to generate ICC values
Write this table to .CSV file

Algorithm 5 for MATLAB: ICC function – developed by Arash Salarian [34]

```

function [r, LB, UB, F, df1, df2, p] = ICC_case_A_1(MSR, MSE, MSC, MSW,
alpha, r0, n, k)
r = (MSR - MSE) / (MSR + (k-1)*MSE + k*(MSC-MSE)/n);

a = (k*r0) / (n*(1-r0));
b = 1 + (k*r0*(n-1))/(n*(1-r0));
F = MSR / (a*MSC + b*MSE);
%df2 = (a*MSC + b*MSE)^2 / ((a*MSC)^2/(k-1) + (b*MSE)^2/((n-1)*(k-1)));

a = k*r/(n*(1-r));
b = 1+k*r*(n-1)/(n*(1-r));
v = (a*MSC + b*MSE)^2 / ((a*MSC)^2/(k-1) + (b*MSE)^2/((n-1)*(k-1)));

df1 = n - 1;
df2 = v;
p = 1-fcdf(F, df1, df2);

Fs = finv(1-alpha/2, n-1, v);
LB = n*(MSR - Fs*MSE)/(Fs*(k*MSC + (k*n - k - n)*MSE) + n*MSR);

Fs = finv(1-alpha/2, v, n-1);
UB = n*(Fs*MSR-MSE)/(k*MSC + (k*n - k - n)*MSE + n*Fs*MSR);

```

Features

original_shape_SurfaceArea	7.666666667	6.25	6.333333	4.25	6.25	6	5	3.714286
original_shape_Volume	7.333333333	6	8.333333	5.75	7	6.666667	5	5.428571
original_glrIm_GrayLevelNonUniformity	7.333333333	6	5.666667	6	6.75	6.666667	5	5.142857
original_shape_MinorAxis	7.333333333	6	5	2.75	4.5	4.666667	2	2.857143
original_shape_Maximum2DDiameterSlice	7	6	4	2.5	2.75	2.333333	3	2
original_glrIm_RunLengthNonUniformity	6.666666667	5.5	9	6.25	6.5	6.333333	5	4.857143
original_gldm_DependenceNonUniformity	6.666666667	5	7.666667	6.25	6.25	6.333333	5	3.714286
original_shape_Maximum3DDiameter	6.666666667	6	4	2	2.75	2	4	2
original_shape_SurfaceVolumeRatio	6	5.25	7.333333	6.5	6.25	6	5	6.285714
original_shape_LeastAxis	5.666666667	5.5	5.666667	5	6.5	6.333333	3	3.142857
original_shape_Maximum2DDiameterColumn	5.666666667	4.5	3.333333	2.5	2.75	3.333333	4	2
original_gldm_GrayLevelNonUniformity	5.666666667	5.5	5.333333	5.75	6	6.666667	5	5.428571
original_firstorder_Energy	5.333333333	6	8	5.75	6.25	6.666667	5	6.285714
original_firstorder_TotalEnergy	5.333333333	6	8	5.75	6.25	6.666667	5	6.285714
original_ngtDm_Coarseness	5.333333333	5	7	6.25	7	6.333333	5	5.714286
original_shape_Maximum2DDiameterRow	5.333333333	5	2.666667	2.25	3.5	5	2	2
original_gldm_DependenceEntropy	5	4.75	5.666667	2.5	4.75	4	3	5.714286
original_glrIm_GrayLevelNonUniformityNormalized	5	3	4.666667	2	5	4	4	2.285714
original_gldm_SumEntropy	4.666666667	3.25	4	2	5	4	4	2
original_shape_MajorAxis	4.666666667	4.75	3.666667	2	4	3.666667	2	2
original_gldm_JointEntropy	4.333333333	2.75	3.666667	2	4.5	4.666667	4	2.285714
original_gldm_DifferenceEntropy	4.333333333	2.25	2.666667	2	3.5	4	3	2
original_firstorder_Entropy	4.333333333	2.75	4.666667	2	5	4.333333	4	2
original_ngtDm_Busyness	4	3	4.666667	3.75	3	3.666667	3	2.857143
original_gldm_LowGrayLevelEmphasis	4	3.25	3.666667	2.5	2	3	2	2.571429
original_firstorder_Skewness	4	3	2.666667	2	2	2.666667	2	2
original_firstorder_Uniformity	4	2.5	4.333333	2.25	5	4.666667	3	2.285714
original_glszm_LargeAreaLowGrayLevelEmphasis	4	3	3.666667	4.75	2	2	2	3.142857
original_gldm_LargeDependenceLowGrayLevelEmphasis	4	3.5	3.666667	2.25	2	3	3	4

original_glszm_LargeAreaHighGrayLevelEmphasis
original_firstorder_MeanAbsoluteDeviation
original_gldm_GrayLevelVariance
original_firstorder_Variance
original_glcmm_SumSquares
original_glrldm_GrayLevelVariance
original_glcmm_ClusterTendency
original_ngtdmm_Contrast
original_glcmm_Correlation
original_glrldm_RunEntropy
original_glrldm_LongRunLowGrayLevelEmphasis
original_glcmm_JointEnergy
original_glszm_ZoneVariance
original_firstorder_RobustMeanMeanAbsoluteDeviation
original_firstorder_RootMeanSquared
original_firstorder_Mean
original_gldm_LargeDependenceEmphasis
original_gldm_DependenceNonUniformityNormalized
original_glrldm_LowGrayLevelRunEmphasis
original_firstorder_10Percentile
original_glcmm_Idn
original_glcmm_Autocorrelation
original_glcmm_JointAverage
original_glcmm_SumAverage
original_glcmm_MaximumProbability
original_glszm_LargeAreaEmphasis
original_glrldm_RunVariance
original_glcmm_ClusterProminence
original_glcmm_DifferenceVariance
original_glszm_ZoneEntropy

4	3.5	2.666667	5.5	2.5	3.666667	2	4.857143
4	3.25	5	2.5	4.75	5.333333	2	2
4	3	4.666667	2.25	4.75	5	2	2.285714
4	3.25	4.666667	2.5	4.75	5	2	2
4	3.5	4	2.5	5.25	4.666667	2	2
4	2.75	5	2.5	4.75	4.333333	2	2.285714
4	3.5	4.333333	2.25	5	4	2	2
3.666666667	2.25	4	2	3.25	2.666667	2	2
3.666666667	3.75	2.666667	2.25	2.5	2	3	2
3.666666667	4.5	5.333333	2	4.5	4	4	4.857143
3.666666667	2.75	3.666667	2.5	2	3	3	2.571429
3.666666667	2.25	3.666667	2.25	4.25	5	3	2.285714
3.666666667	2.75	3.333333	5.25	2	2	2	4.571429
3.666666667	3	5.666667	2.75	5.5	5.333333	2	2.285714
3.666666667	2.5	3.666667	2	3.5	5	4	2
3.666666667	2.5	4	2	3.75	4.666667	5	2
3.666666667	2.5	3	2.5	2.25	4	2	5.714286
3.666666667	3	3.333333	2	2.75	3.666667	2	6.285714
3.666666667	3	3.666667	3.75	2.25	3	2	2.285714
3.666666667	3	5	2	2.75	2	2	2.571429
3.666666667	2.75	2.666667	2	2	2	2	2
3.333333333	3.25	3	2	3.5	5	5	2.857143
3.333333333	3.25	3	2.75	3.75	4.333333	4	2.857143
3.333333333	3.25	3	2.75	3.75	4.333333	4	2.857143
3.333333333	2.25	3.666667	2	4	4.333333	2	2.571429
3.333333333	2.75	3.333333	5.25	2	2	2	4.571429
3.333333333	2.5	2.666667	2.5	2	3.333333	2	3.714286
3.333333333	2.25	3.333333	2	4.25	3.333333	2	2
3.333333333	2	3.666667	2	3.5	3.333333	3	2
3	2.5	3.333333	2	4	4	2	4

original_glcm_Idm	3	2.25	3	2	3.25	4	2	2.571429
original_glcm_Id	3	2.25	3	2.25	3.25	4	2	2.571429
original_firstorder_Median	3	2.5	4	3	3.25	3	4	2
original_glcm_Idmn	3	2.5	2.666667	2	2	2	2	2
original_grlm_LongRunHighGrayLevelEmphasis	3	2.5	3	2	3	4.666667	5	4
original_grlm_ShortRunEmphasis	3	2.5	3	2	2.5	4	2	6
original_shape_Sphericity	3	2	2.666667	5	2	4	5	5.714286
original_grlm_LongRunEmphasis	3	2	2.666667	2.5	2	3.666667	2	4.857143
original_glcm_ClusterShade	3	2	3.666667	2.5	3.5	2	2	2.285714
original_glszm_GrayLevelNonUniformity	3	4	5.333333	4.25	5	5.333333	5	2.285714
original_firstorder_Kurtosis	3	3	2.666667	2	2.5	3.666667	2	2
original_glszm_LowGrayLevelZoneEmphasis	3	3	2.666667	4.25	2.25	2.666667	3	2.857143
original_grlm_RunPercentage	2.666666667	2.25	3	2.5	2.5	4	2	6
original_glcm_DifferenceAverage	2.666666667	2	2.666667	2	3.5	4	3	2.285714
original_glcm_InverseVariance	2.666666667	2	4	2	3.5	4	4	2
original_glcm_Contrast	2.666666667	2	3.333333	2	3.5	4	4	2
original_ngtdm_Complexity	2.666666667	2.25	2.666667	2	2.25	2	2	2
original_gldm_HighGrayLevelEmphasis	2.666666667	2.5	3.333333	2	3.5	5	4	2
original_grlm_ShortRunLowGrayLevelEmphasis	2.666666667	4.25	4	3.25	2.25	3.333333	2	2
original_gldm_DependenceVariance	2.666666667	2.75	2.666667	2	2.25	3	2	4.571429
original_firstorder_Maximum	2.666666667	2.5	2.666667	2	2	2.333333	2	2
original_glszm_ZonePercentage	2.666666667	3	2.666667	5.5	2.5	2	2	6.285714
original_grlm_ShortRunHighGrayLevelEmphasis	2.333333333	2.5	5	2.5	3.5	5.333333	4	2.285714
original_grlm_HighGrayLevelRunEmphasis	2.333333333	2.5	4	2	3.5	5	4	2
original_glcm_Imc2	2.333333333	3.75	3	2.5	3	2.666667	2	2
original_gldm_LargeDependenceHighGrayLevelEmphasis	2.333333333	2	2.666667	2.25	2.25	3.666667	2	4
original_glcm_Imc1	2.333333333	3.25	2.666667	2	2.75	2	2	2.285714
original_firstorder_Minimum	2.333333333	2	2.666667	2	2	2	2	2
original_firstorder_90Percentile	2	2	4.333333	2	4.25	5.333333	3	2.285714
original_glszm_SizeZoneNonUniformity	2	2.25	6	4.25	4.5	4.333333	4	2.857143

original_glcm_Idm

original_glcm_Id

original_firstorder_Median

original_glcm_Idmn

original_grlm_LongRunHighGrayLevelEmphasis

original_grlm_ShortRunEmphasis

original_shape_Sphericity

original_grlm_LongRunEmphasis

original_glcm_ClusterShade

original_glszm_GrayLevelNonUniformity

original_firstorder_Kurtosis

original_glszm_LowGrayLevelZoneEmphasis

original_grlm_RunPercentage

original_glcm_DifferenceAverage

original_glcm_InverseVariance

original_glcm_Contrast

original_ngtdm_Complexity

original_gldm_HighGrayLevelEmphasis

original_grlm_ShortRunLowGrayLevelEmphasis

original_gldm_DependenceVariance

original_firstorder_Maximum

original_glszm_ZonePercentage

original_grlm_ShortRunHighGrayLevelEmphasis

original_grlm_HighGrayLevelRunEmphasis

original_glcm_Imc2

original_gldm_LargeDependenceHighGrayLevelEmphasis

original_glcm_Imc1

original_firstorder_Minimum

original_firstorder_90Percentile

original_glszm_SizeZoneNonUniformity

original_gIrm_RunLengthNonUniformityNormalized
original_ngtdm_Strength
original_glszm_HighGrayLevelZoneEmphasis
original_glszm_SmallAreaLowGrayLevelEmphasis
original_gldm_SmallDependenceLowGrayLevelEmphasis
original_gldm_SmallDependenceHighGrayLevelEmphasis
original_shape_Flatness
original_glszm_GrayLevelVariance
original_glszm_SmallAreaHighGrayLevelEmphasis
original_glszm_GrayLevelNonUniformityNormalized
original_glszm_SizeZoneNonUniformityNormalized
original_shape_Elongation
original_gldm_SmallDependenceEmphasis
original_glszm_SmallAreaEmphasis
original_firstorder_InterquartileRange
original_firstorder_Range

2	2.5	3	2	2.5	4	2	6
2	2	3.666667	2.5	2.25	4	2	3.714286
2	2	3.666667	2.75	4	3.666667	4	2.571429
2	2	2.666667	3.25	2	3.333333	2	2.285714
2	2.25	2.666667	5.5	2	3	2	6
2	2	3.333333	2	4	3	2	6
2	2.75	2.666667	4.25	2	3	2	2.285714
2	2	4.666667	2	2.75	2.666667	2	3.714286
2	2	4.333333	2.5	3.5	2.666667	2	2.285714
2	2	4.333333	2	2.75	2.666667	2	2.285714
2	2	2.666667	2	2.25	2.333333	2	2.857143
2	2	2.666667	2.5	2	2.333333	2	2.571429
2	2	2.666667	4.5	2.5	2	2	6.285714
2	2	2.666667	2	2.25	2	2	2.285714
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0