

RESEARCH

Open Access



Simulation CT-based radiomics for prediction of response after neoadjuvant chemo-radiotherapy in patients with locally advanced rectal cancer

Pierluigi Bonomo¹, Jairo Socarras Fernandez², Daniela Thorwarth², Marta Casati³, Lorenzo Livi¹, Daniel Zips⁴ and Cihan Gani^{4*}

Abstract

Background: To report on the discriminative ability of a simulation Computed Tomography (CT)-based radiomics signature for predicting response to treatment in patients undergoing neoadjuvant chemo-radiation for locally advanced adenocarcinoma of the rectum.

Methods: Consecutive patients treated at the Universities of Tübingen (from 1/1/07 to 31/12/10, explorative cohort) and Florence (from 1/1/11 to 31/12/17, external validation cohort) were considered in our dual-institution, retrospective analysis. Long-course neoadjuvant chemo-radiation was performed according to local policy. On simulation CT, the rectal Gross Tumor Volume was manually segmented. A feature selection process was performed yielding mineable data through an in-house developed software (written in Python 3.6). Model selection and hyper-parametrization of the model was performed using a fivefold cross validation approach. The main outcome measure of the study was the rate of pathologic good response, defined as the sum of Tumor regression grade (TRG) 3 and 4 according to Dworak's classification.

Results: Two-hundred and one patients were included in our analysis, of whom 126 (62.7%) and 75 (37.3%) cases represented the explorative and external validation cohorts, respectively. Patient characteristics were well balanced between the two groups. A similar rate of good response to neoadjuvant treatment was obtained in both cohorts (46% and 54.7%, respectively; $p = 0.247$). A total of 1150 features were extracted from the planning scans. A 5-metafeature complex consisting of Principal component analysis (PCA)-clusters (whose main components are LHL Grey-Level-Size-Zone: Large Zone Emphasis, Elongation, HHH Intensity Histogram Mean, HLL Run-Length: Run Level Variance and HHH Co-occurrence: Cluster Tendency) in combination with 5-nearest neighbour model was the most robust signature. When applied to the explorative cohort, the prediction of good response corresponded to an average Area under the curve (AUC) value of 0.65 ± 0.02 . When the model was tested on the external validation cohort, it ensured a similar accuracy, with a slightly lower predictive ability (AUC of 0.63).

Conclusions: Radiomics-based, data-mining from simulation CT scans was shown to be feasible and reproducible in two independent cohorts, yielding fair accuracy in the prediction of response to neoadjuvant chemo-radiation.

*Correspondence: Cihan.Gani@med.uni-tuebingen.de

⁴ Department of Radiation Oncology, Medical Faculty and University Hospital, Eberhard Karls University Tübingen, Tübingen, Germany
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords: Radiomics, Simulation computed tomography, Rectal cancer, Pathologic response, Radiotherapy, Chemotherapy

Background

Preoperative radiation therapy (RT) is the mainstay of multidisciplinary treatment [1] in locally advanced rectal cancer (LARC). A very high rate ($\geq 90\%$) of sustained loco-regional control is obtained with neoadjuvant concurrent chemo-radiotherapy (CRT) followed by total mesorectal excision (TME) [2]. After surgery, a pathologic complete response (pCR) can be found in 10% to 25% of specimens. A long-term survival benefit [3] was reported for patients for whom no viable tumor cells were detectable following treatment. Thus, pCR is considered a surrogate marker for favorable outcome. In view of its positive prognostic impact but also of the morbidity [4] commonly associated with standard trimodality treatment, a growing interest emerged [5] in pursuing organ preservation strategies. Indirectly, no significant disease-free survival difference was shown between patients with a complete clinical remission on a “watch and wait” policy [6] and those with a proven pCR after TME. Conventional magnetic resonance (MR) imaging plays a pivotal role in staging and prognostication [7] of LARC, however it is characterized by very limited sensitivity [8] in assessing minimal residual disease after CRT. In this perspective, suboptimal diagnostic accuracy was also reported for functional imaging modalities, such as ^{18}F -fluorodeoxyglucose (FDG) Positron Emission Tomography/Computed tomography (PET-CT) and diffusion-weighted MR [9]. The inability to predict the response to standard preoperative treatment is a major limitation in clinical practice. The lack of biomarkers allowing for personalized radiation oncology [10] is an unmet need in LARC. Radiomics is a complex process [11] that consists of the high-throughput extraction of multidimensional features from images, their conversion into mineable data and ensuing support for clinical decision-making. In LARC, radiomics is still in its infancy. Promisingly, the additive value of quantitative data analysis combined with clinical information in identifying disease remission was reported in preliminary studies [12–22] centered on MR imaging. A major advantage of a CT-based radiomics approach is the standardized nature of the imaging information (Hounsfield unit). However, far less information are available in respect to CT-based radiomics in rectal cancer [23–28]. We therefore aimed to evaluate the potential accuracy of a radiation planning CT-based radiomics signature in the prediction of response to neoadjuvant CRT in patients with LARC.

Methods

Patients’ and treatment characteristics

We performed a dual-institution, observational, retrospective study. Consecutive patients treated for histologically-confirmed, locally advanced adenocarcinoma of the rectum at the Universities of Tübingen (TU) and Florence (FL) in two subsequent time frames (1/1/07–31/12/10 and 1/1/11–31/12/17 in TU and FL, respectively) were considered for our analysis. In general, staging included Gadolinium-enhanced pelvic MR, iodinated contrast-enhanced CT of the chest and abdomen, and colonoscopy. Clinical stage was defined according to UICC/TNM 7th edition. After multidisciplinary discussion, all patients with UICC stage II/III rectal cancer deemed amenable to undergo a full course of pre-operative radiation-based treatment followed by curatively-intended surgery could be included in our study. No tumor upper distance limit from the anal verge was specified. The primary tumor location was identified based on pelvic MR, in accordance with consensus definition [29]. No upper age limit was defined. Neoadjuvant chemotherapy, unresectable primary tumor, previous RT to the pelvis or previous surgical manipulation of the rectum were exclusion criteria. In addition, patients with unrecognizable rectal Gross Tumor Volume (GTV) on the planning CT or with image artifacts induced by hip prosthesis or rectal stent could not be included. As per local practice, standard of care for neoadjuvant treatment differed between the two centers. In TU, 50.4 Gy were delivered in 28 fractions of 1.8 Gy each (5 fractions per week). Radio-sensitizing chemotherapy consisted of 120-h continuous infusion of 5-fluorouracil during the first and fifth weeks of radiation (daily dose of 1000 mg/m^2 on days 1 through 5 and 29 through 33, respectively). Selected patients also received deep regional hyperthermia within a clinical trial, which was administered with a Sigma Eye or Sigma-60 applicator up to twice weekly for at least 60 min to a target temperature of 40.5°C , as previously described [30–32]. In FL, a total dose of 45 Gy was delivered with standard fractionation over 5 weeks (25 fractions of 1.8 Gy per day). Chrono-modulated capecitabine was prescribed for the whole RT course at a daily dose of 825 mg/m^2 BID. After restaging and due time interval (usually 6 to 10 weeks from the end of CRT), surgery was performed. Abdomino-perineal resection or rectal anterior resection (RAR) with TME were the procedures of choice. Pathologic response evaluation was assessed in accordance with Dworak’s tumor regression grade (TRG) [33] in

both institutions. Dworak's 5-point scale was as follows: 0, 1, 2, 3 and 4 scores were indicative of no regression, predominantly tumor with significant fibrosis and/or vasculopathy, predominantly fibrosis with scattered tumor cells, only scattered tumor cells in the space of fibrosis with/without acellular mucin, and no vital tumor cells detectable, respectively. Capecitabine or 5-fluorouracil (plus folinic acid)—based, adjuvant chemotherapy was offered to selected patients in case of unfavorable pathologic findings.

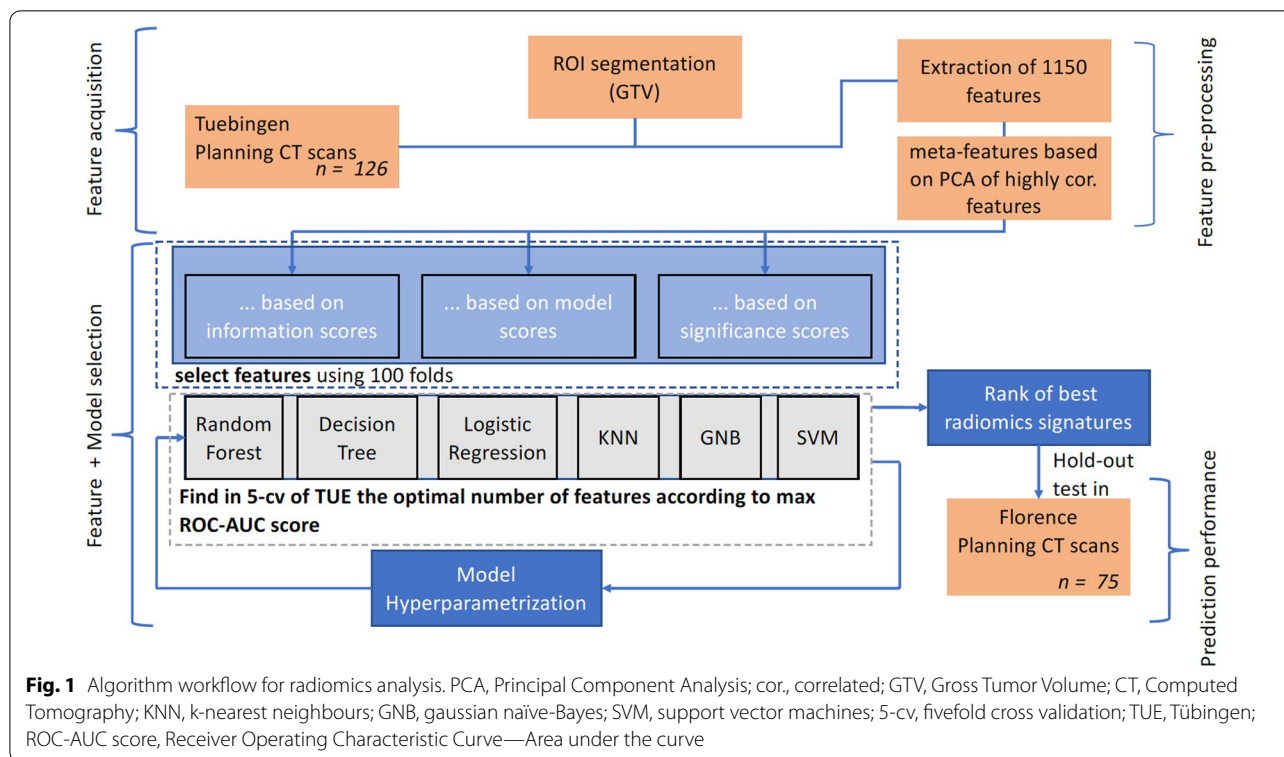
Imaging analysis and radiomics protocol

In terms of CT acquisition, treatment planning and delivery, the following procedures were performed, according to local standard of practice. A CT scan (Big Bore, Philips Medical Systems, Cleveland, OH, USA) was acquired at 3 mm slice thickness for planning purpose. The same CT model was used in both institutions. Most patients were immobilized in the prone position with an ankle-holder. In order to displace the small bowel loops from the irradiation field, a belly board device was used. RT was delivered by a linear accelerator (Elekta, Crawley, UK) with standard 3-field box technique or intensity modulated radiotherapy. In terms of delineation, the same procedures were followed in both institutions. The following organs at risk (OAR's) were contoured: femoral heads, bladder, small bowel, penile bulb and anal canal (if not infiltrated). Typically, the clinical target volume (CTV) consisted of the mesorectum and internal iliac, pre-sacral and obturator lymph nodes. In the definition of CTV, no volume modulation was used. For the purpose of this study all primary tumors were manually segmented by either of two experienced radiation oncologists (CG and PB) in a blinded fashion. For selected cases, such as those with difficult visualization of the GTV, a consensus segmentation between the two physicians was performed. Staging MR T2-weighted sequences was used to aid target definition. Imaging characteristics such as intensity distributions, texture patterns, shape features and wavelets (coiflet 1) kernel based features were extracted from planning CTs of both institutions through volume-averaged and voxelized methods. Features definitions were obtained from the Imaging Biomarker Standardisation Initiative (IBSI) [34]. For the texture features, we used the grey-level co-occurrence (GLCM), grey-level run length (GRLM), neighbourhood grey tone difference (NGTDM), grey-level size zone (GLZSM) and grey-level distance zone (GLDZM) matrix. They were computed in 3 dimensions regardless of differences between in-plane and in-slice voxel dimensions. One level undecimated wavelet features were obtained as follows. Firstly, the original images were filtered using high (H) or low-pass (L) "Coiflet 1" filter in every image (x, y, z) direction.

Different filter combinations resulted in 8 filtered images. Subsequently, intensity and texture features were computed for each filtered image [35]. All filtering and feature computations were implemented in-house in Python 3.6. Several of the radiomics features described by the IBSI are highly correlated and therefore redundant. Hence, in the training phase, we clustered correlated features (more than 95% correlation in Pearson correlation coefficient), in order to optimise the feature selection process. To do so, features were first scaled according to the quartile range (interquartile range, IQR), which ranges between the first quartile (25% quantile) and the third quartile (75% quantile). This was performed to avoid strong influence of noisy observations (for instance to imaging artefacts). Then, they were clustered hierarchically according to Pearson correlation coefficient. Finally, every cluster was reduced to one single feature using principal component analysis (PCA) to conserve the maximum possible variance inside the cluster [36]. Moreover, all features with variance lower than 0.3 were excluded from the final feature set. A feature selection process was performed whereby features with low correlation were excluded, highly correlated features were reduced to a single meta-feature and several feature selection algorithms were applied, yielding mineable data through Python 3.6. After feature selection, model hyperparameters such as the number of neighbours for Random Forest (RF) were optimised using grid search (Additional file 1 for full radiomics protocol) and fivefold cross validation. All methods and algorithms were implemented in-house in Python 3.6 using the packages *Pandas*, *Scikit-learn* and *mlxtend* for machine learning. A schematic overview of the algorithmic workflow used in this study is represented in Fig. 1.

Outcome measures and statistical analysis

After treatment, all patients were followed-up in accordance with international guidelines [37]. The main outcome measure of the study was the rate of pathologic good response (GR), defined as the sum of TRG 3 and 4 according to Dworak's classification. Baseline demographics, patients' characteristics and treatment features were summarized using descriptive statistics. Continuous variables (medians) were analyzed with Mann-Whitney test, while Fisher's exact chi-square test was employed for categorical variables. A p value < 0.05 was considered statistically significant. In order to assess the predictive power of the radiomics signature in estimating the development of GR, Receiver Operating Characteristic (ROC) curves were generated to calculate sensitivity, specificity and area under the curve (AUC) values. All statistical analyses were performed by using the statistical software SPSS (SPSS Inc, Chicago, IL, USA) for Windows (version



22). ROC-AUC curves were extracted with support of the Python 3.6 package scikit-learn and matplotlib.

Results

A total of 222 imaging datasets of patients treated for LARC in TU and FL in the considered time frame were delineated. Ten patients from TU and eleven from FL were excluded due to poor visibility of the primary tumor on CT. Two-hundred and one patients with visible tumors complying with our inclusion criteria were included in our analysis. A total of 126 (62.7%) and 75 (37.3%) of them represented the training and external validation cohorts, respectively. Patients’ characteristics are shown in Table 1. All patients had UICC/TNM stage II or III LARC.

With the exception of a slightly younger median age in the German cohort (63 vs 67 years; $p=0.042$) the remaining patient and tumor features were not significantly different between the two groups. Of note, 78 out of 126 subjects (61.9%) from TU underwent deep regional hyperthermia combined with standard CRT. All included patients underwent surgery at a median interval of 7 weeks (IQR, 6–9) from the end of CRT. A longer median waiting time to intervention was observed in the Italian cohort (9 vs 6 weeks in TU, $p<0.0001$). In terms of response to treatment, the pCR and GR rates in the overall, TU and FL cohorts were 12.4%, 12.6%, 12%, and

49.2%, 46% and 54.7%, respectively (Table 2). Overall, 201 image datasets from planning CTs were studied and 1150 features were extracted. The most robust 5-metafeature signature consisted of PCA-clusters characterized by LHL Grey-Level-Size-Zone: Large Zone Emphasis (Wavelet-texture family), Elongation (Shape family), HHH Intensity Histogram Mean (wavelet-intensity family), HLL Run-Length: Run Level Variance (wavelet texture family) and HHH Co-occurrence: Cluster Tendency (Wavelet—texture family) in combination with 5-nearest neighbour model. When applied to the explorative cohort, the prediction of GR was corresponding to an average AUC value of 0.65 ± 0.02 in 5 cross validation approach. When the model was tested on the external validation cohort, it ensured a similar accuracy, with a slightly lower prediction ability (AUC of 0.63) (Fig. 2).

Discussion

To the best of our knowledge, with just over 200 patients our study represents the largest planning CT-based radiomics investigation involving two independent institutions for the prediction of pathologic response after neoadjuvant treatment in rectal cancer. Taking into account the well-known inability of standard clinical parameters in anticipating the response of LARC to neoadjuvant treatment [5], retaining a similar predictive accuracy when shifting from internal to

Table 1 patient characteristics

| Characteristic | No. of patients (%), n = 201 | Tübingen cohort (%), n = 126 | Florence cohort (%), n = 75 | p value |
|--|------------------------------|------------------------------|-----------------------------|---------|
| Median age | | | | |
| years (IQR) | 65 (56–72) | 63 (53.7–71.2) | 67 (58–74) | 0.041 |
| Sex | | | | |
| Male | 133 (66.1%) | 81 (64.3%) | 52 (69.3%) | 0.538 |
| Female | 68 (33.9%) | 45 (35.7%) | 23 (30.7%) | |
| Staging (TNM/AJCC 7th ed.) | | | | |
| cT2N1/N2 | 11 (5.5%) | 7 (5.5%) | 4 (5.3%) | 0.655 |
| cT3N0 | 41 (20.4%) | 22 (17.5%) | 19 (25.3%) | |
| cT3N1/N2 | 127 (63.2%) | 81 (64.3%) | 46 (61.4%) | |
| cT4anyN | 22 (10.9%) | 16 (12.7%) | 6 (8%) | |
| Primary location | | | | |
| Low rectum | 89 (44.3%) | 50 (39.7%) | 39 (52%) | 0.148 |
| Middle rectum | 101 (50.2%) | 70 (55.5%) | 31 (41.3%) | |
| High rectum | 11 (5.5%) | 6 (4.8%) | 5 (6.7%) | |
| Median time interval (end of RT- surgery) weeks (IQR) | 7 (6–9) | 6 (5–7) * | 9 (8–11) | <0.0001 |

IQR, Interquartile range; RT, Radiotherapy; *: 9 missing values; °: N1 included N1a, N1b, N1c; N2 included N2a and N2b

Table 2 pattern of response to neoadjuvant therapy

| Characteristic | No. of patients (%), n = 201 | Tübingen cohort (%), n = 126 | Florence cohort (%), n = 75 | p value |
|----------------|------------------------------|------------------------------|-----------------------------|---------|
| TRG | | | | |
| 0 | 2 (1%) | 2 (1.5%) | 0 (0%) | 0.459 |
| 1 | 31 (15.5%) | 18 (14.4%) | 13 (17.3%) | |
| 2 | 69 (34.3%) | 48 (38.1%) | 21 (28%) | |
| 3 | 74 (36.8%) | 42 (33.4%) | 32 (42.7%) | |
| 4 | 25 (12.4%) | 16 (12.6%) | 9 (12%) | |
| GR | | | | |
| TRG 3 + 4 | 99 (49.2%) | 58 (46%) | 41 (54.7%) | 0.247 |

TRG, Tumor Regression Grade; GR, Good Response

external validation underlines the potential generalizability of our hypothesis-generating data in the clinic. With radiomics, it is hypothesized that selected quantitative features might represent imaging biomarkers [38, 39] able to meaningfully provide information on the prognosis of the disease or the prediction of response to therapies, on top of genomic and metabolic factors. By focusing on simulation CT, we sought to address whether an integral imaging component of radiation workflow may be informative of patient response to treatment and be associated with a favorable phenotype. MR imaging has undisputed relevance in the management of rectal cancer, representing the gold standard modality [40] in this disease. In a

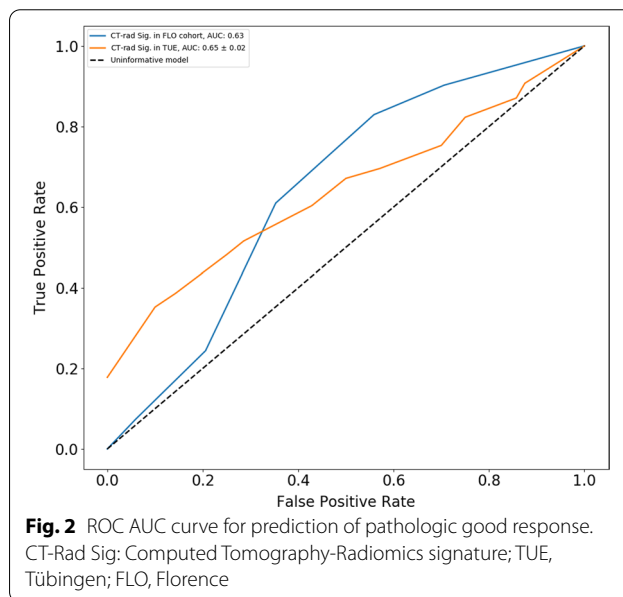


Fig. 2 ROC AUC curve for prediction of pathologic good response. CT-Rad Sig: Computed Tomography-Radiomics signature; TUE, Tübingen; FLO, Florence

single-center, retrospective experience on 48 patients, Nie et al. [12] were the first to show that an artificial neural network approach applied to multi-parametric MR images was able to significantly improve their predictive value of response to CRT, in comparison with conventional parameters. In a more recent multi-center study on 226 cases, Dinapoli and colleagues [14] demonstrated that a histogram-based radiomics signature could be associated with the development of pCR and

that the model was equally informative and reproducible in 3 independent cohorts, independently from the type of MR scanner used. Taking all data on MR-radiomics together [12–22], albeit promising, the large heterogeneity and complexity observed in terms of imaging standardization [41] restrain from direct clinical application in LARC.

In colo-rectal cancer, CT-based radiomics modelling was thus far explored in regards to preoperative stage discrimination [42], identification of pathologic lymph nodes [43], and grading differentiation [44]. To the best of our knowledge, only few groups [23–28] focused on the potential predictive value in respect to pathologic response to treatment. In a single-center, retrospective analysis [23] on 95 patients who received neoadjuvant CRT, homogeneous texture features (≤ 6.7 for entropy, ≥ 0.0118 for uniformity, and ≤ 28.06 for standard deviation) extracted from diagnostic, contrast-enhanced CT scans correlated with better disease-free survival. Vandendorpe et al. [24] performed a similar explorative analysis on baseline staging CTs of 121 patients, split between a training and validation cohorts (79 and 42 subjects, respectively). Of note, 84.3% of the whole sample was selected from a single institution. Texture analysis was based on a single slice of portal-phase images where the ROI was delineated by only one radiologist. The model showed good discriminative ability to predict a downstaging response in the training cohort (AUC of 0.90, 95% CI 0.83–0.97) which however did not hold in the test data set (AUC of 0.70, 95% CI 0.48–0.92). The use of simulation CT for radiomics modelling was previously described in a single experience. Bibault and colleagues [25] performed a high-dimensional quantitative analysis by creating a deep neural network (DNN) based on a combination of clinical information (primary tumor stage) and 28 features. The latter were extrapolated from planning CT scans of 95 patients with LARC treated in 3 different institutions. Through their innovative approach, the authors showed that the tested DNN yielded an 80% accuracy in the prediction of pCR, with a mean AUC value of 0.72 (95% CI 0.65–0.87). Other than powerful calculation algorithms, the methodological quality of investigations on radiomics must be considered of utmost importance, particularly when—as in ours and many previously published experiences—no “ground truth” correspondence between imaging and biology is available. For this purpose, Lambin and colleagues [45] proposed a composite metrics, the so called “quality radiomics score” (QRS), to evaluate the overall quality of a radiomic workflow. Secondly only to a prospective study design registered in a clinical trial database, the presence of validation in at least two distinct datasets from independent institutions is considered the single most important factor, with a score of 4 out of a total of 36 points. When taking into account the TRIPOD

statement [46] as a tool to assess the value of predictive models, our work should be regarded as a type 3 study, ranking therefore among the most reliable investigations. In the context of quantitative imaging analysis, the use of a highly standardized imaging modality such as simulation CT is also something worth highlighting. In view of its inherent reproducibility, a potential cross-validation among radiation oncology centers from different countries may not represent a critical issue as in the case of more complex diagnostic procedures, such as MR or FDG PET-CT. In this perspective, planning CT scans could well be viewed as truly “theragnostic” [47, 48] images characterized by reduced inter-operator variability, widespread accessibility, and cost-effectiveness. Pending extensive validation, the versatility of CT-based radiomics could pave the way for longitudinal analysis [49], integration of 3D dose distribution [50–52] (or dosiomics) and more complex applications in the frame of hybrid [53, 54] machines. In analogy with previous published experiences [23, 25], the presence of radiomics features portending lower heterogeneity in addition to an elongated shape in our 5-metafeature combination may be indicative of better response to neoadjuvant treatment. In view of very similar results in both TU and FL cohorts, pointing towards its reproducibility, the fair predictive accuracy of GR of the identified radiomics signature ($AUC > 0.6 < 0.7$) could depend on several reasons. In our retrospective study, the presence of unrecognized confounding variables can't be overlooked. In particular, factors such as treatment heterogeneity, waiting time to surgery [55], different RT total dose and type of systemic agents may have well influenced the pattern of response after treatment. However, patient features were not significantly different between the two cohorts, with the exception of a slightly younger median age in the German group (63 vs. 67 years; $p = 0.042$). A longer median waiting time to intervention was observed in the Italian cohort (9 vs. 6 weeks in Tuebingen, $p < 0.0001$). The shorter time interval in the Tuebingen cohort may be due to the fact that patients were treated in an older period compared with those in the Florence one, mirroring a different attitude towards what has become part of routine practice more recently. However, still today the interplay between pathologic response after neoadjuvant treatment and waiting time to surgery is among the most controversial research topics in rectal cancer, without definitive answers [56, 57]. Both median time intervals in our study groups fell within the acceptable standard of care, and no significant difference was reported in terms of pCR and GR between the two cohorts. Deep regional hyperthermia is a well known strategy with potential radiosensitizing effect, with relatively limited applications in the clinic in experienced centers. Albeit promising, relatively limited prospective data are available in rectal cancer

[5]. A recently published single-arm, prospective phase 2 trial [32] showed optimal outcome and compliance by adding hyperthermia to concomitant CRT in patients with LARC, with an overall pCR rate of 14%, very similar to what was found in our work. Although we can't rule out that the use of hyperthermia in 61.9% of patients from the Tuebingen cohort may have contributed to a better response, a marked unbalance on our results is unlikely, in our opinion. Overall, no clinical variables emerged in our model as significant, so that no stratification for T and N stage was performed. Further limitations have to be acknowledged in the interpretation of our results. First, the relatively small sample size and the reported pCR rate at the lower end of the expected range limit the strength of our findings. In view of the small proportion of patients with pCR in our study, we decided to cluster patients with TRG 3 and 4. However, since TRG 3 is indicative of only few scattered tumor cells left in the specimen, the biological meaning of it can be very similar to TRG 4. In the first ever published experience on radiomics in rectal cancer [12] a similar approach was applied to discriminate good responders from the rest. Second, the intrinsically low contrast resolution of native CTs and the overall image quality may have hampered the accuracy of GTV segmentation in some challenging cases, in spite of our attempt to solve discrepancies by consensus between the two radiation oncologists involved in target delineation. In this perspective, tools for semi-automatic segmentation may provide benefit, particularly for the definition of cranio-caudal extension [58]. Third, tumor volume was not predictive of GR in our study and thus excluded from the final model; however, it cannot be overlooked that our radiomics signature could be just a surrogate of tumor volume [59] and as such associated with pathologic response, as already shown in other investigations with similar discriminative power. Fourth, when dealing with the "classical" rule of thumb introduced by Hosmer and Lemeshow [60] on how to score the discriminative ability of a ROC curve, a minimum AUC score of 0.7 defines the lower end of what can provide an acceptable discrimination. However, in a hypothesis-generating study such as ours even an AUC value $>0.6 < 0.7$ may still be deemed fair and worthy of further investigation. For instance, to name one of the most studied examples in radiation oncology involving quantitative imaging, when applying radiomics analysis to NTCP models for the prediction of radiation pneumonitis, AUC values between 0.6 and 0.7 have been regarded as promising [61, 62]. Albeit the discriminative power by itself in our study was not very high, we believe that our CT-based radiomics model is strengthened by the independent, dual-institution validation approach we performed and very similar performance we obtained in both cohorts, which may represent a benchmark for future

investigations on simulation CT in LARC. Ultimately, recognizing the central role of MR imaging in rectal cancer and the fact that by itself CT imaging may not be ideal in terms of GTV definition, we believe that in terms of radiomics perspective, the use of simulation CT for radiomics purposes may be of extreme interest, given the fact that it is essentially imaging acquired in the treatment position, in contrast to diagnostic MR. Our study lends support to CT-based radiomics in LARC as a hypothesis-generating approach, adding value to the very limited available evidence in this topic, pending further external validation in a larger cohort of patients.

Conclusions

In our hypothesis-generating study, a process of data-mining from pre-treatment simulation CT scans was shown to be feasible and reproducible in two independent cohorts. The potential predictive ability of a CT-based signature in both datasets in identifying patients with pathologic GR after surgery warrants further investigations.

Abbreviations

CT: Computed tomography; PCA: Principal component analysis; AUC: Area under the curve; RT: Radiation therapy; LARC: Locally advanced rectal cancer; CRT: Chemo-radiotherapy; TME: Total mesorectal excision; pCR: Pathologic complete response; MR: Magnetic resonance; FDG: 18F-fluorodeoxyglucose; PET-CT: Positron Emission Tomography/Computed tomography; TU: Tuebingen; FL: Florence; UICC: Union Internationale Contre Le Cancer; GTV: Gross tumor volume; RAR: Rectal anterior resection; TRG: Tumor regression grade; OH: Ohio; UK: United Kingdom; USA: United States of America; OARs: Organs at risk; CTV: Clinical target volume; CG: Cihan Gani; PB: Pierluigi Bonomo; IBSI: Imaging Biomarker Standardisation Initiative; GLCM: Grey-level co-occurrence; GRLM: Grey-level run length; NGTDM: Neighbourhood grey tone difference; GLZSM: Grey-level size zone; GLDZM: Grey-level distance zone; H: High; L: Low; IQR: Interquartile range; RF: Random forest; GR: Good response; ROC: Receiver operating characteristic; AUC: Area under the curve; DNN: Deep neural network; QRS: Quality radiomics score; TRIPOD: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis; 3D: 3 Dimensional.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13014-022-02053-y>.

Additional file 1. Full radiomics protocol.

Acknowledgements

Not applicable.

Author contributions

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. No funding was provided.

Availability of data and materials

The present data is summarized in this paper (METHODS). The complete dataset can be retrieved from the authors upon formal request from interested readers.

Declarations

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by Medizinische Fakultät Tübingen and University of Florence. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Radiation Oncology, Azienda Ospedaliero-Universitaria Careggi, University of Florence, Florence, Italy. ²Section for Biomedical Physics, Department of Radiation Oncology, Eberhard Karls University Tübingen, Tübingen, Germany. ³Medical Physics, Azienda Ospedaliero-Universitaria Careggi, University of Florence, Florence, Italy. ⁴Department of Radiation Oncology, Medical Faculty and University Hospital, Eberhard Karls University Tübingen, Tübingen, Germany.

Received: 12 February 2022 Accepted: 11 April 2022

Published online: 28 April 2022

References

1. Van der Velde CJ, Aristei C, Boelens PG, et al. EURECCA colorectal: multi-disciplinary mission statement on better care for patients with colon and rectal cancer in Europe. *Eur J Cancer*. 2013;49:2784–90. <https://doi.org/10.1016/j.ejca.2013.04.032>.
2. Sauer R, Liersch T, Merkel S, 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. <https://doi.org/10.1200/JCO.2011.40.1836>.
3. Maas M, Nelemans PJ, Valentini V, 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. [https://doi.org/10.1016/S1470-2045\(10\)70172-8](https://doi.org/10.1016/S1470-2045(10)70172-8).
4. Sprenger T, Beissbarth T, Sauer R, et al. Long-term prognostic impact of surgical complications in the German Rectal Cancer Trial CAO/ARO/AIO-94. *Br J Surg*. 2018;105:1510–8. <https://doi.org/10.1002/bjs.10877>.
5. Gani C, Bonomo P, Zwirner K, et al. Organ preservation in rectal cancer - Challenges and future strategies. *Clin Transl Radiat Oncol*. 2017;23:9–15. <https://doi.org/10.1016/j.ctro.2017.02.002>.
6. Van der Valk MJ, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWW): an international multicentre registry study. *Lancet*. 2018;391:2537–45. [https://doi.org/10.1016/S0140-6736\(18\)31078-X](https://doi.org/10.1016/S0140-6736(18)31078-X).
7. Horvath N, Carlos Tavares Rocha C, Clemente Oliveira B, Petkovska I, Gollub MJ. MRI of rectal cancer: tumor staging, imaging techniques, and management. *Radiographics*. 2019;39:367–87. <https://doi.org/10.1148/rg.2019180114>.
8. Van der Paardt MP, Zagers MB, Beets-Tan RG, Stoker J, Bipat S. Patients who undergo preoperative chemoradiotherapy for locally advanced rectal cancer restaged by using diagnostic MR imaging: a systematic review and meta-analysis. *Radiology*. 2013;269:101–12. <https://doi.org/10.1148/radiol.13122833>.
9. Joye I, Deroose CM, Vandecaveye V, Haustermans K. The role of diffusion-weighted MRI and (18)F-FDG PET/CT in the prediction of pathologic complete response after radiochemotherapy for rectal cancer: a systematic review. *Radiother Oncol*. 2014;113:158–65. <https://doi.org/10.1016/j.radonc.2014.11.026>.
10. Zhang B, Wang J, Wang X, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014;513:382–7. <https://doi.org/10.1038/nature13438>.
11. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures. *They Are Data Radiology*. 2016;278:563–77. <https://doi.org/10.1148/radiol.2015151169>.
12. Nie K, Shi L, Chen Q, et al. Rectal cancer: assessment of neoadjuvant chemo-radiation outcome based on radiomics of multi-parametric MRI. *Clin Cancer Res*. 2016;22:5256–64. <https://doi.org/10.1158/1078-0432.CCR-15-2997>.
13. Liu Z, Zhang XY, Shi YJ, 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. <https://doi.org/10.1158/1078-0432.CCR-17-1038>.
14. Dinapoli N, Barbaro B, Gatta R, 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. <https://doi.org/10.1016/j.ijrobp.2018.04.065>.
15. Horvat N, Veeraraghavan H, Khan M, et al. MR imaging of rectal cancer: radiomics analysis to assess treatment response after neoadjuvant therapy. *Radiology*. 2018;287:833–43. <https://doi.org/10.1148/radiol.2018172300>.
16. Yi X, Pei Q, Zhang Y, et al. MRI-based radiomics predicts tumor response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Front Oncol*. 2019;9:552. <https://doi.org/10.3389/fonc.2019.00552>.
17. Cui Y, Yang X, Shi Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol*. 2019;29:1211–20. <https://doi.org/10.1007/s00330-018-5683-9>.
18. Giannini V, Mazzetti S, Bertotto I, 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;46:878–88. <https://doi.org/10.1007/s00259-018-4250-6>.
19. Bulens P, Couwenberg A, Intven M, et al. Predicting the tumor response to chemoradiotherapy for rectal cancer: model development and external validation using MRI radiomics. *Radiother Oncol*. 2019;142:246–52. <https://doi.org/10.1016/j.radonc.2019.07.033>.
20. Shaish H, Aukerman A, Vanguri R, et al. Radiomics of MRI for pretreatment prediction of pathologic complete response, tumor regression grade, and neoadjuvant rectal score in patients with locally advanced rectal cancer undergoing neoadjuvant chemoradiation: an international multicenter study. *Eur Radiol*. 2020;30:6263–73. <https://doi.org/10.1007/s00330-020-06968-6>.
21. Delli Pizzi A, Chiarelli AM, Chiacchiaretta P, et al. MRI-based clinical-radiomics model predicts tumor response before treatment in locally advanced rectal cancer. *Sci Rep*. 2021;11:5379. <https://doi.org/10.1038/s41598-021-84816-3>.
22. Coppola F, Mottola M, Lo Monaco S, et al. The heterogeneity of skewness in T2W-based radiomics predicts the response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Diagnostics (Basel)*. 2021;11:795. <https://doi.org/10.3390/diagnostics11050795>.
23. Chee CG, Kim YH, Lee KH, et al. CT texture analysis in patients with locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy: a potential imaging biomarker for treatment response and prognosis. *PLoS ONE*. 2017;12: e0182883. <https://doi.org/10.1371/journal.pone.0182883>.
24. Vandendorpe B, Durot C, Lebellec L, et al. Prognostic value of the texture analysis parameters of the initial computed tomographic scan for response to neoadjuvant chemoradiation therapy in patients with locally advanced rectal cancer. *Radiother Oncol*. 2019;135:153–60. <https://doi.org/10.1016/j.radonc.2019.03.011>.
25. Bibault JE, Giraud P, Housset M, et al. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep*. 2018;8(1):12611. <https://doi.org/10.1038/s41598-018-30657-6>.
26. Wang J, Shen L, Zhong H, et al. Radiomics features on radiotherapy treatment planning CT can predict patient survival in locally advanced

- rectal cancer patients. *Sci Rep.* 2019;9:15346. <https://doi.org/10.1038/s41598-019-51629-4>.
27. Hamerla G, Meyer HJ, Hamsch P, et al. Radiomics model based on non-contrast CT shows no predictive power for complete pathological response in locally advanced rectal cancer. *Cancers (Basel).* 2019;11:1680. <https://doi.org/10.3390/cancers11111680>.
 28. Yuan Z, Frazer M, Zhang GG, et al. CT-based radiomic features to predict pathological response in rectal cancer: a retrospective cohort study. *J Med Imaging Radiat Oncol.* 2020;64:444–9. <https://doi.org/10.1111/1754-9485.13044>.
 29. Van de Velde CJH, Boelens PG, Borras JM, et al. EURECCA colorectal: multidisciplinary management: European consensus conference colon & rectum. *Eur J Cancer* 2014; 50(1):1e1–1.e34. <https://doi.org/10.1016/j.ejca.2013.06.048>
 30. Schroeder C, Gani C, Lamprecht U, et al. Pathological complete response and sphincter-sparing surgery after neoadjuvant radiochemotherapy with regional hyperthermia for locally advanced rectal cancer compared with radiochemotherapy alone. *Int J Hyperthermia.* 2012;28:707–14. <https://doi.org/10.3109/02656736.2012.722263>.
 31. Gani C, Schroeder C, Heinrich V, et al. Long-term local control and survival after preoperative radiochemotherapy in combination with deep regional hyperthermia in locally advanced rectal cancer. *Int J Hyperthermia.* 2016;32:187–92. <https://doi.org/10.3109/02656736.2015.1117661>.
 32. Gani C, Lamprecht U, Ziegler A, et al. Deep regional hyperthermia with preoperative radiochemotherapy in locally advanced rectal cancer, a prospective phase II trial. *Radiother Oncol.* 2021;159:155–60. <https://doi.org/10.1016/j.radonc.2021.03.011>.
 33. Dworak O, Keilholz L, Hoffman A. Pathological features of rectal cancer after preoperative radiochemotherapy. *Int J Colorectal Dis.* 1997;12:19–23. <https://doi.org/10.1007/s003840050072>.
 34. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295:328–38. <https://doi.org/10.1148/radiol.2020191145>.
 35. Gonzalez RC, Woods RE, Eddins SL. *Digital Image Processing Using MATLAB®*. Gatesmark Publishing; 2009.
 36. Walker M, Kublin JG, Zunt JR. Fast R functions for robust correlations and hierarchical clustering. *J Stat Softw.* 2009;42:115–25.
 37. NCCN guidelines: Rectal Cancer, version 1.2021. Available at: <http://www.nccn.org>.
 38. Aerts HJ, Rios Velazquez E, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006. <https://doi.org/10.1038/ncomms5006>.
 39. O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol.* 2017;14:169–86. <https://doi.org/10.1038/nrclinonc.2016.162>.
 40. Lambregts DM, Maas M, Stokkel MP, Beets-Tan RG. Magnetic resonance imaging and other imaging modalities in diagnostic and tumor response evaluation. *Semin Radiat Oncol.* 2016;26:193–8. <https://doi.org/10.1016/j.semradonc.2016.02.001>.
 41. Di Re AM, Sun Y, Sundaresan P, et al. MRI radiomics in the prediction of therapeutic response to neoadjuvant therapy for locoregionally advanced rectal cancer: a systematic review. *Expert Rev Anticancer Ther.* 2021;21:425–49. <https://doi.org/10.1080/14737140.2021.1860762>.
 42. Liang C, Huang Y, He L, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage III and stage III-IV colorectal cancer. *Oncotarget.* 2016;7:31401–12. <https://doi.org/10.18632/oncotarget.8919>.
 43. Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol.* 2016;34:2157–64. <https://doi.org/10.1200/JCO.2015.65.9128>.
 44. Huang X, Cheng Z, Huang Y, et al. CT-based radiomics signature to discriminate high-grade from low-grade colorectal adenocarcinoma. *Acad Radiol.* 2018;25:1285–97. <https://doi.org/10.1016/j.acra.2018.01.020>.
 45. Lambin P, Leijenaar RT, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14:749–62. <https://doi.org/10.1038/nrclinonc.2017.141>.
 46. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Cancer.* 2015;112:251–9. <https://doi.org/10.1038/bjc.2014.639>.
 47. Ling CC, Humm J, Larson S, et al. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys.* 2000;47:551–60. [https://doi.org/10.1016/s0360-3016\(00\)00467-3](https://doi.org/10.1016/s0360-3016(00)00467-3).
 48. Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol.* 2005;6:112–7. [https://doi.org/10.1016/S1470-2045\(05\)01737-7](https://doi.org/10.1016/S1470-2045(05)01737-7).
 49. Van Timmeren JE, van Elmpst W, Leijenaar RT, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: evaluation of the added prognostic value for overall survival and locoregional recurrence. *Radiother Oncol.* 2019;136:78–85. <https://doi.org/10.1016/j.radonc.2019.03.032>.
 50. Rossi L, Bijman R, Schillema W, et al. Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in radiotherapy. *Radiother Oncol.* 2018;129:548–53. <https://doi.org/10.1016/j.radonc.2018.07.027>.
 51. Chiloiro G, Rodriguez-Carnero P, Lenkiewicz J, et al. Delta Radiomics Can Predict Distant Metastasis in Locally Advanced Rectal Cancer: The Challenge to Personalize the Cure. *Front Oncol.* 2020; 10:595012. <https://doi.org/10.3389/fonc.2020.595012>. eCollection 2020.
 52. Placidi L, Gioscio E, Garibaldi C, et al. A multicentre evaluation of dosiomics features reproducibility, stability and sensitivity. *Cancers.* 2021;13:3835. <https://doi.org/10.3390/cancers13153835>.
 53. Gani C, Boldrini L, Valentini V. Online MR guided radiotherapy for rectal cancer. New opportunities *Clin Transl Radiat Oncol.* 2019;18:66–7. <https://doi.org/10.1016/j.ctro.2019.04.005>.
 54. Boldrini L, Intven M, Bassetti M, Valentini V, Gani C. MR-guided radiotherapy for rectal cancer: current perspective on organ preservation. *Front Oncol.* 2021;11: 619852. <https://doi.org/10.3389/fonc.2021.619852>.
 55. Macchia G, Gambacorta MA, Masciocchi C, et al. Time to surgery and pathological complete response after neoadjuvant chemoradiation in rectal cancer: a population study on 2094 patients. *Clin Transl Radiat Oncol.* 2017;4:8–14. <https://doi.org/10.1016/j.ctro.2017.04.004>.
 56. Lefevre JH, Mineur L, Kotti S, et al. Effect of interval (7 or 11 weeks) between neoadjuvant radiochemotherapy and surgery on complete pathologic response in rectal cancer: a multicenter, randomized, controlled trial (GRECCAR-6). *J Clin Oncol.* 2016;34(31):3773–80. <https://doi.org/10.1200/JCO.2016.67.6049>.
 57. Gambacorta MA, Masciocchi C, Chiloiro G, et al. Timing to achieve the highest rate of pCR after preoperative radiochemotherapy in rectal cancer: a pooled analysis of 3085 patients from 7 randomized trials. *Radiother Oncol.* 2021;154:154–60. <https://doi.org/10.1016/j.radonc.2020.09.026>.
 58. Roach D, Holloway LC, Jameson MG, et al. Multi-observer contouring of male pelvic anatomy: Highly variable agreement across conventional and emerging structures of interest. *J Med Imaging Radiat Oncol.* 2019;63:264–71. <https://doi.org/10.1111/1754-9485.12844>.
 59. Palmisano A, Di Chiara A, Esposito A, et al. MRI predictor of pathological response in locally advanced rectal cancer: when apparent diffusion coefficient radiomics meets conventional volumetry. *Clin Radiol.* 2020;75:798. e1–798.e-11. <https://doi.org/10.1016/j.crad.2020.06.023>
 60. Hosmer DW, Lemeshow S. *Applied logistic regression*, second edition. Chapter 5. New York, NY: Wiley, 2000.
 61. Krafft SP, Rao A, Stingo F, et al. The utility of quantitative CT radiomics features for improved prediction of radiation pneumonitis. *Med Phys.* 2018;45(11):5317–24. <https://doi.org/10.1002/mp.13150>.
 62. Desideri I, Loi M, Francolini G, Becherini C, Livi L, Bonomo P. Application of radiomics for the prediction of radiation-induced toxicity in the IMRT era: current-state-of-the-art. *Front Oncol.* 2020;10:1708. <https://doi.org/10.3389/fonc.2020.01708>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.