Predicting the Tumor Response to Chemoradiotherapy for Rectal Cancer: Model Development and External Validation using MRI Radiomics.

Philippe Bulens[1,*], **Alice Couwenberg**[2,*], **Martijn Intven**[2], **Annelies Debucquoy**[1], **Vincent Vandecaveye**[3], **Eric Van Cutsem**[4], **André D'Hoore**[5], **Albert Wolthuis**[5], **Pritam Mukherjee**[6], **Olivier Gevaert**[6], **Karin Haustermans**[1]

[1]Department of Radiation Oncology, University Hospitals Leuven, Leuven, Belgium [2]Department of Radiation Oncology, University Medical Center Utrecht, Utrecht, The Netherlands [3]Department of Radiology, University Hospitals Leuven, Leuven, Belgium [4]Department of Digestive Oncology, University Hospitals Leuven, Leuven, Belgium [5]Department of Abdominal Surgery, University Hospitals Leuven, Leuven, Belgium [6]Stanford Center for Biomedical Informatics Research, Department of Medicine and Biomedical Data Science, Stanford University, Stanford, USA

Abstract

Background—In well-responding patients to chemoradiotherapy for locally advanced rectal cancer (LARC), a watch-and-wait strategy can be considered. To implement organ-sparing strategies, accurate patient selection is needed. We investigate the use of MRI-based radiomics models to predict tumor response to improve patient selection.

Materials and Methods—Models were developed in a cohort of 70 patients and validated in an external cohort of 55 patients. Patients received chemoradiation followed by surgery and underwent T2-weighted and diffusion-weighted MRI (DW-MRI) before and after chemoradiation. The outcome measure was (near-)complete pathological tumor response (ypT0–1N0).

Tumor segmentation was done on T2-images and transferred to b800-images and ADC maps, after which quantitative and four semantic features were extracted. We combined features using principal component analysis and built models using LASSO regression analysis. The best models based on precision and performance were selected for validation.

Results—21/70 patients (30%) achieved ypT0–1N0 in the development cohort versus 13/55 patients (24%) in the validation cohort. Three models (t2_dwi_pre_post, semantic_dwi_adc_pre, semantic_dwi_post) were identified with an area-under-the-curve (AUC) of 0.83 (95% CI 0.70–

Corresponding author: Karin Haustermans, Dept. Radiation Oncology, University Hospital Leuven, Herestraat 49, B-3000 Leuven, Belgium, Tel.: +32 16/346902, karin.haustermans@uzleuven.be.
*PB and AC contributed equally to the manuscript

The authors declare no potential conflicts of interest.

0.95), 0.86 (95% CI 0.75–0.98) and 0.84 (95% CI 0.75–0.94) respectively. Two models (t2_dwi_pre_post, semantic_dwi_post) validated well in the external cohort with AUCs of 0.83 (95% CI 0.70 – 0.95) and 0.86 (95% CI 0.76 – 0.97). These models however did not outperform a previously established four-feature semantic model.

Conclusion—Prediction models based on MRI radiomics non-invasively predict tumor response after chemoradiation for rectal cancer and can be used as an additional tool to identify patients eligible for an organ-preserving treatment.

Keywords

rectal cancer; magnetic resonance imaging; radiomics; response prediction

INTRODUCTION

The current standard treatment for patients with locally advanced rectal cancer (LARC) is total mesorectal excision (TME) after neoadjuvant chemoradiation (CRT) [1,2]. Despite this homogeneous treatment schedule, the response of these tumors is very heterogeneous and approximately 20% of patients will achieve a pathological complete response (pCR) defined as no residual tumor and complete nodal response as reported at histology after a standard resection [3]. Given the good oncological outcome of these patients, organ-sparing treatment approaches such as a watch-and-wait policy or a local excision could be considered [4–6]. By avoiding extensive surgery, patients are likely to develop less treatment-related symptoms and to maintain a better quality of life [7,8].

Accurate detection of a clinical complete response (cCR) as a surrogate for pCR, is essential to select patients for organ-sparing strategies. Currently, qualitative assessment of anatomical and functional T2-weighted and diffusion-weighted magnetic resonance imaging (T2-MRI and DW-MRI) followed by digital rectal examination and endoscopy is the optimal strategy to define a good response to chemoradiation [9–11]. In addition to this qualitative assessment, prediction models built on manually annotated semantic parameters on pre- and posttreatment images, such as volumetric T2-MRI parameters, DW-MRI parameters and positron-emission tomography (PET) parameters, were developed with strong predicting performance for (near-)complete response in patients with locally advanced rectal cancer [12,13]. More recently, others have explored the potential of the novel field of radiomics, i.e. the automatic extraction of numerous quantitative image features for the comprehensive quantification of tumor phenotypes [14–16], for prediction of response of patients with rectal cancer to chemoradiation, with promising results. To date however, radiomics models have not yet been combined or compared with already established semantic quantitative MRI models [17,18].

In this study, we therefore aim to improve the strategy for selection of well-responding patients with rectal cancer undergoing chemoradiation that are eligible for organ-preserving treatment strategies. First, we develop novel radiomics prediction models for rectal cancer and validate them on an external patient cohort. Additionally, we investigate the addition of radiomics features to update a previously reported four-feature semantic model to further

enhance the accurate prediction of response to chemoradiation in patients with locally advanced rectal cancer [13].

MATERIALS AND METHODS

Study design and patients

In the development cohort, data were collected in a prospective clinical trial () under approval of the institutional ethical committee after obtaining written informed consent. Between January 2012 and February 2015, eighty-five patients with LARC defined as histologically proven adenocarcinoma of the rectum, clinical stage T3–4N0 or T1–4N1–2 were consecutively included in the cohort. Patients with distant metastases, prior chemotherapy or radiotherapy for rectal cancer, previous or concurrent malignancies, and known allergies to intravenous contrast agents or other contraindications for MRI acquisition were excluded. Patients were treated with chemoradiation (45 Gy in 25 fractions of 1.8 Gy; continuous infusion of 5-Fluorouracil 225 mg/m^2/d) followed by TME after an interval of eight weeks from completion of chemoradiation. Patients received multiparametric MRI prior to chemoradiation and prior to surgery (6 weeks after chemoradiation)(Ingenia, Philips Healthcare, Best, The Netherlands).

In the validation cohort, patient data were collected in an observational prospective study under approval of the institutional ethical committee after obtainment of written informed consent. For the validation cohort, accrual of fifty-five patients was done consecutively between November 2008 and December 2011. LARC was defined as primary histologically proven adenocarcinoma of the rectum, clinical stage T4Nx or TxN2 or T3 with threatened mesorectal fascia. Exclusion criteria were non-resectable and/or metastatic disease, insufficient MR image quality and patients with a mucinous rectal tumor. All patients gave written informed consent prior to study entry. Patients received chemoradiation (50 Gy in 25 fractions of 2 Gy; capecitabin 825 mg/m^2 bid) followed by TME after an interval of seven to ten weeks from completion of chemoradiation. Multiparametric MRI was obtained prior to chemoradiation and one to two weeks prior to surgery at a 3 Tesla MRI scanner (AchievaTX, Philips Medical System, Best, the Netherlands).

In both the development and validation cohort, only patients with availability of pre-treatment and pre-surgery axial T2-weighted and DW-MRI images were included. Exclusion criteria were incomplete MRI data or insufficient MRI quality. All patients have been previously reported. This prior article dealt with the development of a semantic model whereas in this manuscript quantitative MRI features were used [13]. The study was conducted using the TRIPOD recommendations for prediction model development and validation [19].

Imaging protocol

For the development cohort, MRIs were performed at a 3T MRI scanner (Ingenia, Philips Medical System, Best, the Netherlands). The MRI protocol consisted of T2-weighted turbo spin-echo (TSE) sequences (repetition time (TR)/echo time (TE) 5000ms/95ms; acquired resolution of 0.69 × 0.82 mm^2, slice thickness 3.0 mm). DWI was performed by using a

single-shot Spin-Echo Echo Planar Imaging (ssSE-EPI) sequence (TR/TE: 5775ms/68ms; EPI factor: 63, acquired resolution of 3.02×3.08 mm^2, slice thickness 5.0 mm), with spectral attenuated inversion-recovery (SPAIR) fat suppression and isotropic diffusion weighting in three directions with six b-values: 0, 50, 100, 300, 600, 1000 sec/mm^2.

For the validation cohort, MRIs were obtained at a 3T MRI scanner (AchievaTX, Philips Medical System, Best, the Netherlands). The MRI protocol consisted of T2 weighted TSE sequence (repetition time (TR)/echo time (TE) 5632ms/120ms; acquired resolution of 0.45×0.45 mm^2, slice thickness 3.0 mm). Four patients received a 3D T2 sequence and were excluded in the analyses. DWI was performed by using a single-shot Spin-Echo Echo Planar Imaging (ssSE-EPI) sequence (TR/TE: 7600ms/63ms; EPI factor: 63, acquired resolution of 2.5×2.5 mm^2, slice thickness 4.0 mm and the number of averages was three), with spectral attenuated inversion-recovery (SPAIR) fat suppression and isotropic diffusion weighting in three directions with three b-values: 0, 200, 800 s/mm^2.

Outcome measure

The primary outcome measure was (near-)complete pathological tumor response (pCR) defined as ypT0–1N0. Histologic evaluation of the resection specimen was performed by experienced pathologists according to the method of Quirke et al. [20]. In the development cohort, two patients had clinical evidence of a clinical complete response (repeated digital rectal examination, endoscopy, MRI) with disease-free survival of more than 4 years in a watch-and-wait protocol, which was considered a surrogate endpoint for pCR.

Image selection and annotation processing

We extracted the imaging protocol descriptions from the Digital Imaging and Communications in Medicine (DICOM) header and standardized annotations for axial T2- and DW-MRI images. Next, a board-certified radiologist (VV) and a board-certified radiation oncologist (MI) manually delineated the region of interest (ROI) on the axial T2-weighted images of the development and validation cohort, respectively. Thereafter, all contours were independently revised by PB and AC and adjustments were made in agreement when deemed necessary.

For DW-MRI analyses, b0 and b800 images were selected. Due to unavailability in the development cohort, b800 images were obtained using the available b-values 0, 50, 100, 300, 600 and 1000, by ordinary least squares regression based on the linear model:

$$lnS_x = lnS_0 - x * ADC$$

where S_x is the voxel intensity for the image with b-value x. ADC maps were constructed in both cohorts using b0 and b800 values. T2-derived ROIs were transferred to b800 DW images and ADC maps after registration.

Next, 2131 image features based on intensity, shape and size, texture, and wavelet and Gabor filters were extracted for each ROI on T2-MRI images, b800 DW-MRI images and ADC maps using our image feature pipeline [21,22]. Furthermore, two volumetric and two ADC

parameters (=semantic features) used in a previously published prediction model were calculated: percentage change in tumor volumetry (Volume%), equivalent sphere diameter post-chemoradiation (Sphere_post), average ADC value post-chemoradiation (ADC_avg_post) and the ratio of average ADC before and after chemoradiation (ADCratio_avg) [13].

Statistical analysis

As our features, we derived quantitative radiomic features extracted from T2-weighted, b800 DW-MRI images, and ADC maps, pre- and post-chemoradiation, along with the four semantic features. We combined various subsets of this feature set and built models on each of them. To combine features from different modalities and time points, we concatenated the raw feature vectors for each patient. The dimensionality of the concatenated feature vectors was very large (for e.g., we had $4 \times 2131 = 8524$ features, if using two modalities and two time points) compared to the number of patients. First, each feature was standardized by removing the mean and scaling to unit variance in the development cohort – on the validation cohort, the means and standard deviations learned on the development cohort were used for standardizing features. Then, principal component analysis (PCA) was used for dimensionality reduction as to linearly combine the radiomics features of T2-MRI images, b800 DW-MRI images and ADC maps with the four semantic features. Finally, logistic regression with a least absolute shrinkage and selection operator (LASSO) was used on the reduced dimensionality features to develop prediction models combining different imaging modalities with and without the four semantic features. We tuned the following hyperparameters: the number of components in PCA, the regularization weight and the penalty in LASSO, using five-fold cross validation on the development cohort – the remaining hyperparameters were left at their default values in Python scikit-learn, version 0.19.2. Finally, the model was refit on the full development cohort using the obtained best hyperparameters. The code used in this project is publicly available at https://github.com/gevaertlab/Colorectal_cancer_radiomics.

For these models, cross-validated performance using the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) curve and the operating points including precision (positive predicting value (PPV)), sensitivity and specificity were calculated. The confidence intervals were obtained via bootstrapping. To safely predict complete response and to overcome the risk of undertreatment, models with the highest positive predictive value and additionally an AUC > 0.75 were selected for validation in the validation cohort. Lastly, model performances were compared using the Delong test with a significance level of $\alpha = 0.05$ [23].

RESULTS

Seventy out of eighty-five patients in the development cohort were available for analysis. Six patients were excluded due to lack of complete MRI data and nine due to insufficient MRI quality. For the validation cohort, data of all fifty-five patients were included. Patient demographics of both cohorts are summarized in Table 1. In both cohorts, median age at diagnosis was 64 years. All patients completed the chemoradiation schedule.

Chemoradiation was followed by surgery after a median interval of 54 days in the development cohort and 55 days in the validation cohort. A total of 21 patients (30%) and 13 patients (24%) were considered to have a ypT0–1N0 response in the development cohort and in the validation cohort respectively.

Using various subsets from the available feature set (pre and post-chemoradiation data of T2-MRI images, b800 DW-MRI images and ADC maps, along with four semantic features), forty-two different models were developed (Supplementary table 1). To safely predict complete response and to overcome the risk of undertreatment, only models with the highest precision (after refitting on the development cohort using best hyperparameters) were selected. Of these most precise models, the additional selection criterion of an AUC > 0.75 resulted in a selection of three prediction models (Figure 1A, Table 2). The first model selected was a radiomics model based on quantitative T2-MRI and b800 DW-MRI features, using both pretreatment and post-chemoradiation data (t2_dwi_pre_post). With a performance of AUC 0.83 (95% CI 0.73 – 0.93), this model predicted (near-)complete response to chemoradiation with a sensitivity of 50% and a specificity of 90% for a positive predictive value (PPV) of 67%. Next, two models combining radiomics features and the four previously defined semantic MRI parameters were selected. The first one combined the semantic parameters with pretreatment b800 and ADC maps radiomics features (semantic_dwi_adc_pre), with a predicting performance of AUC 0.86 (95% CI 0.75 – 0.98), resulting in a sensitivity of 52%, specificity of 98% and a PPV of 92%. The second model combined semantic features with post-chemoradiation b800 radiomics features (semantic_dwi_post). This model predicted response with AUC 0.84 (95% confidence interval (CI) 0.75 – 0.94), resulting in a sensitivity of 63% and specificity of 88% for a PPV of 67%. Furthermore, since the delineation method of the previously published model using only semantic MRI parameters was slightly different [13], we recalculated the performance and operating points for this model using two volumetric parameters and two ADC parameters (Supplementary Figure 1, Table 3). In the development cohort, this four-feature semantic model predicted response with an AUC of 0.86 (95% CI 0.77 – 0.95) with a sensitivity of 53%, specificity of 94% and a PPV of 77%.

Two out of three selected models validated well, thereby identifying these models as independent predictors of tumor response to chemoradiation (Figure 1B, Table 2). The radiomics model t2_dwi_pre_post had an AUC of 0.83 (95% CI 0.70 – 0.95) with a PPV of 80% for a 33% sensitivity and 97% specificity. Likewise, the combination signature semantic_dwi_post yielded an AUC of 0.86 (95% CI 0.75 – 0.97) with a PPV of 80% for a 33% sensitivity and 97% specificity. On thethe other hand, combination model based on the four manually-derived features and pretreatment imaging validated poorly, given the mean predicting performance of AUC 0.49 (95% confidence interval 0.24 – 0.74). Additionally, the previously reported validation of the four-feature semantic model was confirmed in this analysis as shown by the AUC of 0.87 (95% CI 0.76 – 0.97) with a PPV of 100% for a sensitivity of 33% and a specificity of 100%. On the validation cohort, performance did not significantly differ between the four-feature semantic model and the radiomics models, except for the semantic_dwi_adc_pre model which predicted tumour response significantly worse (p = 0.009) (supplementary table 2).

DISCUSSION

In this study we showed that MRI radiomics can be used to predict (near-)complete tumor response to chemoradiation in patients with rectal cancer. Two of the three selected models using pre- and posttreatment imaging combined or posttreatment imaging alone were externally validated, thereby making these models a tool for clinical decision making when considering organ-preserving strategies. Nevertheless, these radiomics models did not outperform a previously validated four-feature semantic model.

Using T2-MRI and DW-MRI data, we identified three models based on radiomics analyses that accurately predicted the tumor response to chemoradiation, of which two validated well in an external patient cohort: the radiomics model t2_dwi_pre_post and the semantic-radiomics combination model semantic_dwi_post. These models had performances and operating points similar to the accepted strategy to select patients with cCR based on the visual judgement of MRI combined with digital rectal examination (DRE) and endoscopy [9]. Both validated models included post-chemoradiation data, which is in compliance with previous reports that indicate the importance of posttreatment MRI findings to assess tumor response [24,25]. In contrast, pretreatment MRI radiomics did not validate well in our cohort. Also, previous reports on pretreatment MRI radiomics to predict pCR or pT-stage in patients with rectal cancer lacked validation [26–28], except for prediction of pCR after chemoradiation using pretreatment MRI intensity histogram analysis, which was validated in a vendor-independent external patient-cohort albeit with a lower AUC (0.75) than the AUCs observed in our study when also using posttreatment data [29]. We therefore believe that up to now posttreatment imaging needs to be included in a prediction model. Nevertheless, as observed from the AUCs on the validation cohort, neither the pure radiomics model (t2_dwi_pre_post; AUC = 0.83) nor the semantic-radiomics combination model (semantic_dwi_post; AUC = 0.86) outperformed the four-feature semantic model (AUC = 0.87) (Table 3, Supplementary Figure 1). Since feature derivation and model building in a radiomics workflow is an abstract technique, the use of the semantic feature model using only four T2-MRI and DW-MRI parameters therefore is more comprehensible and favorable from the clinician's point of view.

The strength of our study lies in the use of an independent patient cohort for external validation, which increases the robustness of our findings and decreases selection bias. The results of our radiomics analyses are concordant with previous reports that did not use such external validation. Although Nie et al. reported on an adequate predicting performance of T2-MRI radiomics (AUC = 0.84), the study was limited by the lack of a validation cohort [30], as was the study of Horvat et al. reporting an AUC of 0.93 for pCR prediction with T2-MRI radiomics [31]. Liu et al. were the first to validate that T2- and DW-MRI radiomics could predict pCR with an AUC of 0.98, however this was not an external cohort and selection bias could be introduced in their results as the data was collected retrospectively [32].

Notwithstanding the external validation, our study had several limitations. First, slight differences in patient and tumor characteristics and radiation dose were observed the development between the development and validation cohort, which could influence the

response outcome given the dose-response relationship in rectal cancer [33]. Nevertheless, response rates were comparable. It would however be interesting to take differences in patient and tumor characteristics into account when developing prediction models, certainly when applying novel techniques such as deep learning [34]. Second, patient numbers are relatively low, which might introduce bias in our results as the risk for overfitting increases. Also, this could influence the decrease in sensitivity observed when applying the models on the validation cohort. Nevertheless, when trying to select patients for organ-preservation strategies, it is important to aim for a high specificity in order to avoid incorrect patient selection, which in turn leads to an expected decrease in model sensitivity. Therefore, ypT0–1N0 was selected over pCR as an endpoint as the higher occurrence rate would allow for more robust prediction model development. Also, local excision is an alternative for ypT1 tumours in order to preserve the rectum after chemoradiotherapy. Additionally, we chose ypT0–1N0 as an endpoint as in the current analysis surgery was performed seven to eight weeks after chemoradiotherapy and it is known that tumor regression proceeds beyond this interval [35]. Another limitation is the potential addition of noise in the data due to the interpolation in the development cohort of b800 images between different b-value images to standardize the images. This could explain the sometimes-higher AUCs on the validation cohort. Lastly, to simplify the use of our models, ROIs were delineated on T2-MRI images and then transferred to b800 images and ADC maps. Potentially, separate ROI definition on DW-MRI could improve performance but this has to be weighed up against the extra effort to delineate this ROI. Furthermore, ROIs were delineated by only one reader. In the future, these obstacles could be overcome by the introduction of automatic delineation tools [36,37].

In conclusion, our study shows that models based on MRI radiomics that predict rectal cancer response to chemoradiotherapy can be validated in an external patient cohort. Therefore, although they do not outperform a simpler four-feature semantic MRI model, they can be used as an additional tool for clinical decision making to adequately and non-invasively select patients for organ-preserving strategies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

Abbreviations:

AUC	Area Under the ROC Curve
cCR	Clinical Complete Remission
CRT	Chemoradiotherapy/Chemoradiation
LARC	Locally Advanced Rectal Cancer
pCR	Pathological Complete Remission
PPV	Positive Predictive Value
TME	Total Mesorectal Excision

REFERENCES

[1]. Kapiteijn E, Marijnen CA, Nagtegaal ID, Putter H, Steup WH, Wiggers T, et al. Preoperative Radiotherapy combined with Total Mesorectal Excision for Resectable Rectal Cancer. N Engl J Med 2001;345:638–46. doi:10.1056/NEJMoa010580. [PubMed: 11547717]

[2]. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: Results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. J Clin Oncol 2012;30:1926–33. doi:10.1200/JCO.2011.40.1836. [PubMed: 22529255]

[3]. Maas M, Beets-Tan RGH, Lambregts DMJ, Lammering G, Nelemans PJ, Engelen SME, et al. Wait-and-See Policy for Clinical Complete Responders After Chemoradiation for Rectal Cancer. J Clin Oncol 2011;29:4633–40. doi:10.1200/JCO.2011.37.7176. [PubMed: 22067400]

[4]. Habr-Gama A, Perez RO, Nadalin W, Sabbaga J, Ribeiro U, Silva e Sousa AH, et al. Operative Versus Nonoperative Treatment for Stage 0 Distal Rectal Cancer Following Chemoradiation Therapy. Ann Surg 2004;240:711–8. doi:10.1097/01.sla.0000141194.27992.32. [PubMed: 15383798]

[5]. Maas M, Nelemans PJ, Valentini V, Das P, Rödel C, Kuo L-J, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. Lancet Oncol 2010;11:835–44. doi:10.1016/S1470-2045(10)70172-8. [PubMed: 20692872]

[6]. Dossa F, Chesney TR, Acuna SA, Baxter NN. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2017;2:501–13. doi:10.1016/S2468-1253(17)30074-2. [PubMed: 28479372]

[7]. Marijnen CAM, Kapiteijn E, Van de Velde CJH, Martijn H, Steup WH, Wiggers T, et al. Acute side effects and complications after short-term preoperative radiotherapy combined with total mesorectal excision in primary rectal cancer: Report of a multicenter randomized trial. J Clin Oncol 2002;20:817–25. doi:10.1200/JCO.20.3.817. [PubMed: 11821466]

[8]. Hupkens BJP, Maas M, Martens MH, van der Sande ME, Lambregts DMJ, Breukink SO, et al. Organ Preservation in Rectal Cancer After Chemoradiation: Should We Extend the Observation Period in Patients with a Clinical Near-Complete Response? Ann Surg Oncol 2018;25:197–203. doi:10.1245/s10434-017-6213-8. [PubMed: 29134378]

[9]. Maas M, Lambregts DMJ, Nelemans PJ, Heijnen LA, Martens MH, Leijtens JWA, et al. Assessment of Clinical Complete Response After Chemoradiation for Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI. Ann Surg Oncol 2015;22:3769–71. doi:10.1093/annonc/mdv223. [PubMed: 26130453]

[10]. Fusco R, Petrillo M, Granata V, Filice S, Sansone M, Catalano O, et al. Magnetic resonance imaging evaluation in neoadjuvant therapy of locally advanced rectal cancer: A systematic review. Radiol Oncol 2017;51:252–62. doi:10.1515/raon-2017-0032. [PubMed: 28959161]

[11]. Joye I, Deroose CM, Vandecaveye V, Haustermans K. The role of diffusion-weighted MRI and 18F-FDG PET/CT in the prediction of pathologic complete response after radiochemotherapy for rectal cancer: A systematic review. Radiother Oncol 2014;113:158–65. doi:10.1016/j.radonc. 2014.11.026. [PubMed: 25483833]

[12]. Joye I, Debucquoy A, Deroose CM, Vandecaveye V, Cutsem E Van, Wolthuis A, et al. Quantitative imaging outperforms molecular markers when predicting response to chemoradiotherapy for rectal cancer. Radiother Oncol 2017;124:104–9. doi:10.1016/j.radonc. 2017.06.013. [PubMed: 28647399]

[13]. Bulens P, Couwenberg A, Haustermans K, Debucquoy A, Vandecaveye V, Philippens M, et al. Development and validation of an MRI-based model to predict response to chemoradiotherapy for rectal cancer. Radiother Oncol 2018;126:437–42. doi:10.1016/j.radonc.2018.01.008. [PubMed: 29395287]

[14]. Aerts HJWL Velazquez ER,Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4644. doi:10.1038/ncomms5006.

[15]. Bakr S, Echegaray S, Shah R, Kamaya A, Louie J. Noninvasive radiomics signature based on quantitative analysis of computed tomography images as a surrogate for microvascular invasion in hepatocellular carcinoma: a pilot study. J Med Imaging 2017;4:1–8. doi:10.1117/1.JMI. 4.4.041303.

[16]. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom KW, et al. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. Am J Neuroradiol 2018;39:208–16. doi:10.3174/ajnr.A5391. [PubMed: 28982791]

[17]. Gevaert O, Echegaray S, Khuong A, Hoang CD, Shrager JB, Jensen KC, et al. Predictive radiogenomics modeling of EGFR mutation status in lung cancer. Sci Rep 2017;7:1–8. doi: 10.1038/srep41674. [PubMed: 28127051]

[18]. Zhou M, Leung A, Echegaray S, Gentles A, Schrager JB, Jensen KC, et al. Non-small cell lung cancer radiogenomics Map identifies relationships between molecular and imaging phenotypes with prognostic implications. Radiology 2018;286:307–15. doi:10.1148/radiol.2016152244. [PubMed: 28727543]

[19]. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. Eur Urol 2015;67:1142–51. doi:10.1016/j.eururo.2014.11.025. [PubMed: 25572824]

[20]. Quirke P, Dixon MF, Durdey P, Williams NS. Local Recurrence of Rectal Adenocarcinoma Due To Inadequate Surgical Resection. Lancet 1986;328:996–9. doi:10.1016/S0140-6736(86)92612-7.

[21]. Gevaert O, Xu J, Hoang CD, Leung AN, Xu Y, Quon A, et al. Non-Small Cell Lung Cancer: Identifying Prognostic Imaging Biomarkers by Leveraging Public Gene Expression Microarray Data - Methods and Preliminary Results. Radiology 2012;264:387–96.doi:10.1148/radiol. 12111607/-/DC1. [PubMed: 22723499]

[22]. Gevaert O, Mitchell LA, Achrol AS, Xu J, Echegaray S, Steinberg GK, et al. Glioblastoma Multiforme: Exploratory Radiogenomic Analysis by Using Quantitative Image Features. Radiology 2014;273:168–74. doi:10.1148/radiol.14131731. [PubMed: 24827998]

[23]. Delong ER, Delong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics 1988;44:837–45. [PubMed: 3203132]

[24]. Intven M, Reerink O, Philippens MEP. Diffusion-weighted MRI in locally advanced rectal cancer: Pathological response prediction after neo-adjuvant radiochemotherapy. Strahlentherapie Und Onkol 2013;189:117–22. doi:10.1007/s00066-012-0270-5.

[25]. Lambregts DMJ, Rao S-X, Sassen S, Martens MH, Heijnen L a, Buijsen J, et al. MRI and Diffusion-Weighted MRI Volumetry for Identification of Complete Tumor Responders After Preoperative Chemoradiotherapy in Patients With Rectal Cancer: A Bi-institutional Validation Study. Ann Surg 2015;262:1034–9. doi:10.1097/SLA.0000000000000909. [PubMed: 25211270]

[26]. Cui Y, Yang X, Shi Z, Yang Z, Du X, Zhao Z, et al. Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced

Rectal Cancer. Eur Radiol 2019;29:1211–20. doi:10.1158/1078-0432.CCR-17-1038. [PubMed: 30128616]

[27]. Sun Y, Hu P, Wang J, Shen L, Xia F, Qing G, et al. Radiomic features of pretreatment MRI could identify T stage in patients with rectal cancer: Preliminary findings. J Magn Reson Imaging 2018;48:615–21. doi:10.1002/jmri.25969.

[28]. Giannini V, Mazzetti S, Bertotto I, Chiarenza C, Cauda S, Delmastro E, et al. Predicting locally advanced rectal cancer response to neoadjuvant therapy with 18F-FDG PET and MRI radiomics features. Eur J Nucl Med Mol Imaging 2019:[Epub ahead of print]. doi:10.1007/s00259-018-4250-6.

[29]. Dinapoli N, Barbaro B, Gatta R, Chiloiro G, Casà C, Masciocchi C, et al. Magnetic Resonance, Vendor-independent, Intensity Histogram Analysis Predicting Pathologic Complete Response After Radiochemotherapy of Rectal Cancer. Int J Radiat Oncol Biol Phys 2018;102:765–74. doi:10.1016/j.ijrobp.2018.04.065. [PubMed: 29891200]

[30]. Nie K, Shi L, Chen Q, Hu X, Jabbour S, Yue N, et al. Rectal Cancer: Assessment of Neoadjuvant Chemo-Radiation Outcome Based on Radiomics of Multi-Parametric MRI. Clin Cancer Res 2016;22:5256–64. doi:10.1158/1078-0432.ccr-15-2997. [PubMed: 27185368]

[31]. Horvat N, Veeraraghavan H, Khan M, Blazic I, Zheng J, Capanu M, et al. MR Imaging of Rectal Cancer: Radiomics Analysis to Assess Treatment Response after Neoadjuvant Therapy. Radiology 2018;287:833–43. doi:10.1148/radiol.2018172300. [PubMed: 29514017]

[32]. Liu Z, Zhang X-Y, Shi Y-J, Wang L, Zhu H-T, Tang Z, et al. Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. Clin Cancer Res 2017;23:7253–62. [PubMed: 28939744]

[33]. Appelt AL, Pløen J, Vogelius IR, Bentzen SM, Jakobsen A. Radiation dose-response model for locally advanced rectal cancer after preoperative chemoradiation therapy. Int J Radiat Oncol Biol Phys 2013;85:74–80. doi:10.1016/j.ijrobp.2012.05.017. [PubMed: 22763027]

[34]. Bibault J-E, Giraud P, Durdux C, Taieb J, Berger A, Coriat R, et al. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. Sci Rep 2018;8:12611. doi:10.1038/s41598-018-30657-6. [PubMed: 30135549]

[35]. Sloothaak DAM, Geijsen DE, van Leersum NJ, Punt CJA, Buskens CJ, Bemelman WA, et al. Optimal time interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer (Br J Surg 2013; 100: 933–939). Br J Surg 2013;100:933–9. doi:10.1002/bjs.9130. [PubMed: 23536485]

[36]. Echegaray S, Gevaert O, Shah R, Kamaya A, Louie J, Kothary N, et al. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. J Med Imaging 2015;2:1–10. doi:10.1117/1.JMI.2.4.041011.

[37]. Wang S, Zhou M, Liu Z, Liu Z, Gu D, Zang Y, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Med Image Anal 2017;40:172–83. doi:10.1016/j.media.2017.06.014. [PubMed: 28688283]

HIGHLIGHTS

- Prediction models based on quantitative imaging using multiparametric MRI can validated were to predict the response of patients with rectal cancer to chemoradiation. These models can either be based on radiomics parameters only (AUC = 0.83) or a combination of radiomics and semantic MRI features (AUC = 0.86).

- The newly developed radiomics-based models do not outperform a relatively simple four-feature semantic MRI model (AUC = 0.87).

- MRI-based models provide the potential for non-invasive selection of patients with complete response following chemoradiation for locally advanced rectal cancer, eligible for an organ-preserving treatment. These findings can be used to tailor the treatment for individual patients with rectal cancer.
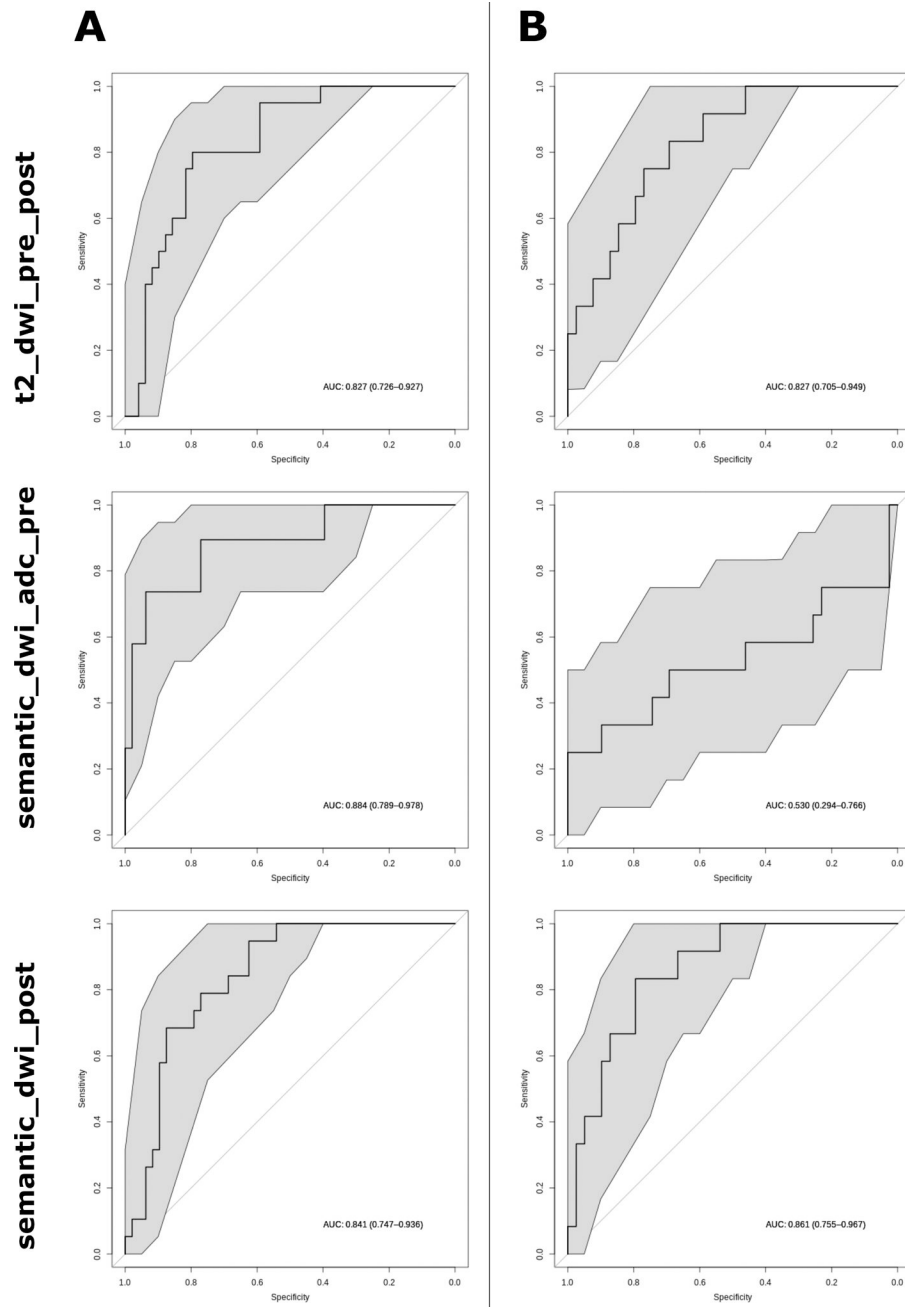
Figure 1. Receiver operating characteristic curves for the three radiomics models on both the training (A) and the validation (B) cohort.
Grey area represents the 95% confidence interval.

Table 1:

Patient characteristics, Surgery and Response Outcome.

Data are presented as n (%) or median (range). Abbreviations: APR = Abdominoperineal Resection; W&W = Watch and Wait; TEM = Transanal Endoscopic Microsurgery; pCR = pathologic Complete Response.

		Development cohort (n = 70)	Validation cohort (n = 55)
Age (years)		64 (56–69)	64 (56–70)
Sex			
	Female	23 (33%)	13 (24%)
	Male	47 (67%)	42 (76%)
Clinical tumor stage			
	T2	12 (17%)	0 (0%)
	T3	55 (79%)	46 (84%)
	T4	3 (4%)	9 (16%)
Clinical nodal stage			
	N0	3 (4%)	5 (9%)
	N1	19 (27%)	15 (27%)
	N2	48 (69%)	35 (64%)
Interval to Surgery (days)		54 (49–56)	55 (48–60)
Type of surgery			
	Sphincter-sparing	64 (91%)	36 (65%)
	APR	4 (6%)	18 (33%)
	W&W	2 (3%)	0 (0%)
	TEM	0 (0%)	1 (2%)
pCR (ypT0N0)			
	No	58 (83%)	47 (85%)
	Yes	12 (17%)	8 (15%)
Near-pCR (ypT0–1N0)			
	No	49 (70%)	42 (76%)
	Yes	21 (30%)	13 (24%)

Table 2:

Performance and operating points of the selected models that predict response to chemoradiation for patients with rectal cancer, for both the development and the validation cohort.

Data between brackets represents the 95% confidence interval. Abbreviations: AUC = Area Under the Curve; PPV = Positive Predictive Value; t2_dwi_pre_post = prediction model based on T2-weighted and b800 images pre- and post-chemoradiation, semantic_dwi_adc_pre = prediction model based on the four semantic features and b800 images and ADC maps pre-chemoradiation; semantic_dwi_post = prediction model based on the four semantic features and b800 images post-chemoradiation.

		Development Cohort	Validation Cohort
t2_dwi_pre_post			
	AUC	0.83(0.73–0.93)	0.83(0.70–0.95)
	PPV	67%	80%
	Sensitivity	50%	33%
	Specificity	90%	97%
	Accuracy	79%	82%
semantic_dwi_adc_pre			
	AUC	0.86(0.75–0.98)	0.49(0.24–0.74)
	PPV	92%	0%
	Sensitivity	52%	0%
	Specificity	98%	100%
	Accuracy	85%	75%
semantic_dwi_post			
	AUC	0.84(0.75–0.94)	0.86(0.76–0.97)
	PPV	67%	80%
	Sensitivity	63%	33%
	Specificity	88%	97%
	Accuracy	81%	82%

Table 3:

Performance and operating points of the four-feature semantic model that predicts response to chemoradiation for patients with rectal cancer, for both the development and the validation cohort.

Data between brackets represents the 95% confidence interval. Abbreviations: AUC = Area Under the Curve; PPV = Positive Predictive Value.

	Development Cohort	Validation Cohort
semantic		
AUC	0.86(0.77–0.95)	0.87(0.76–0.97)
PPV	77%	100%
Sensitivity	53%	33%
Specificity	94%	100%
Accuracy	82%	84%