Published in final edited form as:
Radiother Oncol. 2019 May 01; 134: 119–126. doi:10.1016/j.radonc.2019.01.022.

MRI heterogeneity analysis for prediction of recurrence and disease free survival in anal cancer

Kasia Owczarczyk[a,*], **Davide Prezzi**[a], **Matthew Cascino**[b], **Robert Kozarski**[a], **Andrew Gaya**[c], **Muhammad Siddique**[a], **Gary J. Cook**[a], **Rob Glynne-Jones**[d], **Vicky Goh**[a]

[a]Cancer Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, United Kingdom

[b]Genentech, South San Francisco, United States

[c]Department of Oncology, Guy's and St Thomas' Hospital NHS Foundation Trust, London

[d]Cancer Centre, Mount Vernon Hospital, Northwood, United Kingdom

Abstract

Background—The aim of this study was to evaluate the role of image heterogeneity analysis of standard care magnetic resonance imaging (MRI) in patients with anal squamous cell carcinoma (ASCC) to predict chemoradiotherapy (CRT) outcome. The ability to predict disease recurrence following CRT has the potential to inform personalized radiotherapy approaches currently being explored in novel clinical trials.

Methods—An IRB waiver was obtained for retrospective analysis of standard care MRIs from ASCC patients presenting between 2010 and 2014. Whole tumor 3D volume-of-interest (VOI) was outlined on T2-weighted (T2w) and diffusion weighted imaging (DWI) of the pre- and post-treatment scans. Independent imaging features most predictive of disease recurrence were added to the baseline clinico-pathological model and the predictive value of respective extended models was calculated using net reclassification improvement (NRI) algorithm. Cross-validation analysis was carried out to determine percentage error reduction with inclusion of imaging features to the baseline model for both endpoints.

[*]Corresponding author at: Cancer Imaging, School of Biomedical Engineering and Imaging Sciences, 4th Floor, Lambeth Wing, St. Thomas' Hospital, Westminster Bridge Road, London SE1 7EH, United Kingdom. kasia.owczarczyk@kcl.ac.uk (K. Owczarczyk).

Conflict of interest statements

KO: Dr. Owczarczyk has nothing to disclose.
DP: Dr. Prezzi has nothing to disclose.
MC: Dr. Cascino is an employee of Roche/Genentech.
RK: Dr Kozarski has nothing to disclose.
AG: Dr Gaya has nothing to disclose.
MS: Dr Siddique has nothing to disclose.
GC: Dr. Cook has nothing to disclose.
RGJ: Dr. Glynne-Jones reports grants from Roche, grants from Merck Serono, personal fees from Roche, personal fees from Amgen, personal fees from Servier, personal fees from Sanofi, personal fees from Merck Serono, personal fees from Eli Lilley, personal fees from Roche, personal fees from Home Nutrition, personal fees from Servier, personal fees from Sanofi, personal fees from Eisai, personal fees from Amgen, personal fees from BMS during the conduct of the study.
VG: Dr. Goh reports grants from Siemens Healthineers, outside the submitted work.

Results—Forty patients who underwent 1.5 T pelvic MRI at baseline and following completion of CRT were included. A combination of two baseline MR heterogeneity features (baseline T2w energy and DWI coefficient of variation) was most predictive of disease recurrence resulting in significant NRI ($p = 0 < 0.001$). This was confirmed in cross-validation analysis with 34.8% percentage error reduction for the primary endpoint and 18.1% reduction for the secondary endpoint with addition of imaging variables to baseline model.

Conclusion—MRI heterogeneity analysis offers complementary information, in addition to clinical staging, in predicting outcome of CRT in anal SCC, warranting validation in larger datasets.

Keywords

Anal cancer; Imaging; Biomarkers; MRI

Definitive chemoradiotherapy (CRT) with concomitant mitomycin C (MMC) and 5-fluorouracil (5-FU) forms the backbone of anal squamous cell carcinoma (ASCC) management based on clinical evidence from three pivotal phase III trials [1–3]. However, 25% of patients will relapse following CRT, with 84% of recurrences occurring within 2-years [4–6].

Currently, there is no reliable way to predict which patients will experience disease recurrence following CRT. The ability to predict response to CRT at baseline would be of significant clinical benefit as it would allow the management to be individualized with personalization of RT dose. The baseline American Joint Committee on Cancer (AJCC) tumor, node and metastasis (TNM) staging system has undoubted prognostic value particularly for ≥T3, node positive tumors [1,7,8]. Tumor involvement of the external sphincter and pre-treatment anal function score also have proven prognostic value [9]. However, for node negative tumors and node positive tumors that are ≤5 cm [8] there is greater variation in reported outcomes and prognostic biomarkers are lacking.

The addition of clinical, biochemical and molecular markers to clinical stage may improve predictive accuracy [10,11,12]. While there is good supporting evidence that p16 negative patients are at higher risk of failure following CRT [12], the fact that 90% of human anal cancers are human papilloma virus (HPV)-driven, limits its prognostic utility [13]. There have been preliminary suggestions that factors such as tumor infiltrating lymphocytes may play a role [13] and that genomic as well as cellular heterogeneity, reflecting the presence of different malignant subclonal and stromal cell populations, may affect clinical outcome [14]; however, no particular clonal somatic mutations or biomarkers have been reported to date to predict response to CRT in ACSS.

In terms of imaging, early post CRT assessment based on magnetic resonance imaging (MRI) morphology has limited value due to confounding effects of treatment related changes [15], although a tumor regression grading system (MRI-TRG) has shown some promise in identifying patients who could benefit from early salvage surgery for recurrence [16]. More recently there has been interest in the potential of quantitative ‘phenotyping’ information from medical images beyond volumetric and descriptive measures [17] with

Europe PMC Funders Author Manuscripts

the aim of improving patient stratification into distinct phenotypic subgroups [18]. Still considered investigational, nevertheless such image analysis has shown promise in cancer management, including cancer screening, diagnosis [19], treatment response assessment [20,21] and more recently, in predicting tumor molecular phenotype [18] and disease outcome [22].

We hypothesize that imaging features capture distinct phenotypic differences and may have prognostic value, supplementing clinical and size/volume data in ASCC. Thus we aimed to assess whether MRI heterogeneity features may predict disease recurrence or 2-year disease free survival (DFS) in ASCC undergoing CRT.

Methods

Patient characteristics, treatment and follow-up

Institutional review board (IRB) waiver of informed consent was obtained for this retrospective analysis of consecutive MRI data obtained as part of the standard care pathway. Inclusion criteria were histological diagnosis of anal SCC, completion of a course of CRT with curative intent and availability of baseline MRI and a follow-up MRI performed within 3–6 months of treatment. Cases were recruited from two tertiary care institutions between 2010 and 2014. Exclusion criteria were the absence of DWI MRI sequences, deviations from a standard MRI protocol, poor image quality and absence of visible tumor (Fig. 1 suppl.).

The clinical variables considered for the purpose of multivariate analysis included patient age at diagnosis, patient gender, binary tumour size (<5cm vs ≥5 cm), N stage (N0/N1/N2/N3) and radiological response (complete response (CR) vs no CR).

Radiotherapy was delivered to a mean dose of 50.86 Gy (range 50.4–54 Gy) using a linear accelerator (Elekta, Crawley or Varian, Palo Alto) applying a 3-D conformal or intensity-modulated technique. Concomitant chemotherapy consisted of mitomycin C (MMC) 12 mg/m^2 day 1 with either 5-fluorouracil 1000 mg/m^2/day (continuous venous infusion) days 1–4 and 29–32 or capecitabine (825 mg/m^2 twice a day on radiation days). Following completion of CRT, patients were assessed at 8–10 weeks, every 3 months for the first two years and every 6 months afterwards. Re-staging MRI was carried out at 3–6 months.

Clinical endpoints

The primary endpoint of this study was disease recurrence, either locoregional (defined as biopsy-confirmed evidence of non-complete response at restaging or locoregional tumour detection during follow-up after initial response) or metastatic (defined as occurrence of distant metastasis during CRT, at re-staging, or during follow-up). Patients who were alive and free of recurrences or died without recurrence were censored for these endpoints. All time-to-event end points were measured from completion of CRT. Disease-free survival (DFS) was defined from completion of CRT to the day of locoregional failure or distant recurrence.

MRI acquisition and imaging response assessment

Pelvic MRI scans were performed at baseline and on completion of CRT. Patients were scanned in the supine position on one of three 1.5 T MRI scanners (MAGNETOM Avanto or Aera, Siemens Healthcare, Erlangen, Germany) using a pelvic phased array coil. The standard MRI acquisition protocol included pelvic T2-weighted (T2w) axial, T2w sagittal, T1-weighted axial and DWI axial sequences (b = 0, 100, 500, 800 s/mm^2); additional high-resolution T2w sequences were undertaken parallel and perpendicular to the anal canal. The protocols in the two centres albeit similar, were not standardised for the purpose of this study; respective acquisition details from participating centres have been summarized in Table 1. Patients did not require any additional preparation prior to the examination.

Axial-oblique T2w and the b800-value DWI images, as well as vendor-produced apparent diffusion coefficient ($ADC_{0–800}$) maps, were downloaded from the PACS system onto a standalone work station for further analysis.

Scans were analyzed by two independent readers (a clinical oncologist and radiologist, with 2 and 10 years' MRI experience, respectively) in consensus, blinded to clinical data. Tumor maximum size, volume, extent and TNM stage were recorded. Response to CRT was evaluated in relation to the tumor and nodal stage and in accordance with Response Evaluation Criteria in Solid Tumors (RECIST 1.1) and for the purpose of this study coded in a binary format (complete response versus incomplete/no response).

Heterogeneity feature extraction and selection

Feature extraction and analysis workflow is outlined on Fig. 1. The tumors were manually delineated on all slices of axial T2w and on ADC parametric maps generating 4 separate whole tumor 3D volumes-of-interest (VOI) per case (baseline T2w, baseline DWI, post-CRT T2w and post-CRT DWI). In cases where no residual tumour was present on post-CRT MRI, the corresponding anatomical region with any residual tumor or therapy related changes was outlined, confined to the area of visible mucosal abnormality.

Seventy-eight statistical and fractal heterogeneity parameters were derived per VOI using in-house developed software implemented in Matlab (Matlab 2013, Mathworks, Natick, MA, USA) [23]. Highly correlated variables (r > 0.9) and those exhibiting linear dependencies were removed resulting in ten first-, second-order and fractal features per sequence per timepoint (Table 2). Conventional tumor volume (cm^3) was included in the analysis alongside heterogeneity variables.

Statistical analysis

Statistical analysis was conducted using R (version 3.5.0) Independent heterogeneity features and conventional ROI volume from baseline and follow-up scans were entered individually into univariate logistic regression. The primary outcome was recurrence (binary measure) and the secondary outcome was 2-year DFS. The Benjamini-Hochberg multiple testing correction procedure was used to control the adjusted false discovery rate (FDR). Nominal and FDR-adjusted p-values and 95% confidence intervals are reported.

We used random forest approach [24], and its R package randomForest (version 4.6–14) implementation, to select variables with a highest discriminatory value (≥20% increase in the prediction mean square error after variable permutation, %IncMSE) in predicting primary and secondary outcome.

C-statistic (a goodness-of-fit measure similar to the receiver operating characteristic area under the curve (AUC) as previously reported [17]) was calculated for 1) baseline model = multivariate clinical model including the following clinical variables (age, gender, T stage, N stage); 2) extended model = clinical model with the addition of top performing imaging features. As this was an exploratory analysis intended to generate hypotheses that could be tested in future cohorts, variables were selected for inclusion in multivariate models based on a FDR-adjusted type I error rate of 20% (adjusted $P < 0.2$).

In addition, we used a net reclassification improvement (NRI) – a criterion designed to quantify improvement in model performance as a result of adding a new marker [25,26] – to assess the impact of extending the baseline clinico-pathological model by imaging features. The NRI calculation procedure is implemented in the R package nricens (version 1.6).

The candidate extended models were built by extending the baseline model by each of the most important variables from the random forest model (number of trees 1e5) as well as by their two-element combinations with respect to the primary and secondary endpoints. Each model was cross-validated (leave-one-out) using the R package boot (version 1.3–20). Outcome for each model was reported as percentage average mean squared error reduction from the baseline model for each of the endpoints with baseline cross validation errors listed in the titles.

Results

Patient and treatment characteristics

40 consecutive patients were included in the analysis (25 female, 15 male). Median age at diagnosis was 60.5 (range 37–83). The majority of cases were T2 (17/40, 45%) and T3 (14/40, 37%) tumors and there was an even distribution of N0 (13/40, 32.5%), N1 (12/40, 30%) and N2+ (15/40, 37.5%) disease. Radiotherapy was delivered to a dose of 50.4 Gy in 28# in 34/40 (85%) cases and to a dose of 54 Gy in 30# in 5/40 (12%) cases over a median of 37 days (RT details were not available for one subject). Patients received fluoropyrimidine-based chemotherapy, either as monotherapy or in combination with a single dose of MMC in 21/40 (53%) cases and in combination with cisplatin in 2/40 (5%) cases.

Clinical follow up

Median follow-up was 34.5 months (range 2–102 months) and median time to recurrence was 25 months (2–102 months). During the course of follow-up 11 patients experienced disease recurrence (28%); 4 patients recurred locally, 5 patients recurred with both local and metastatic disease and 2 recurred with metastatic disease only (cervical lymph node and liver metastases, respectively). 4 patients underwent early salvage abdomino-perineal excision of

Europe PMC Funders Author Manuscripts

rectum (APER). 1 year, 2 year- and 3 year-disease free survival (DFS) rate was 90%, 83% and 80%, respectively.

Imaging response assessment

Post therapy MRI was performed at a median of 15 weeks (range 5–19 weeks) from the start of radiotherapy with 90% of scans performed within 12 weeks of completion of treatment. 31/40 (73%) patients achieved a radiological complete response (CR) on both T2-w and DWI. Radiological CR at this time point from the start of CRT did not predict outcome (CR rate: 7/11 (64%) and 24/29 (83%), $p = 0.2$ in patients who subsequently recurred versus patients who did not recur).

Association between MR image heterogeneity and disease recurrence

The imaging features most strongly associated with disease recurrence in univariate analysis were T2w energy and DWI coefficient of variation (CoV) from baseline MRI scans (Table 1 suppl). Patients who developed disease recurrence post CRT demonstrated higher baseline T2w energy and higher DWI CoV. In comparison, there was no statistically significant difference in either baseline tumor or post-treatment residual volume in patients who experienced recurrence versus those that did not, although the post-treatment volume may have been over-represented as it included radiotherapy- induced mucosal change and fibrosis ($p = 0.35$ and $p = 0.06$, respectively).

Predictive accuracy of the baseline clinic-pathological multivariate logistic regression model for disease recurrence including age, gender, T stage (dichotomised) and N stage was numerically increased with the addition of these two key imaging characteristics (AUC = 0.71 for baseline vs AUC = 0.83 for extended model; $p = 0.25$ for comparison), and baseline T2w energy remained associated with disease recurrence after adjusting for clinical characteristics (Fig. 2a; Table 3a).

We applied NRI analysis to the baseline and extended model (including T2w energy and DWI CoV). Resulting NRI = 1.29 statistic value indicates a significant (p-value <0.001) net reclassification improvement by the addition of the 2 imaging variables to the baseline model.

Association between MR image heterogeneity and disease free survival

In the univariate analysis, the imaging features associated with 2-year DFS are listed in Table 2 suppl. When T2w energy and DWI CoV were included in the multivariate regression model for 2-year DFS based on clinical characteristics only, predictive accuracy was improved from AUC 0.71 to AUC 0.95 ($p = 0.09$ for comparison) (Fig. 2b; Table 3b).

Model cross-validation for disease recurrence and 2y-DFS

The following variables were selected from the random forest model which fulfilled the importance criterion: (Baseline T2w Energy, Baseline T2w Entropy, Baseline T2w GLCM: Cluster Prominence, Post CRT T2w GLRL: Low Gray-Level Run Emphasis, Post CRT T2w GLRL: Run Percentage, baseline DWI Coefficient of Variation, baseline DWI Standard Deviation). These variables, as well as all their two-element combinations, were taken

forward to be included in the extended model (n = 7 + 21 models considered) as compared to the baseline clinic-pathological model for both the primary and secondary endpoints. For the primary endpoint the model extended with T2w energy and DWI CoV resulted in the greatest error reduction (34.8%) as compared to the baseline model baseline (CV error = 0.24) (Fig. 3a, suppl Table 3).

For the secondary endpoint the model extended with T2w Energy resulted in the greatest error reduction (30.3%) from baseline (CV error = 0.174) whereas the combined model (T2w energy and DWI CoV) provided reduction by 18.1% (Fig. 3b, suppl Table 3).

Discussion

In this exploratory analysis of imaging heterogeneity features derived from standard care MRI acquired at baseline and following CRT in patients with anal cancer, we identified two imaging features, namely baseline T2w energy and DWI CoV, which appeared to be predictive of CRT outcome, independent of clinical characteristics alone. The addition of these two imaging features to multivariate logistic regression models based on clinical characteristics including age, gender, T and N stage yielded numeric increases in the predictive accuracy for both, disease recurrence as well as 2 year-DFS, when using both, conventional *C*-statistic as well as recently described NRI algorithm.

The extended model (incorporating T2w energy and DWI CoV) demonstrated 34.8% error reduction beyond baseline clinical model in terms of disease recurrence prediction and 18.1% error reduction in terms of 2y-DFS post CRT in an independent cross validation analysis meriting its replication in a larger external validation cohort.

As we await the results of a prospective multi-center Australian trial with the aim of determining whether the addition of functional MRI sequences to morphological data has the potential to predict later disease recurrence [27], to our knowledge, our study is the first to explore the potential of functional MRI (including DWI) heterogeneity analysis to predict clinical outcome following CRT in anal cancer.

An imaging “radiomics” approach, extracting a large number of quantitative texture features from diagnostic images, has the potential to derive more in-depth characterization of the tumor, helping to stratify patients into distinct phenotypic subgroups [18] with the added advantage of being non-invasive and potentially unaffected by sampling bias [11].

To our knowledge, only one previous study investigated the association between baseline MR imaging heterogeneity features and outcome in patients with ASCC undergoing CRT [28]. The authors identified two baseline imaging factors, extracted from baseline T2w sequences, predictive of tumour related event occurrence, highlighting the potential usefulness of MR texture analysis as a predictive factor of outcome for ASSC. The potential advantage of our study is the larger sample size as well as functional (DWI) sequences in the analysis.

Our results highlight the complementary nature of anatomical and functional MR assessment in ASCC. This has also been appreciated in other tumour types, in particular gliomas, where

predictive models combining radiomic features from multiple sequences (T1 and ADC) had higher diagnostic accuracy in discriminating low grade from high grade gliomas than models based on features extracted from a single sequence possibly due to the complexity of the microenvironment in tumors [29]. In other studies, radiomic "signatures" employing heterogeneity analysis of multiple sequences have been shown to correlate with molecular signatures, such as multigene assays in breast cancer [30]

First-order energy is a measure of the overall uniformity (homogeneity) of voxel gray levels while first-order coefficient of variation is a measure of intensity variability (heterogeneity) within an image. We therefore hypothesise that the higher baseline anatomical uniformity combined with higher variation in diffusion (which is considered an indirect indication of cellularity) may reflect unfavourable tumour and host environment characteristics (i.e. aggressive heterogenous tumour combined with less pronounced immune response, immune cell influx etc) leading to resistance to treatment. A study of patients with head and neck squamous cell carcinoma, found that high ADC histogram range measures (ADC_{diff}), reflecting more tumor heterogeneity on DWI, was associated with poorer outcomes which is consistent with our study [31]. In addition, in another study of advanced squamous cell carcinoma of the head and neck, higher imaging uniformity was predictive of local failure [32] which is consistent with the T2w feature analysis in our study.

Our study has several limitations. Firstly, this is an exploratory analysis and the results, whilst interesting and consistent with previous reports in squamous cell cancers, will require validation in a larger cohort. Due to the large number of imaging features and small number of patients with this rare cancer coupled with relatively small number of events the risk of model overfitting is high although we have made every effort to correct for this using two separate indices (*C*-statistic as well as NRI) as well as internal cross-validation analysis.

Secondly, we included both, local as well as metastatic/combined recurrence, as the primary outcome. Due to the limited number of events in our cohort we could not carry out separate logistic regression analyses for these outcomes. Further studies are required to find factors specifically associated with local or distant tumour progression as this may carry therapeutic implication.

Thirdly, signal intensity variability from the MR acquisition protocols and reconstruction algorithms were inevitable due to intrinsic differences in the 1.5 T systems used. However, sequences were the same type and standard quality assurance was performed. In addition, first-order heterogeneity features, which have been found to be most predictive in our study, have been shown in other trials to be more reproducible [23] and less reliant on acquisition parameters [33] than higher order texture features.

Furthermore, there were challenges in defining the optimal VOI on the post CRT scans given the overlap between post CRT change and residual tumor, if present. This was minimized by using a radiation oncologist and radiologist in consensus as well as by inclusion of functional sequences which have, in our previous published work, been demonstrated to improve outlining confidence [34]. Finally, we were unable to independently validate our

findings, however, going forward prospective investigation of model performance in the multicenter PLATO trial (ISRCTN 88455282) is planned.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors acknowledge financial support from the King's College London/University College London Comprehensive Cancer Imaging Centre funded by Cancer Research UK and Engineering and Physical Sciences Research Council (EPSRC) in association with the Medical Research Council and Department of Health; from the Department of Health via the National Institute for Health Research Comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust; and Wellcome EPSRC Centre for Medical Engineering at King's College London (WT 203148/Z/16/Z).

References

[1]. Ajani JA, Winter KA, Gunderson LL, Pedersen J, Benson AB 3rd, Thomas CR Jr, et al. Prognostic factors derived from a prospective database dictate clinical biology of anal cancer: the intergroup trial (RTOG 98–11). Cancer. 2010; 116: 4007–13. DOI: 10.1002/cncr.25188 [PubMed: 20564111]

[2]. James RD, Glynne-Jones R, Meadows HM, Cunningham D, Myint AS, Saunders MP, et al. Mitomycin or cisplatin chemoradiation with or without maintenance chemotherapy for treatment of squamous-cell carcinoma of the anus (ACT II): a randomised, phase 3, open-label, 2 x 2 factorial trial. Lancet Oncol. 2013; 14: 516–24. [PubMed: 23578724]

[3]. Peiffert D, Tournier-Rangeard L, Gerard JP, Lemanski C, Francois E, Giovannini M, et al. Induction chemotherapy and dose intensification of the radiation boost in locally advanced anal canal carcinoma: final analysis of the randomized UNICANCER ACCORD 03 trial. J Clin Oncol. 2012; 30: 1941–8. [PubMed: 22529257]

[4]. Gunderson LL, Winter KA, Ajani JA, Pedersen JE, Moughan J, Benson AB 3rd, et al. Long-term update of US GI intergroup RTOG 98–11 phase III trial for anal carcinoma: survival, relapse, and colostomy failure with concurrent chemoradiation involving fluorouracil/mitomycin versus fluorouracil/cisplatin. J Clin Oncol. 2012; 30: 4344–51. DOI: 10.1200/JCO.2012.43.8085 [PubMed: 23150707]

[5]. Spithoff K, Cummings B, Jonker D, Biagi JJ. Gastrointestinal Cancer Disease Site G.. Chemoradiotherapy for squamous cell cancer of the anal canal: a systematic review. Clin Oncol (R Coll Radiol). 2014; 26: 473–87. [PubMed: 24721444]

[6]. Northover J, Glynne-Jones R, Sebag-Montefiore D, James R, Meadows H, Wan S, et al. Chemoradiation for the treatment of epidermoid anal cancer: 13-year follow-up of the first randomised UKCCCR Anal Cancer Trial (ACT I). Br J Cancer. 2010; 102: 1123–8. DOI: 10.1038/sj.bjc.6605605 [PubMed: 20354531]

[7]. Glynne-Jones R, Nilsson PJ, Aschele C, Goh V, Peiffert D, Cervantes A, et al. Anal cancer: ESMO-ESSO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014; 25: iii10–20. [PubMed: 25001200]

[8]. Gunderson LL, Moughan J, Ajani JA, Pedersen JE, Winter KA, Benson AB 3rd, et al. Anal carcinoma: impact of TN category of disease on survival, disease relapse, and colostomy failure in US Gastrointestinal Intergroup RTOG 98–11 phase 3 trial. Int J Radiat Oncol Biol Phys. 2013; 87: 638–45. DOI: 10.1016/j.ijrobp.2013.07.035 [PubMed: 24035327]

[9]. Deniaud-Alexandre E, Touboul E, Tiret E, Sezeur A, Houry S, Gallot D, et al. Results of definitive irradiation in a series of 305 epidermoid carcinomas of the anal canal. Int J Radiat Oncol Biol Phys. 2003; 56: 1259–73. [PubMed: 12873670]

[10]. Glynne-Jones R, Sebag-Montefiore D, Adams R, Gollins S, Harrison M, Meadows HM, et al. Prognostic factors for recurrence and survival in anal cancer: generating hypotheses from the

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

mature outcomes of the first United Kingdom Coordinating Committee on Cancer Research Anal Cancer Trial (ACT I). Cancer. 2013; 119: 748–55. [PubMed: 23011911]

[11]. Jones CM, Goh V, Sebag-Montefiore D, Gilbert DC. Biomarkers in anal cancer: from biological understanding to stratified treatment. Br J Cancer. 2017; 116: 156–62. DOI: 10.1038/bjc.2016.398 [PubMed: 27923035]

[12]. Serup-Hansen E, Linnemann D, Skovrider-Ruminski W, Hogdall E, Geertsen PF, Havsteen H. Human papillomavirus genotyping and p16 expression as prognostic factors for patients with American Joint Committee on Cancer stages I to III carcinoma of the anal canal. J Clin Oncol. 2014; 32: 1812–7. [PubMed: 24821878]

[13]. Gilbert DC, Serup-Hansen E, Linnemann D, Hogdall E, Bailey C, Summers J, et al. Tumour-infiltrating lymphocyte scores effectively stratify outcomes over and above p16 post chemo-radiotherapy in anal cancer. Br J Cancer. 2016; 114: 134–7. DOI: 10.1038/bjc.2015.448 [PubMed: 26730577]

[14]. Liu SV, Lenkiewicz E, Evers L, Holley T, Kiefer J, Ruiz C, et al. Genomic analysis and selected molecular pathways in rare cancers. Phys Biol. 2012; 9: 065004. [PubMed: 23196986]

[15]. Goh V, Gollub FK, Liaw J, Wellsted D, Przybytniak I, Padhani AR, et al. Magnetic resonance imaging assessment of squamous cell carcinoma of the anal canal before and after chemoradiation: can MRI predict for eventual clinical outcome? Int J Radiat Oncol Biol Phys. 2010; 78: 715–21. [PubMed: 20171812]

[16]. Kochhar R, Renehan AG, Mullan D, Chakrabarty B, Saunders MP, Carrington BM. The assessment of local response using magnetic resonance imaging at 3- and 6-month post chemoradiotherapy in patients with anal cancer. Eur Radiol. 2017; 27: 607–17. DOI: 10.1007/s00330-016-4337-z [PubMed: 27090113]

[17]. Aerts HJ, Grossmann P, Tan Y, Oxnard GG, Rizvi N, Schwartz LH, et al. Defining a radiomic response phenotype: a pilot study using targeted therapy in NSCLC. Sci Rep. 2016; 6 doi: 10.1038/srep33860 [PubMed: 27645803]

[18]. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014; 5 doi: 10.1038/ncomms5006 [PubMed: 24892406]

[19]. Cameron A, Khalvati F, Haider MA, Wong A. MAPS: a quantitative radiomics approach for prostate cancer detection. IEEE Trans Biomed Eng. 2016; 63: 1145–56. [PubMed: 26441442]

[20]. Yip C, Tacelli N, Remy-Jardin M, Scherpereel A, Cortot A, Lafitte JJ, et al. Imaging tumor response and tumoral heterogeneity in non-small cell lung cancer treated with antiangiogenic therapy: comparison of the prognostic ability of RECIST 1.1, an alternate method (Crabb), and image heterogeneity analysis. J Thorac Imaging. 2015; 30: 300–7. [PubMed: 26164165]

[21]. Yip CS, Davnall F, Kozarski R, Landau D, Mason R, Lagergren J, et al. CT tumoral heterogeneity as a prognostic marker in primary esophageal cancer following neoadjuvant chemotherapy. Pract Radiat Oncol. 2013; 3: S3. [PubMed: 24674540]

[22]. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol. 2015; 114: 345–50. DOI: 10.1016/j.radonc.2015.02.015 [PubMed: 25746350]

[23]. Gourtsoyianni S, Doumou G, Prezzi D, Taylor B, Stirling JJ, Taylor NJ, et al. Primary rectal cancer: repeatability of global and local-regional MR imaging texture features. Radiology. 2017; doi: 10.1148/radiol.2017161375 [PubMed: 28481194]

[24]. Breiman. Random forests. Mach Learn. 2001; 45: 5–32.

[25]. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011; 30: 11–21. DOI: 10.1002/sim.4085 [PubMed: 21204120]

[26]. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008; 27: 157–72. [PubMed: 17569110]

[27]. Jones M, Hruby G, Stanwell P, Gallagher S, Wong K, Arm J, et al. Multiparametric MRI as an outcome predictor for anal canal cancer managed with chemoradiotherapy. BMC Cancer. 2015; 15: 281. doi: 10.1186/s12885-015-1244-7 [PubMed: 25885556]

[28]. Hocquelet A, Auriac T, Perier C, Dromain C, Meyer M, Pinaquy JB, et al. Pre-treatment magnetic resonance-based texture features as potential imaging biomarkers for predicting event free survival in anal cancer treated by chemoradiotherapy. Eur Radiol. 2018; 28: 2801–11. [PubMed: 29404766]

[29]. Qin JB, Liu Z, Zhang H, Shen C, Wang XC, Tan Y, et al. Grading of gliomas by using radiomic features on multiple magnetic resonance imaging (MRI) sequences. Med Sci Monit. 2017; 23: 2168–78. DOI: 10.12659/MSM.901270 [PubMed: 28478462]

[30]. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. Radiology. 2016; 281: 382–91. DOI: 10.1148/radiol.2016152110 [PubMed: 27144536]

[31]. Choi JW, Lee D, Hyun SH, Han M, Kim JH, Lee SJ. Intratumoural heterogeneity measured using FDG PET and MRI is associated with tumour-stroma ratio and clinical outcome in head and neck squamous cell carcinoma. Clin Radiol. 2017; 72: 482–9. [PubMed: 28285707]

[32]. Jansen JF, Lu Y, Gupta G, Lee NY, Stambuk HE, Mazaheri Y, et al. Texture analysis on parametric maps derived from dynamic contrast-enhanced magnetic resonance imaging in head and neck cancer. World J Radiol. 2016; 8: 90–7. DOI: 10.4329/wjr.v8.i1.90 [PubMed: 26834947]

[33]. Zhao B, Tan Y, Tsai WY, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. Sci Rep. 2016; 6 doi: 10.1038/srep23428 [PubMed: 27009765]

[34]. Prezzi D, Mandegaran R, Gourtsoyianni S, Owczarczyk K, Gaya A, Glynne-Jones R, et al. The impact of MRI sequence on tumour staging and gross tumour volume delineation in squamous cell carcinoma of the anal canal. Eur Radiol. 2018; 28: 1512–9. DOI: 10.1007/s00330-017-5133-0 [PubMed: 29134349]

Europe PMC Funders Author Manuscripts

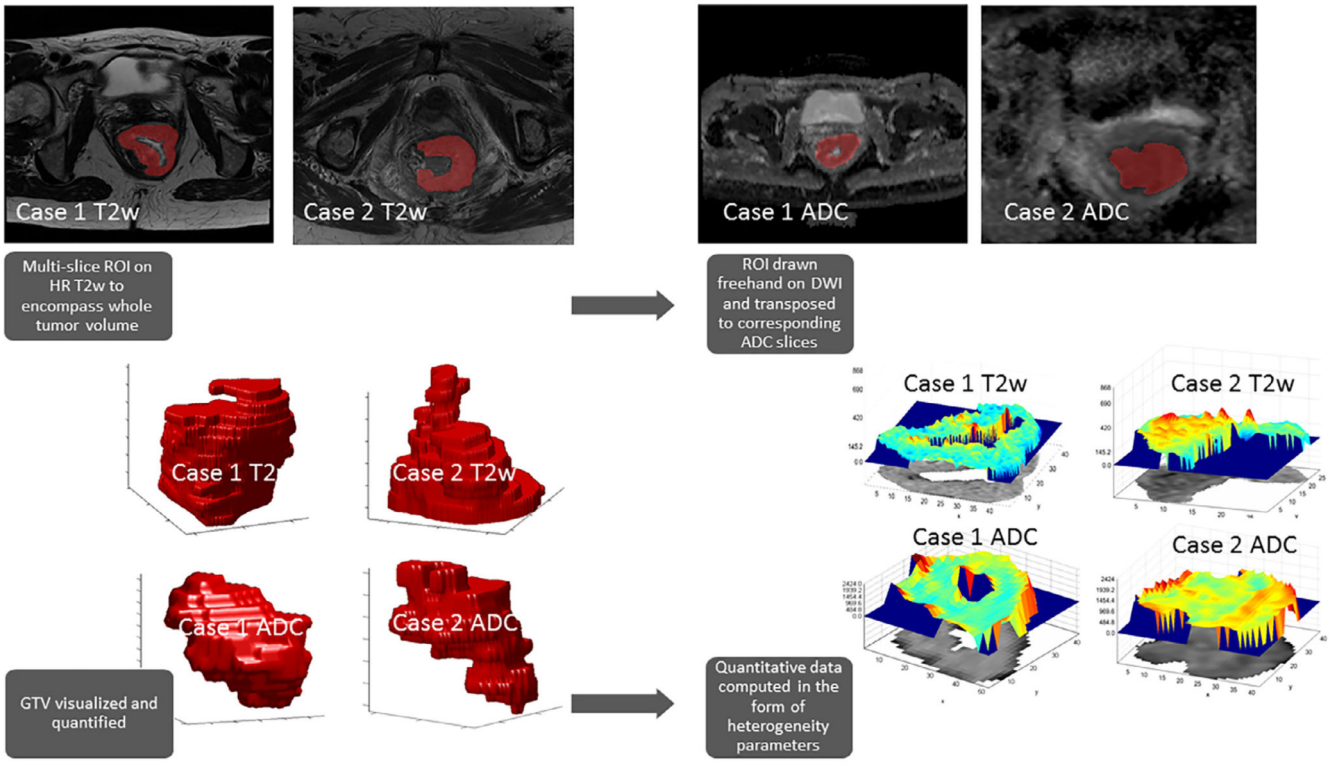
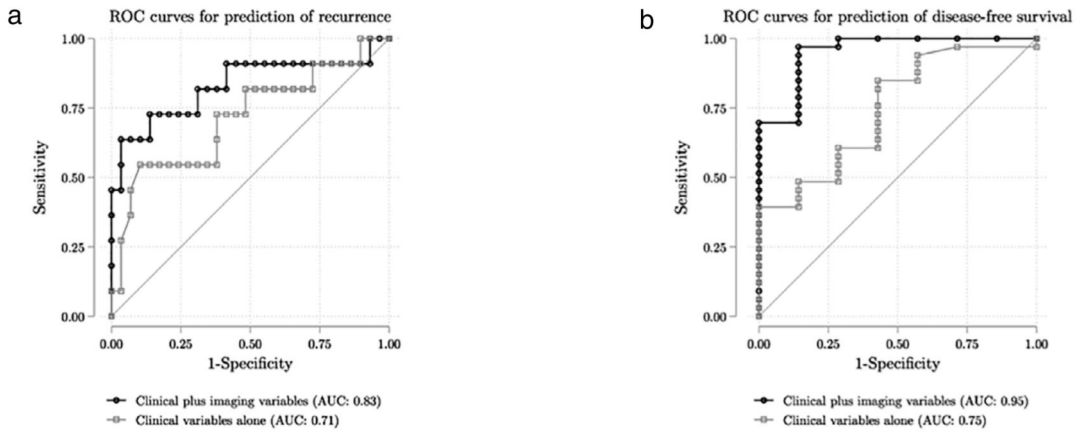


Fig 1.

Analysis workflow. Multi-slice regions of interest (ROI) were drawn around the tumor on high resolution T2w images as well as on Apparent Diffusion Coefficient (ADC) parametric maps with reference to high b value diffusion weighted images (DWI), generating a 3D whole tumor volume. From this volume, first-, second- and high-order statistical and fractal parameters were derived using in-house software. Both illustrated cases are of T4 anal canal tumors (Case 1 – Partial response post CRT, delayed recurrence; Case 2 – Partial response post CRT, subsequently CR and no recurrence at last follow-up).

Europe PMC Funders Author Manuscripts



	Obs	ROC Area	Std. Err.	-Asymptotic Normal-- [95% Conf.	Interval]
clinimaging	40	0.8307	0.0906	0.65318	1.00000
clinical	40	0.7147	0.1023	0.51422	0.91525

Ho: area(clinimaging) = area(clinical)
chi2(1) = 1.44 Prob>chi2 = 0.2297

	Obs	ROC Area	Std. Err.	-Asymptotic Normal-- [95% Conf.	Interval]
clinimagin~s	40	0.9524	0.0448	0.86450	1.00000
clindfs	40	0.7468	0.1036	0.54373	0.94978

Ho: area(clinimagin~s) = area(clindfs)
chi2(1) = 2.79 Prob>chi2 = 0.0950

Fig 2.

Receiver operator curves (ROC) for prediction of recurrence (a) and disease free survival (b) comparing model using clinical variables alone with model using clinical and imaging variables as described in text.

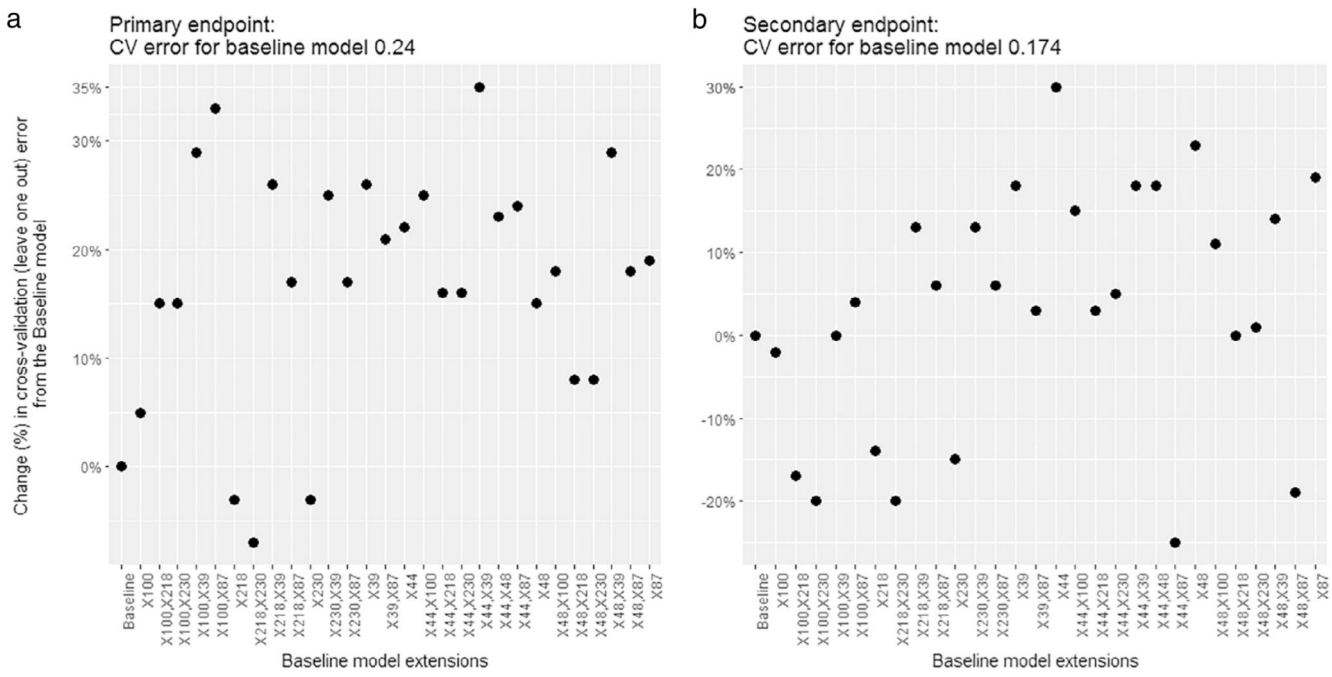


Fig 3.

Percentage change in cross-validation error from the baseline model for prediction of disease recurrence (a) and disease free survival (b).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

T2-weighted and diffusion-weighted MRI acquisition parameters utilized.

Sequence	Acquisition parameters (Center 1)	Acquisition parameters (Center 2)
T2 axial TSE	TR/TE: 5010/137 ms;	TR/TE: 5290/97 ms;
Tumor	Flip angle 137° NSA 4; ST 3 mm FoV 220*220 mm	Flip angle 150° NSA 2; ST 3 mm FoV 200*200 mm
Diffusion axial SS-EPI	TR/TE: 3000/65 ms;	TR/TE: 3100/77 ms;
Pelvis	Flip angle 90; NSA 4	Flip angle 90; NSA 4
b = 0, 100, 500, 800, (1100) s/mm^2	ST 6 mm Fov 260*260 mm	ST 6 mm Fov 250*250 mm

TSE, turbo spin echo; TR, repetition time; TE, echo time; NSA, number of signals averaged; ST, slice thickness; FoV, field of view; SS-EPI, single-shot echo planar imaging.

Table 2

Summary description and relevant formulae for the independent first, second and fractal features analyzed.

Parameter	Description	Formula
First order: histogram statistics provide an indication of central tendency (coefficient of variation) and variability (kurtosis, energy and entropy)		
Coefficient of Variation (CoV)	Indicates how large the standard deviation is in relation to the mean	$\frac{\sigma}{\mu}$
Kurtosis	Describes the "peak" of a distribution. Kurtosis >3: sharper peak than a normal distribution Kurtosis <3: flatter peak than a normal distribution Kurtosis = 3: normal distribution	$\frac{n}{(n-1)(n-2)(n-3)}\frac{\sum_{(x,y)\in R}[a(x,y)-a]^4}{[sd(a)]^4}-3\frac{(n-1)^2}{(n-2)(n-3)}$ where n = the total number of voxels in the region-on-interest, R within the image $a(x,y)$; sd = standard deviation; ā is the mean value within R
Energy	Measures voxel signal distribution. High energy is noted in homogeneous voxels	$\sum_{i=1}^{imax}[p(i)]^2$ where i is the voxel value (between $i=1$ to i_{max} in the region of interest and $p(i)$ the probability of the occurrence of that voxel value
Entropy	Measures voxel randomness. Low entropy is noted in homogeneous voxels	$\sum_{i=1}^{imax}[p(i)ln[p(i)]$ where i is the voxel value (between $i=1$ to i_{max} in the region of interest and $p(i)$ the probability of the occurrence of that voxel value
Second order: Gray Level Co-occurrence matrix (GLCM) statistics are computed after the original texture image D is re-quantized into an image G with reduced number of gray level, Ng by scanning the intensity of each voxel and its neighbour, defined by displacement d and angle θ. A displacement, d could take a value of 1,2,3,…n whereas an angle, θ is limited 0°, 45°, 90° and 135°. The GLCM $p(i; j|d; \theta)$ is a second order joint probability density function of gray level pairs in the image for each element in the co-occurrence matrix by dividing each element with Ng. Finally, scalar secondary features are extracted from this co-occurrence matrix		
GLCM: Correlation	Measures gray level intensity linear dependence between the voxels (i,j) at the specified positions relative to each other	$\sum_i\sum_j(ij)p(i,j)$ where i is the voxel value (between $i=1$ to i_{max} in the region of interest; j is the voxel value (between $j=1$ to j_{max} in the region of interest; and $p(i,j)$ the probability of the occurrence of that voxel value i relative to j
GLCM: Cluster prominence	Measures asymmetry. A low cluster prominence value indicates small variations in gray-scale	$\sum_i\sum_j(i+j-\mu_x-\mu_y)^4p(i,j)$ where i is the voxel value (between $i=1$ to i_{max} in the region of interest; j is the voxel value (between $j=1$ to j_{max} in the region of interest; $p(i,j)$ is the probability of the occurrence of that voxel value i relative to j; μ_x is the mean of p_x and μ_y is the mean of p_y
Fractal features describe self-similar fractal shapes		
Mean fractal dimension	Measures the texture of a fractal, a self similar pattern. A higher fractal dimension corresponds to greater roughness	$\overline{D}=\frac{\sum_{i=1}^{N}D_i}{N}$ where N is the number of slices and D_i is the fractal dimension for the i^{th} slice
Standard deviation	Measures the standard deviation of a fractal computed by a differential box counting algorithm	$\sigma=\sqrt{\sum_{i=1}^{N}D_i^2/N-\left(\frac{\sum_{i=1}^{N}D_i}{N}\right)^2}$ where N is the number of slices and D_i is the fractal dimension for the i^{th} slice
Lacunarity	Measures the amount of "gaps" in the image/object. If a fractal has large "gaps", it has high lacunarity	$\frac{\left[\sum_{i=1}^{N}D_i^2/N\right]}{\left[\left(\frac{\sum_{i=1}^{N}D_i}{N}\right)^2\right]}-1$ where N is the number of slices and Di is the fractal dimension for the i^{th} slice
Hurst component	Measures the density of the image/object i.e. how much the image/object occupies	$H=3-\overline{D}$ where $\overline{D}$ is the mean fractal dimension

Parameter	Description	Formula
	the space that contains it. A small value corresponds to coarse texture	

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 3

Results of the cross-validation analysis for extended models based on most predictive imaging variables. Error from each model cross-validation was reported together with its percentage change from the baseline model one for each of the endpoints.

Model	Primary endpoint (Disease recurrence)		Secondary endpoint (2y-DFS)	
	CV error	CV error change (%) from the Baseline Model	CV error	CV error change (%) from the Baseline Model
Baseline Model	0.240	0.0	0.174	0.0
Baseline model extensions:				
Baseline T2w Energy, Baseline T2w Entropy	0.186	−22.8	0.142	−18.3
Baseline T2w Energy, Baseline T2w GLCM: Cluster Prominence	0.181	−24.5	0.148	−14.9
Baseline T2w Energy, Post CRT T2w GLRL: Low Gray-Level Run Emphasis	0.201	−16.2	0.168	−3.2
Baseline T2w Energy, Post CRT T2w GLRL: Run Percentage	0.203	−15.6	0.166	−4.6
Baseline T2w Energy, Baseline DWI Coefficient of Variation	0.157	−34.8	0.142	−18.1
Baseline T2w Energy, Baseline DWI Standard Variation	0.183	−23.7	0.217	24.7
Baseline T2w Entropy, Baseline T2w GLCM: Cluster Prominence	0.196	−18.4	0.155	−10.6
Baseline T2w Entropy, Post CRT T2w GLRL: Low Gray-Level Run Emphasis	0.221	−8.2	0.173	−0.2
Baseline T2w Entropy, Post CRT T2w GLRL: Run Percentage	0.222	−7.8	0.172	−1.3
Baseline T2w Entropy, Baseline DWI Coefficient of Variation	0.170	−29.1	0.150	−13.9
Baseline T2w Entropy, Baseline DWI Standard Variation	0.196	−18.4	0.208	19.4
Baseline T2w GLCM: Cluster Prominence, Post CRT T2w GLRL: Low Gray-Level Run Emphasis	0.203	−15.3	0.204	17.1
Baseline T2w GLCM: Cluster Prominence, Post CRT T2w GLRL: Run Percentage	0.205	−14.6	0.208	19.8
Baseline T2w GLCM: Cluster Prominence, Baseline DWI Coefficient of Variation	0.171	−28.7	0.173	−0.3
Baseline T2w GLCM: Cluster Prominence, Baseline DWI Standard Variation	0.162	−32.7	0.167	−3.8
Post CRT T2w GLRL: Low Gray-Level Run Emphasis, Post CRT T2w GLRL: Run Percentage	0.257	6.8	0.208	19.7
Post CRT T2w GLRL: Low Gray-Level Run Emphasis, Baseline DWI Coefficient of Variation	0.178	−25.7	0.151	−13.0
Post CRT T2w GLRL: Low Gray-Level Run Emphasis, Baseline DWI Standard Variation	0.199	−17.3	0.163	−6.5
Post CRT T2w GLRL: Run Percentage, Baseline DWI Coefficient of Variation	0.179	−25.3	0.152	−12.8
Post CRT T2w GLRL: Run Percentage, Baseline DWI Standard Variation	0.200	−16.6	0.163	−6.3
Baseline DWI Coefficient of Variation, Baseline DWI Standard Variation	0.189	−21.5	0.168	−3.1
Baseline T2w Energy	0.188	−21.7	0.121	−30.3
Baseline T2w Entropy	0.203	−15.4	0.134	−22.6
Baseline T2w GLCM: Cluster Prominence	0.228	−5.0	0.177	2.0
Post CRT T2w GLRL: Low Gray-Level Run Emphasis	0.247	3.0	0.198	14.1
Post CRT T2w GLRL: Run Percentage	0.248	3.3	0.201	15.4
Baseline DWI Coefficient of Variation	0.178	−26.0	0.142	−18.2

Europe PMC Funders Author Manuscripts

Model	Primary endpoint (Disease recurrence)		Secondary endpoint (2y-DFS)	
	CV error	CV error change (%) from the Baseline Model	CV error	CV error change (%) from the Baseline Model
Baseline DWI Standard Variation	0.195	−18.8	0.140	−19.3