HHS Public Access

Author manuscript

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2021 April 01.

Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2020 April 01; 106(5): 998–1009. doi:10.1016/j.ijrobp.2019.12.038.

Preoperative prediction of pathologic response to neoadjuvant chemoradiotherapy in patients with esophageal cancer using ^{18}F-FDG PET/CT and DW-MRI: a prospective multicenter study

Alicia S. Borggreve, MD[*,1,2], **Lucas Goense, MD PhD**[*,1,2], **Peter S.N. van Rossum, MD PhD**[1], **Sophie E. Heethuis, PhD**[1], **Richard van Hillegersberg, MD PhD**[2], **Jan J.W. Lagendijk, PhD**[1], **Marnix G.E.H. Lam, MD PhD**[3], **Astrid L.H.M.W. van Lier, PhD**[1], **Stella Mook, MD PhD**[1], **Jelle P. Ruurda, MD PhD**[2], **Marco van Vulpen, MD PhD**[1], **Francine E.M. Voncken, MD**[4], **Berthe M.P. Aleman, MD PhD**[4], **Annemarieke Bartels-Rutten, MD PhD**[5], **Jingfei Ma, PhD**[6], **Penny Fang, MD**[7], **Benjamin C. Musall, MSc**[6], **Steven H. Lin, MD PhD**[7], **Gert J. Meijer, PhD**[1]

[1]Department of Radiation Oncology, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands [2]Department of Surgery, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands [3]Department of Nuclear Medicine, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands [4]Department of Radiation Oncology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands [5]Department of Radiology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands [6]Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas, United States of America [7]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas, United States of America

Abstract

Corresponding author during review process: Alicia S. Borggreve, MD, A.S.Borggreve@umcutrecht.nl. **Corresponding author after review process (should also be addressed for reprints):** Gert J. Meijer, PhD, Address: Department of Radiation Oncology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands, G.J.Meijer@umcutrecht.nl, Telephone: +31 (0)88-7551503.

Authors responsible for statistical analyses:
Alicia S. Borggreve and Lucas Goense

*Both first authors contributed equally to this manuscript.

Conflict of interest

This research was partially funded by Elekta Inc. and by National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30CA016672.

R.v.H. and J.P.R. are proctoring surgeons for Intuitive Surgical Inc. and train other surgeons in robot-assisted minimally invasive esophagectomy. J.J.W. receives research funding from Elekta Inc. S.H.L. receives research funding from Elekta Inc., Genentech, Hitachi Chemicals, New River Labs, Beyond Spring Pharmaceuticals, and is a member of the Advisory Board of AstraZeneca. All of the above are not in conflict with the research in question. All other authors have nothing to disclose.

Trial registration: The article reports on a health care intervention on human participants and was prospectively registered on April 29, 2014 under ClinicalTrials.gov Identifier: NCT02125448

Background—Accurate preoperative prediction of pathologic response to neoadjuvant chemoradiotherapy (nCRT) in esophageal cancer patients could enable to forgo esophagectomy in patients with a pathologic complete response (pCR). This study aimed to evaluate the individual and combined value of ^{18}F-fluorodeoxyglucose positron emission tomography with integrated computed tomography (^{18}F-FDG PET/CT) and diffusion-weighted magnetic resonance imaging (DW-MRI) during and after nCRT to predict pathologic response in esophageal cancer patients.

Methods—In this multicenter prospective study, patients scheduled to receive nCRT followed by esophagectomy for esophageal cancer underwent ^{18}F-FDG PET/CT and DW-MRI scanning prior to start of nCRT, during nCRT, and before esophagectomy. Response to nCRT was based on histopathological evaluation of the resection specimen. Relative changes in ^{18}F-FDG PET/CT and DW-MRI parameters were compared between patients with pCR and non-pCR groups. Multivariable Ridge regression analyses with bootstrapped c-indices were performed to evaluate the individual and combined value of ^{18}F-FDG PET/CT and DW-MRI.

Results—pCR was found in 26.1% of 69 patients. Relative changes in ^{18}F-FDG PET/CT parameters *after* nCRT ($\Delta SUV_{mean,post}$ p=0.016, and ΔTLG_{post} p=0.024), as well as changes in DW-MRI parameters *during* nCRT (ΔADC_{during} p=0.008) were significantly different between pCR and non-pCR. A c-statistic of 0.84 was obtained for a model with ΔADC_{during}, $\Delta SUV_{mean,post}$ and histology in classifying patients as pCR (versus 0.82 for ΔADC_{during} and 0.79 for $\Delta SUV_{mean,post}$ alone).

Conclusions—Changes on ^{18}F-FDG PET/CT after nCRT and early changes on DW-MRI during nCRT can help identify pCR to nCRT in esophageal cancer. Moreover, ^{18}F-FDG PET/CT and DW-MRI might be of complementary value in the assessment of pCR.

Summary

Changes on ^{18}F-FDG PET/CT after neoadjuvant chemoradiotherapy and early changes on DW-MRI during neoadjuvant chemoradiotherapy can help identify pathologic complete responders to neoadjuvant chemoradiotherapy in esophageal cancer. Accurate response prediction would enable physicians to individualize and adapt therapy for each individual esophageal cancer patient.

Keywords

esophageal cancer; neoadjuvant chemoradiotherapy; pathologic complete response; ^{18}F-FDG PET/CT; DW-MRI

Introduction

Neoadjuvant chemoradiotherapy (nCRT) followed by surgery is the preferred treatment with curative intent for patients with locally advanced esophageal cancer.[1] Through tumor downsizing and downstaging, nCRT improves locoregional control and overall survival rates compared to surgery alone.[2–4] Many studies have reported that the degree of tumor regression in response to nCRT is directly related to long-term survival, with pathologic complete responders having the most favorable long-term prognosis.[3,5] The absence of viable tumor cells at the site of the primary tumor in pathologic complete responders raises the question whether a surgical resection is of benefit for patients who might have already

been cured locoregionally by nCRT alone.[6] Accurate prediction of pathologic complete response (pCR) before surgery could potentially allow patients to forgo surgery and would enable researchers to study the feasibility and outcome of an organ-preserving strategy after chemoradiotherapy. Conversely, patients with a poor response to nCRT are likely to benefit less or not at all from nCRT, but are exposed to its side-effects. A reliable identification of poor responders during nCRT may thus be also beneficial, as ineffective therapy could be stopped and/or alternative treatment strategies (e.g. additional neoadjuvant treatment or upfront surgery) could be explored.

Unfortunately, most studied modalities – including endoscopic biopsy, endoscopic ultrasonography (EUS) and computed tomography (CT) – yield unsatisfactory results for the evaluation of tumor response to nCRT.[7–11] Metabolic and functional imaging modalities such as 18F-fluorodeoxyglucose positron emission tomography with integrated computed tomography (^{18}F-FDG PET/CT) and diffusion-weighted magnetic resonance imaging (DW-MRI) may be more promising because they allow biological and microstructural characterization of tumors and visualization of treatment-induced changes before volumetric changes become apparent.[9,12–18] However, the discriminatory ability of ^{18}F-FDG PET/CT alone has previously shown in a multitude of studies to be insufficient to guide clinical decision-making.[14,19,20] The performance of DW-MRI and the quantitative apparent diffusion coefficient (ADC) are promising in predicting response to nCRT although they have been demonstrated only in two small single-center pilot studies.[16,17]

A multimodality imaging approach, in comparison to that of a single-modality, may provide complementary value for predicting pathologic response, with the ultimate goal of optimally guiding individualized treatment decision-making. Therefore, the aim of this prospective multicenter study was to evaluate the individual and combined value of ^{18}F-FDG PET/CT and DW-MRI during and after nCRT to predict pathologic response in patients who undergo nCRT for esophageal cancer, as well as to validate these findings for the prediction of survival.

Methods

Three high volume institutions participated in this multicenter prospective study including *Institution 1*, *Institution 2*, and *Institution 3*. The study has been approved by the institutional review board of each institution separately, and all patients provided their written informed consent. The study was registered with ClinicalTrials.gov, number NCT02125448.

Study population

Patients presenting at any of the 3 participating institutions from October 2013 to July 2017 with newly diagnosed biopsy-proven esophageal cancer who were scheduled to receive nCRT followed by surgery were eligible for inclusion. Exclusion criteria are outlined in the Supplementary Methods. The results of first 20 of the 26 included patients from *Institution 3* have been published previously.[1]

Treatment protocol

At *Institution 1* and *Institution 2*, the neoadjuvant treatment regimen consisted of carboplatin/paclitaxel with concurrent radiotherapy (41.4 Gy in 23 fractions of 1.8 Gy).[4] At *Institution 3*, the regimen consisted of 5-fluorouracil with either platinum or taxane-based chemotherapy and concurrent radiotherapy (50.4 Gy in 28 fractions of 1.8 Gy). All patients were treated with an intensity-modulated radiation therapy (IMRT) technique. At a median of 8 weeks (interquartile range [IQR]: 7–10 weeks) after completion of nCRT, patients underwent a transhiatal or transthoracic esophagectomy with two-field lymphadenectomy and gastric conduit reconstruction with either cervical or intrathoracic anastomosis, depending on patient characteristics and local preference.

Histopathological assessment

Histopathological tumor regression of the surgical resection specimen was assessed by specialized gastrointestinal pathologists who were blinded for the results of the ^{18}F-FDG PET/CT and DW-MRI scans. Patients were staged in accordance with the 7th edition of the Union for International Cancer Control (UICC).[21] The following tumor regression groups were scored: absence of residual cancer cells (TRG 1), 1–10% residual cancer cells (TRG 2), 10–50% residual cancer cells (TRG 3) and >50% residual cancer cells (TRG 4).[22]

Image acquisition

Patients underwent ^{18}F-FDG PET/CT scanning at 3 time points: within a median of 3 weeks (IQR: 2 – 4 weeks) before nCRT, during nCRT at a median of 13 days (IQR: 10 – 15 days) after initiation of nCRT, and 5 weeks (IQR: 4.5 – 7 weeks) after completion of nCRT and before surgery (median: 14 days [IQR: 9 – 32 days] between ^{18}F-FDG PET/CT scan and surgery).

Patients underwent DW-MRI scanning using 3 different b-values (b = 0, 200 and 800 s/mm^2) at the same 3 time points as the ^{18}F-FDG PET/CT scans. Only the baseline DW-MRI scan was generally not performed on the exact same day as the baseline ^{18}F-FDG PET/CT scan, since the latter was part of the diagnostic work-up and the DW-MRI scan was acquired after patients were included in the study. The baseline DW-MRI scan was in all cases acquired within 2 weeks (median: 5 days, IQR: 1 – 6 days) before nCRT. Detailed image acquisition parameters are presented in Supplementary Methods.

Image analysis

Details on the ^{18}F-FDG PET/CT and DW-MRI tumor delineation methodology are included in the Supplementary Methods. At each of the 3 time points, the following 4 quantitative metrics were extracted from the ^{18}F-FDG PET/CT primary tumor delineations: mean and maximum standardized uptake value (SUV_{mean}, SUV_{max}), metabolic tumor volume (MTV), and total lesion glycolysis (TLG) calculated as the product of MTV and SUV_{mean}.[23] From the DW-MRI primary tumor delineations, the mean ADC values were extracted from the ADC map. The relative changes in percent (Δ%) of these ^{18}F-FDG PET/CT and DW-MRI parameters between baseline scans and the scans during and after nCRT were calculated and considered as possible predictors in the analysis based on prior knowledge (i.e. ΔADC, ΔSUV_{mean}, ΔSUV_{max} and ΔTLG).[16,17,20,24–27]

Statistical analysis

The relative changes of the ^{18}F-FDG PET/CT and DW-MRI parameters were compared between patients with pCR (TRG 1) and non-pCR (TRG 2–4), and between good responders (TRG 1–2, GR) and poor responders (TRG 3–4, non-GR) to validate findings of previous pilot-studies using the Mann-Whitney U test with Benjamin-Hochberg corrections.[16,17,20,24–26,28,29] The ability of single modality 18F-FDG PET/CT and DW-MRI parameters to discriminate between different pathologic response groups was quantified using the area under the receiver operating characteristic curve (c-statistic). The complementary value of ^{18}F-FDG PET/CT and DW-MRI parameters, in addition to histopathological tumor type which known to impact pathologic response to nCRT, was assessed using multivariable penalized Ridge regression models to reduce model overfitting in a situation with few events per variable.[30] The best model for pCR prediction was validated for the prediction of overall survival (OS) and disease free survival (DFS) using multivariable Cox regression analysis. Detailed statistical methods are presented in Supplementary Methods.

Results

Study population

Between October 2013 and July 2017, a total of 82 consecutive patients with newly diagnosed esophageal cancer who underwent standard diagnostic work-up signed informed consent. A total of 13 patients were excluded from the analysis (withdrawal from study participation [n=4], unexpected distant metastatic disease [n=3], no tumor signal on the baseline ^{18}F-FDG PET/CT or DW-MRI [n=3], small tumor volume [<2ml, n=2] and refusal of surgery [n=1]). The remaining 69 patients, with a total of 203 ^{18}F-FDG PET/CT scans and 199 DW-MRI scans, were eligible for analysis (Supplementary Figure 1).

Patients had a mean age of 61.0 years (SD ± 9.2 years), and 88.4% (n=61) were male. Histologic tumor types included adenocarcinoma (n=57, 82.6%), squamous cell carcinoma (n=11, 15.9%) or undifferentiated large cell carcinoma type (n=1, 1.4%). Histopathological assessment of the surgical resection specimen demonstrated pCR (i.e. TRG 1) in 18 patients (26.1%), and GR (i.e. TRG 1–2) in 43 patients (62.3%) after nCRT (see also Supplementary Table 1). Patients with a squamous cell carcinoma had a significantly higher probability of pCR and GR compared to patients with adenocarcinoma (Table 1). No significant differences were observed in clinical T stages, nCRT regimens or time intervals from nCRT to surgery between patients with pCR and non-pCR, or between patients with GR and non-GR (Table 1). Detailed patient and tumor characteristics are outlined in Table 1.

Relative changes in ^{18}F-FDG PET/CT and DW-MRI parameters

The relative change in tumor SUV_{mean}, SUV_{max}, TLG and ADC_{mean} during and after nCRT for pCR and GR patients are presented in Table 2 and illustrated in Figure 1. In Figure 2 ^{18}F-FDG PET/CT and MRI scans of a patient with pCR are presented.

The changes in SUV_{mean}, SUV_{max} and TLG from baseline ^{18}F-FDG PET/CT scans to scans acquired during nCRT were not significantly different between patients with pCR versus

non-pCR (Table 2). Changes in SUV_{mean} and TLG from baseline ^{18}F-FDG PET/CT scans to scans acquired after completion of nCRT, however, were significantly different between pCR and non-pCR ($\Delta SUV_{mean,post}$ [median, IQR]: −63% [−68%, −49%] for pCR versus −42% [−58%, −16%] for non-pCR, p = 0.016, and ΔTLG_{post} [median, IQR]: −86% [−93%, −81%] for pCR versus −65% [−88%, −32%] for non-pCR, p = 0.024). Regarding GR, none of the changes in ^{18}F-FDG PET/CT parameters from baseline to either during or after completion of nCRT were significantly different between GR and non-GR (Table 2).

The relative increase in tumor ADC from baseline DW-MRI scans to scans acquired during nCRT (ΔADC_{during}) was significantly associated with pCR (median, IQR: 28% [15%, 39%] for pCR versus 11% [4%, 17%] for non-pCR, p = 0.008). In contrast, relative changes in ADC from baseline to scans after completion of nCRT (ΔADC_{post}) were not significantly different between pCR and non-pCR groups. These findings were consistent with the association between changes in ADC and GR versus non-GR (ΔADC_{during} [median, IQR]: 20% [13%, 33%] for GR versus 8% [0%, 11%] for non-GR, p = 0.008, and ΔADC_{post} [median, IQR]: 24% [13%, 44%] for GR versus 21% [6%, 30%] for non-GR, p = 0.201, Table 2).

Complementary value ^{18}F-FDG PET/CT and DW-MRI parameters

To evaluate the complementary value of ^{18}F-FDG PET/CT and DW-MRI parameters for pCR prediction, receiver operating characteristic (ROC) analyses for ΔADC_{during} combined with $\Delta SUV_{mean,post}$ and histology showed a superior bootstrapped c-statistic in comparison with their individual values and histology (0.83 [95% CI: 0.74 – 0.94] for ΔADC_{during} and $\Delta SUV_{mean,post}$, versus 0.81 [95% CI: 0.70 – 0.93] for ΔADC_{during} and 0.79 [95% CI: 0.67 – 0.90] for $\Delta SUV_{mean,post}$, all models with histology, Table 3 and Figure 3a). Furthermore, the model with ΔADC_{during} $\Delta SUV_{mean,post}$ and histology also demonstrated the best global model fit in terms of the lowest AIC (Supplementary Table 2). ROC curve analyses of ^{18}F-FDG PET/CT and DW-MRI parameters acquired during treatment for prediction of GR did not demonstrate complementary value of ΔADC_{during} and $\Delta SUV_{max,during}$, but demonstrated a high bootstrapped c-statistic for ΔADC_{during} with histology (0.82 [95% CI: 0.75 – 0.94] for ΔADC_{during} and $\Delta SUV_{max,during}$, versus 0.83 [95% CI: 0.73 – 0.92] for ΔADC_{during} and 0.74 [95% CI: 0.60 – 0.88] for $\Delta SUV_{max,during}$, all models with histology, Table 3 and Figure 3b). The model with $\Delta SUV_{max,during}$ and histology demonstrated the best global model fit in terms lowest AIC (Supplementary Table 2).

Model calibration plots of all 8 models as described in Table 3 are depicted in Supplementary Figure 2, and show variable calibration upon visual inspection. In general, the best model for pCR in terms of highest c-statistic and lowest AIC (the model with DW-MRI, PET-CT and histology, i.e. model 4 in Table 3) was slightly overoptimistic and tended to underestimate the observed probability of pCR, whereas the best model for GR in terms of highest c-statistic (the model with DW-MRI and histology, i.e. model 2 in Table 3) was better calibrated.

Survival

The results of the multivariable Cox regression analysis for OS and DFS including the parameters from the most accurate model for pCR prediction are depicted in Table 4. None of the included imaging parameters (ΔADC_{during} and $\Delta SUV_{mean,post}$), nor histology was significantly associated with OS or DFS, as all 95% CI of the hazard ratio's crossed the null hypothesis.

Discussion

This international multicenter, prospective study was designed to assess the predictive value of quantitative changes on ^{18}F-FDG PET/CT and DW-MRI scans acquired during and after nCRT in patients with esophageal cancer. Changes in ^{18}F-FDG PET/CT parameters *after* nCRT ($\Delta SUV_{mean,post}$ and ΔTLG_{post}), as well as changes in DW-MRI parameters *during* nCRT (ΔADC_{during}) demonstrate to discriminate well between pathologic complete responders (TRG 1) and non-pCR (TRG 2–4) in esophageal cancer. Moreover, ^{18}F-FDG PET/CT and DW-MRI might be of complementary value in the assessment of histopathological response. However, the prediction model including both imaging features did not correspond with OS or DFS in a subsequent experimental analysis on long-term clinical outcomes.

This study provides encouraging results for the potential value of multimodal imaging in discrimination of patients likely to have pCR to nCRT. This is of important clinical value, as accurate response prediction would enable physicians to individualize and adapt therapy for each individual patient. Considering the postoperative morbidity and mortality associated with surgery, and the effect of surgery on quality of life, an organ-sparing approach might improve outcomes for patients with pCR.[31–33] Whereas in patients with a poor response to nCRT, alternative treatment strategies could be explored such as intensification or additional neoadjuvant treatment, or ineffective therapy could be stopped. For this latter purpose, a model was developed based on changes in ^{18}F-FDG PET/CT and DW-MRI parameters *during* nCRT, with GR (TRG 1–2) as outcome variable, demonstrating a high predictive performance of a DW-MRI parameter and tumor histology (c-statistic 0.83).

Until now, most imaging studies in esophageal response prediction have focused on a single modality. Two meta-analyses on the value of ^{18}F-FDG PET/CT suggested that the decrease in mean or maximum metabolic activity within the first two weeks of nCRT are so far the best available predictor of treatment response.[14,19] However, the conclusion of these meta-analyses, including a more recent extensive analysis, is that the discriminatory ability of ^{18}F-FDG PET/CT was insufficient for clinical decision-making.[20] The results of this study are consistent with these findings, with a maximum c-statistic of 0.74 for an ^{18}F-FDG PET/CT imaging parameter (ΔSUV_{mean} after nCRT). Our study results are also consistent with the findings of more recent studies that showed no significant associations between relative changes ^{18}F-FDG PET/CT imaging parameters for pCR, or only for ΔTLG_{during}.[17,27]

A pilot study using DW-MRI reported that the change in the tumor ADC during the first 2–3 weeks of nCRT for esophageal cancer seemed highly predictive for pCR.[16] These findings have recently been externally validated in another single-center pilot study[17], but have not

yet been validated in a larger cohort. The current, larger and multicenter study confirms these findings, reporting the highest c-statistic for prediction of pCR and GR for ΔADC_{during} (0.77 and 0.84, respectively). Addition of ^{18}F-FDG PET/CT parameters to the DW-MRI model raised the ability to discriminate between pCR (TRG 1) and non-pCR (TRG 2–4) in 83% of the patients (c-statistic: 0.83).

As an illustration, the patient with an squamous cell carcinoma – whose scans are depicted in Figure 2 – and a ΔADC_{during} of 25% and $\Delta SUV_{mean,post}$ of −79%, has a predicted probability of 0.69 for a pCR as based on the model including ^{18}F-FDG PET/CT, DW-MRI parameters and histology. A patient with an adenocarcinoma and ΔADC_{during} of 10% and $\Delta SUV_{mean,post}$ of −20%, would have only had a predicted probability 0.12 for pCR based on the model. It is unclear yet what predictive power is required for clinical decision making as this might vary between patients. It could well be argued that for high risk surgical patients with significant comorbidities, a lower accuracy of a predicted pCR might be deemed sufficient to postpone standard esophagectomy, whereas for younger, low risk surgical patients, a higher accuracy might be desired. Other considerations to be taken into account are how many patients will require salvage surgery, and how many patients will have missed an opportunity for cure because surgery was initially avoided, as death of locoregional disease will be considered as the ultimate failure of therapy. Eventually, future improvement in prognostication will likely result in stratification of patients by risk of surgery balanced with the predicted probability of pCR.

Other strategies to identify good responders are currently subject of research in multiple trials. The recently published preSANO study focused on accurate detection of locoregional residual disease rather than pathologic complete response, and reported favorable results for response evaluations using EUS, bite-on-bite biopsies, fine-needle aspiration of suspicious lymph nodes and ^{18}F-FDG PET/CT.[34] Furthermore, it has been recently recognized that the combination of clinical and imaging parameters with molecular biomarkers might be the key to success. Besides molecular markers — such as ALDH1 and GLI1/HH — which might predict response before nCRT, other molecular markers, such as circulating tumor DNA (ctDNA), that evaluate response after completion of nCRT should be further explored.[35,36] A large Dutch multicenter study is currently investigating the additional value of ctDNA to ^{18}F-FDG PET/CT and MRI for accurate pCR prediction (PRIDE study).[37]

Since ^{18}F-FDG PET/CT and DW-MRI scanning during and after completion of nCRT are currently not part of standard imaging evaluation, the predictive value of multimodal imaging evaluation should be considered in light of the associated costs and physical burden to the patients of repeated imaging procedures. Previous studies have demonstrated that ^{18}F-FDG PET/CT and MRI are generally well-tolerated imaging procedures for the assessment of response to treatment in esophageal cancer patients.[38] Furthermore, long-term treatment outcomes (survival and long-term quality of life) appear to outweigh short-term attributes (short-term quality of life and burden of surveillance examinations with endoscopy and PET/CT) in another study.[39] In this perspective, further investigation of multimodal strategies for reliable response prediction to nCRT seems justified.

Several challenges exist before routine ^{18}F-FDG PET/CT and DW-MRI scanning for treatment response assessment can be implemented for clinical practice. First, quantitative imaging values such as SUV_{max} from ^{18}F-FDG PET/CT imaging are influenced by vendor specific characteristics, which differ between the scanner platforms. However, since our study only included relative changes in the analyses, the impact of these differences on the results is limited. Standardization of the DW-MRI protocols is more challenging than for ^{18}F-FDG PET/CT protocols, as they are comprised of more nuanced sequences that are more likely to differ between institutions.[40,41] The MRI scan protocol applied in the current study was clinically developed specifically for esophageal cancer patients, and extensive efforts have been undertaken to synchronize the protocols before initiation of the study. Even though standardization of MRI protocols appears to be challenging, previous studies have found similar changes in mean tumor ADC during nCRT in esophageal cancer to be predictive of pCR, despite major differences in scanner platforms (3.0T versus 1.5T, GE versus Philips).[16,17] Furthermore, our study included both scanner types and found consistent results with the earlier studies.

Second, it became apparent during the assessment of the DW-MRI scans in our study that there is still room for improvement in the quality of these scans. The thorax is a challenging region for MR imaging, given the differences in magnetic susceptibility between the organs of the thorax (e.g., pulmonary parenchyma, vertebrae and ribs), as well as the effects of cardiac and respiratory motion.[42] As such, MR imaging in the thoracic region remains under continuous technical development, and better imaging quality may have the potential to further improve the performance of DW-MRI for treatment response assessment.[43]

Third, even though semi-automatic contouring methods were used, differences between contouring techniques and inter-reader variability could still be an issue. However, the reproducibility of ADC measurements between readers has been studied previously, and was considered reproducible among 5 readers in a relatively small dataset of 20 esophageal cancer patients.[17] These high correlations were achieved when ADC values were assessed with whole volume tumor ADC measurements, as in our study. As semi-automatic tumor delineation on ^{18}F-FDG PET/CT and DW-MRI scans remains a time-consuming process, the clinical applicability of ^{18}F-FDG PET/CT and DW-MRI for response assessment might benefit from a validated fully automatic contouring tool.

Fourth, the prediction model including both imaging features that provided good performance for prediction of pCR did not correspond with OS or DFS in a subsequent experimental analysis. We believe that the survival analyses in the current study should be interpreted with caution, as these survival analyses were not predefined and hence may suffer from a lack of power. Also, it is essential to first develop a clinically useful prediction model to identify complete responders, before drawing any conclusions regarding survival. Especially since subsequent analyses with survival will only make sense if identification of pCR is accurate and treatment is altered based on these predictions, as the burden of 3 additional MRI’s and 1 additional PET/CT (as in the current study) can only be justified if this results in treatment alterations, and not just for prognostic purposes.

Lastly, limitations of the current study include the absence of an external validation cohort to validate our results and the various nCRT regimens applied. However, consistent findings in a slightly heterogeneous, multicenter study population confers external validity to the results of the study (Supplementary Table 1). Other strengths of the current study include the availability of histopathological evaluation as a reference standard, instead of surrogate endpoints such as radiological response criteria (RECIST), and the inclusion of histologic subtypes in the multivariable analyses.

In conclusion, our study shows that quantitative ADC changes from baseline to interim DW-MRI scans and SUV_{mean} changes from baseline to follow-up ^{18}F-FDG PET/CT scans can help identify pCR to nCRT in esophageal cancer patients. However, additional larger prospective studies, as well as other combined multimodal approaches are needed to validate these results, especially regarding the potentially complementary value of ^{18}F-FDG PET/CT and DW-MRI imaging parameters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lordick F, Mariette C, Haustermans K, et al. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016;27:v50–v57. [PubMed: 27664261]
2. Sjoquist KM, Burmeister BH, Smithers BM, et al. Survival after neoadjuvant chemotherapy or chemoradiotherapy for resectable oesophageal carcinoma: an updated meta-analysis. Lancet Oncol. 2011;12:681–92. [PubMed: 21684205]
3. Shapiro J, Van Lanschot JJB, Hulshof MCCM, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. Lancet Oncol. 2015;16:1090–1098. [PubMed: 26254683]
4. van Hagen P, Hulshof MCCM, van Lanschot JJB, et al. Preoperative Chemoradiotherapy for Esophageal or Junctional Cancer. N Engl J Med. 2012;366:2074–2084. [PubMed: 22646630]
5. Steffen T, Dietrich D, Schnider A, et al. Recurrence Patterns and Long-Term Results After Induction Chemotherapy, Chemoradiotherapy, and Curative Surgery in Patients With Locally Advanced Esophageal Cancer. Ann Surg. 2019;269:83–87. [PubMed: 28742685]
6. Noordman BJ, Wijnhoven BPL, Lagarde SM, et al. Active surveillance in clinically complete responders after neoadjuvant chemoradiotherapy for esophageal or junctional cancer. Dis Esophagus. 2017;30:1–8.
7. Westerterp M, van Westreenen HL, Reitsma JB, et al. Esophageal Cancer: CT, Endoscopic US, and FDG PET for Assessment of Response to Neoadjuvant Therapy—Systematic Review. Radiology. 2005;236:841–851. [PubMed: 16118165]
8. van Rossum PSN, Goense L, Meziani J, et al. Endoscopic biopsy and EUS for the detection of pathologic complete response after neoadjuvant chemoradiotherapy in esophageal cancer: a systematic review and meta-analysis. Gastrointest Endosc. 2016;83:866–879. [PubMed: 26632523]
9. Wang L, Liu L, Han C, et al. The diffusion-weighted magnetic resonance imaging (DWI) predicts the early response of esophageal squamous cell carcinoma to concurrent chemoradiotherapy. Radiother Oncol. 2016;121:246–251. [PubMed: 27838148]
10. Yip C, Cook GJR, Landau DB, et al. Performance of different imaging modalities in assessment of response to neoadjuvant therapy in primary esophageal cancer. Dis Esophagus. 2016;29:116–130. [PubMed: 25604614]

11. Ngamruengphong S, Sharma VK, Nguyen B, et al. Assessment of response to neoadjuvant therapy in esophageal cancer: an updated systematic review of diagnostic accuracy of endoscopic ultrasonography and fluorodeoxyglucose positron emission tomography. Dis Esophagus. 2010;23:216–31. [PubMed: 19515185]
12. Koh D-M, Collins DJ. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. Am J Roentgenol. 2007;188:1622–1635. [PubMed: 17515386]
13. Lambrecht M, Vandecaveye V, De Keyzer F, et al. Value of Diffusion-Weighted Magnetic Resonance Imaging for Prediction and Early Assessment of Response to Neoadjuvant Radiochemotherapy in Rectal Cancer: Preliminary Results. Int J Radiat Oncol. 2012;82:863–870.
14. Kwee RM. Prediction of Tumor Response to Neoadjuvant Therapy in Patients with Esophageal Cancer with Use of ^{18} F FDG PET: A Systematic Review. Radiology. 2010;254:707–717. [PubMed: 20177086]
15. Aoyagi T, Shuto K, Okazumi S, et al. Apparent diffusion coefficient values measured by diffusion-weighted imaging predict chemoradiotherapeutic effect for advanced esophageal cancer. Dig Surg. 2011;28:252–7. [PubMed: 21654173]
16. van Rossum PSN, van Lier ALHMW, van Vulpen M, et al. Diffusion-weighted magnetic resonance imaging for the prediction of pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer. Radiother Oncol. 2015;115:163–170. [PubMed: 26002307]
17. Fang P, Musall BC, Son JB, et al. Multimodal Imaging of Pathologic Response to Chemoradiation in Esophageal Cancer. Int J Radiat Oncol Biol Phys. 2018;102:996–1001. [PubMed: 29685377]
18. De Cobelli F, Giganti F, Orsenigo E, et al. Apparent diffusion coefficient modifications in assessing gastro-oesophageal cancer response to neoadjuvant treatment: Comparison with tumour regression grade at histology. Eur Radiol. 2013;23:2165–2174. [PubMed: 23588582]
19. Chen Y-M, Pan X, Tong L-J, et al. Can 18F-fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. Nucl Med Commun. 2011;32:1005–10. [PubMed: 21886014]
20. van Rossum PSN, Fried DV, Zhang L, et al. The Incremental Value of Subjective and Quantitative Assessment of 18F-FDG PET for the Prediction of Pathologic Complete Response to Preoperative Chemoradiotherapy in Esophageal Cancer. J Nucl Med. 2016;57:691–700. [PubMed: 26795288]
21. Rice TW, Blackstone EH, Rusch VW. 7th Edition of the AJCC Cancer Staging Manual: Esophagus and Esophagogastric Junction. Ann Surg Oncol. 2010;17:1721–1724. [PubMed: 20369299]
22. Chirieac LR, Swisher SG, Ajani JA, et al. Posttherapy pathologic stage predicts survival in patients with esophageal carcinoma receiving preoperative chemoradiation. Cancer. 2005;103:1347–1355. [PubMed: 15719440]
23. Roedl JB, Harisinghani MG, Colen RR, et al. Assessment of treatment response and recurrence in esophageal carcinoma based on tumor length and standardized uptake value on positron emission tomography-computed tomography. Ann Thorac Surg. 2008;86:1131–8. [PubMed: 18805147]
24. Li Q-W, Qiu B, Wang B, et al. Prediction of pathologic responders to neoadjuvant chemoradiotherapy by diffusion-weighted magnetic resonance imaging in locally advanced esophageal squamous cell carcinoma: a prospective study. Dis Esophagus. 2018;31(2).
25. Tan S, Kligerman S, Chen W, et al. Spatial-Temporal [18F]FDG-PET Features for Predicting Pathologic Response of Esophageal Cancer to Neoadjuvant Chemoradiation Therapy. Int J Radiat Oncol. 2013;85:1375–1382.
26. Beukinga RJ, Hulshoff JB, Mul VEM, et al. Prediction of Response to Neoadjuvant Chemotherapy and Radiation Therapy with Baseline and Restaging 18 F-FDG PET Imaging Biomarkers in Patients with Esophageal Cancer. Radiology. 2018;287:983–992. [PubMed: 29533721]
27. Elimova E, Wang X, Etchebehere E, et al. 18-fluorodeoxy-glucose positron emission computed tomography as predictive of response after chemoradiation in oesophageal cancer patients. Eur J Cancer. 2015;51:2545–2552. [PubMed: 26321501]
28. van Heijl M, Omloo JM, van Berge Henegouwen MI, et al. Fluorodeoxyglucose Positron Emission Tomography for Evaluating Early Response During Neoadjuvant Chemoradiotherapy in Patients With Potentially Curable Esophageal Cancer. Ann Surg. 2011;253:56–63. [PubMed: 21233607]
29. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc. 1995;57:289–300.

30. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. BMJ. 2015;351:h3868. [PubMed: 26264962]
31. Low DE, Kuppusamy MK, Alderson D, et al. Benchmarking Complications Associated with Esophagectomy. Ann Surg. 2019;269:291–298. [PubMed: 29206677]
32. Kauppila JH, Johar A, Lagergren P. Postoperative Complications and Health-related Quality of Life 10 Years After Esophageal Cancer Surgery. Ann Surg. Epub ahead of print 7 10, 2018.
33. Noordman BJ, Van Klaveren D, van Berge Henegouwen MI, et al. Impact of Surgical Approach on Long-term Survival in Esophageal Adenocarcinoma Patients With or Without Neoadjuvant Chemoradiotherapy. Ann Surg. 2017;267:892–897.
34. Noordman BJ, Spaander MCW, Valkema R, et al. Detection of residual disease after neoadjuvant chemoradiotherapy for oesophageal cancer (preSANO): a prospective multicentre, diagnostic cohort study. Lancet Oncol. 2018;19:965–974. [PubMed: 29861116]
35. Creemers A, Ebbing EA, Pelgrim TC, et al. A systematic review and meta-analysis of prognostic biomarkers in resectable esophageal adenocarcinomas. Sci Rep. 2018;8:13281. [PubMed: 30185893]
36. Ajani JA, Bhutani MS, Swisher SG. Oesophageal preservation in locally advanced oesophageal cancer. Lancet Oncol. 2018;19:e430. [PubMed: 30191840]
37. Borggreve AS, Mook S, Verheij M, et al. Preoperative image-guided identification of response to neoadjuvant chemoradiotherapy in esophageal cancer (PRIDE): a multicenter observational study. BMC Cancer. 2018;18:1006. [PubMed: 30342494]
38. Goense L, Borggreve AS, Heethuis SE, et al. Patient perspectives on repeated MRI and PET/CT examinations during neoadjuvant treatment of oesophageal cancer. Br J Radiol. 2018;91:20170710. [PubMed: 29498535]
39. Noordman BJ, de Bekker-Grob EW, Coene PPLO, et al. Patients' preferences for treatment after neoadjuvant chemoradiotherapy for oesophageal cancer. Br J Surg. 2018;105:1630–1638. [PubMed: 29947418]
40. Verma V, Simone CB, Krishnan S, et al. The Rise of Radiomics and Implications for Oncologic Management. J Natl Cancer Inst. 2017;109(7).
41. deSouza NM. Diffusion-weighted MRI in Multicenter Trials of Breast Cancer: A Useful Measure of Tumor Response? Radiology. 2018;289:628–629. [PubMed: 30179102]
42. Andre JB, Bammer R. Advanced diffusion-weighted magnetic resonance imaging techniques of the human spinal cord. Top Magn Reson imaging. 2010;21:367–78. [PubMed: 22158130]
43. van Rossum PS, van Lier AL, Lips IM, et al. Imaging of oesophageal cancer with FDG-PET/CT and MRI. Clin Radiol. 2015;70:81–95. [PubMed: 25172205]

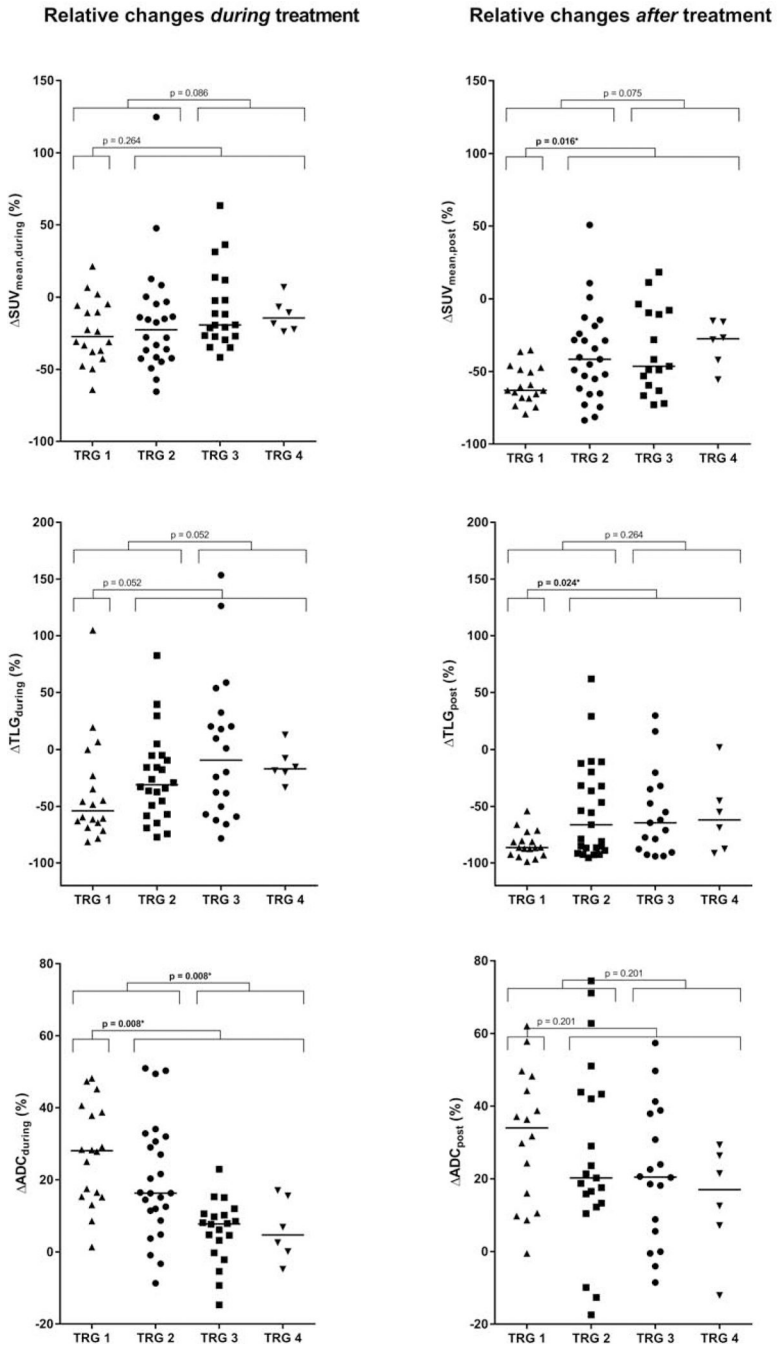

Figure 1.

Relative changes in ^{18}F-FDG PET/CT and DW-MRI parameters during and after neoadjuvant chemoradiotherapy in all patients, categorized by histopathologic tumor regression grades (TRG). The solid horizontal lines represent group medians. Significant differences between the groups based on the Mann-Whitney U test followed by Benjamini-Hochberg adjustment are marked with an asterisk (*).

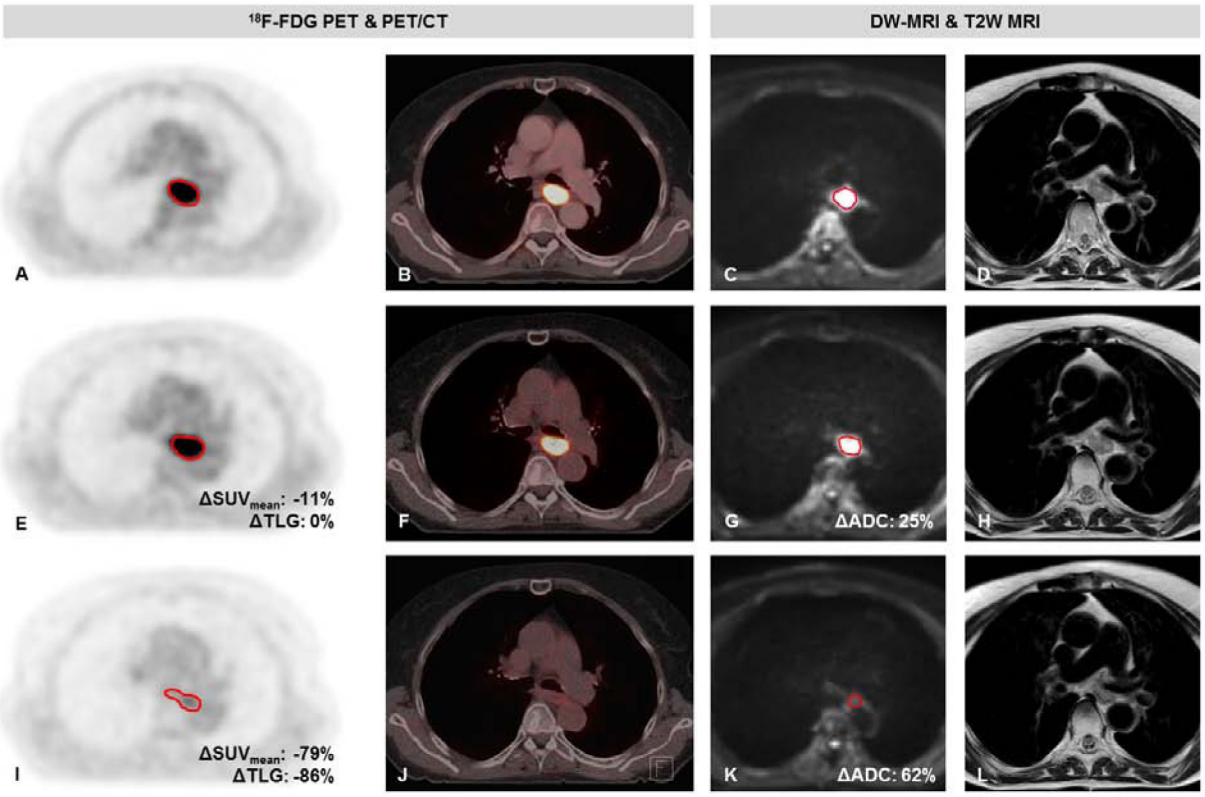

Figure 2.

Patient with a cT3N2M0 mid-esophageal squamous cell carcinoma with a pathologic complete response to neoadjuvant chemoradiotherapy (TRG 1). ^{18}F-FDG PET images (A, E, I), fused PET/CT images (B, F, J), diffusion-weighted images (b-value = 200 s/mm2) (C, G, K) and T2 weighted images (D, H, L) on a 1.5T MR scanner before neoadjuvant chemoradiotherapy (nCRT) (A-D), during nCRT (E-H), and after nCRT (I-L).

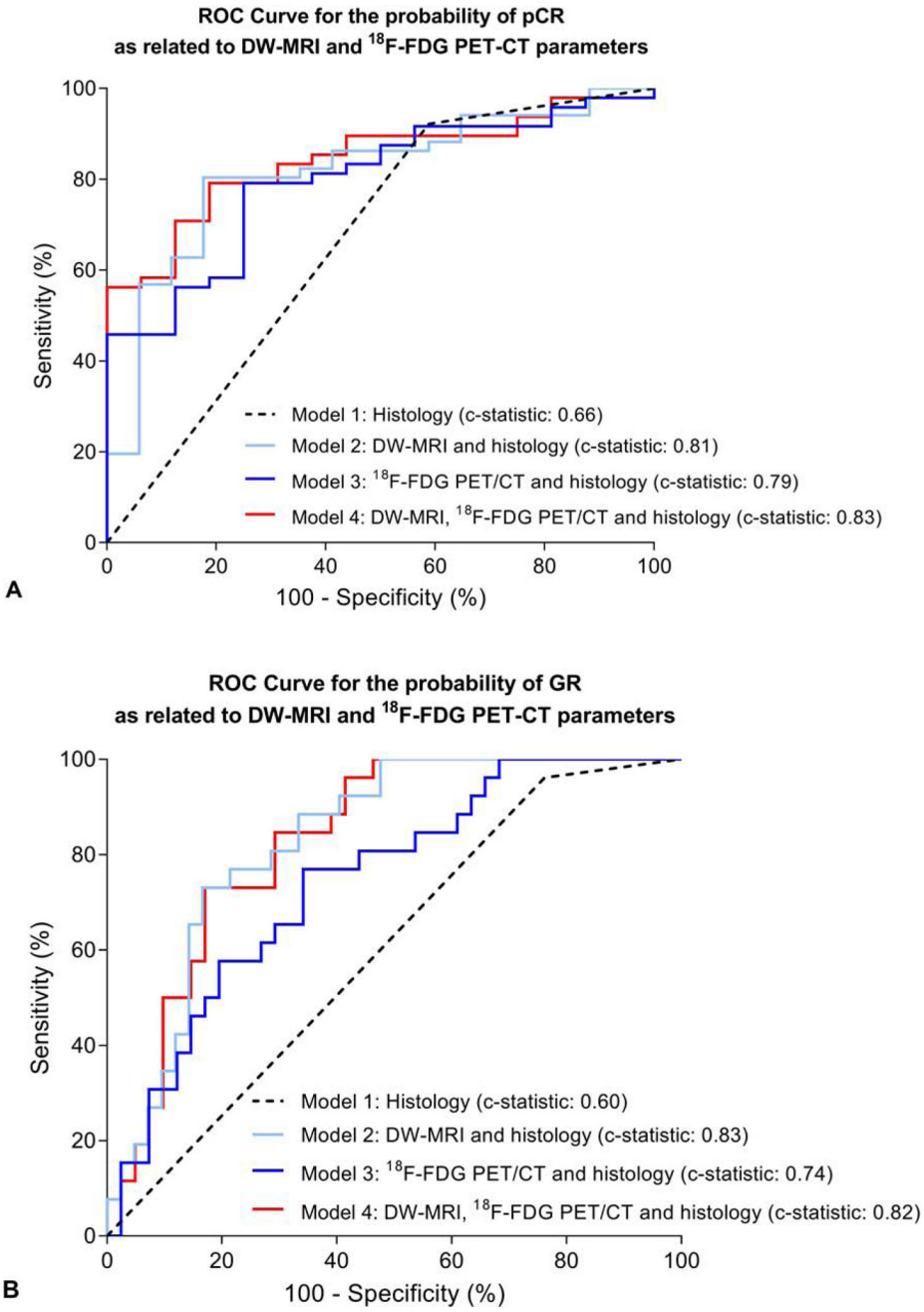

Figure 3.

Receiver-operating-characteristic curve analysis for the regression models with DW-MRI and ^{18}F-FDG PET/CT parameters, as well as histopathological tumor type, for discriminating between pathologic complete response (TRG 1)and non-pathologic complete response (TRG 2–4) patients (A) as well as between good responders (TRG 1–2, GR) and poor responders (TRG 3–4, non-GR) (B).

Table 1.

Clinical characteristics of the study population and univariable analyses between predetermined clinical predictors and response groups.

Characteristics	All	pCR (n=18)	non-pCR (n=51)	p-value	GR (n=43)	non-GR (n=26)	p-value
	n (%)	n (%)	n (%)		n (%)	n (%)	
Age at diagnosis, in years (mean ± SD)	61.0 ± 9.2	60.1 ± 8.8	61.4 ± 9.4	NA	61.1 ± 9.4	60.9 ± 9.0	NA
Sex				NA			NA
Male	61 (88)	3 (17)	5 (10)		5 (12)	3 (12)	
Female	8 (12)	15 (83)	46 (90)		38 (88)	23 (88)	
BMI, in kg/m^2 (mean ± SD)	25.7 ± 3.8	24.7 ± 5.2	26.1 ± 3.1	NA	25.2 ± 4.2	26.7 ± 2.8	NA
WHO performance status at diagnosis				NA			NA
0	27 (39)	5 (28)	22 (43)		18 (42)	9 (35)	
1	40 (58)	12 (67)	28 (55)		24 (56)	16 (62)	
2	2 (3)	1 (6)	1 (2)		1 (2)	1 (4)	
Tumor location				NA			NA
Proximal esophagus	1 (1)	0 (0)	1 (2)		1 (2)	0 (0)	
Middle esophagus	9 (13)	5 (28)	4 (8)		8 (19)	1 (4)	
Distal esophagus	40 (58)	6 (33)	34 (67)		22 (51)	18 (69)	
Gastroesophageal junction	19 (28)	7 (39)	12 (24)		12 (28)	7 (27)	
Clinical T status *				0.233			0.698
T2	9 (13)	2 (11)	7 (14)		6 (14)	3 (12)	
T3	59 (86)	15 (83)	44 (86)		36 (84)	23 (88)	
T4	1 (1)	1 (6)	0 (0)		1 (2)	0 (0)	
Clinical N status *				NA			NA
N0	23 (33)	6 (33)	17 (33)		15 (35)	8 (31)	
N1	26 (38)	5 (28)	21 (41)		12 (28)	14 (54)	
N2	18 (26)	7 (39)	11 (22)		15 (35)	3 (12)	
N3	2 (3)	0 (0)	2 (4)		1 (2)	1 (4)	
Histology				0.001			0.069
Adenocarcinoma	57 (83)	10 (56)	47 (92)		32 (74)	25 (96)	
Squamous cell carcinoma	11 (16)	7 (39)	4 (8)		10 (23)	1 (4)	
Undifferentiated large cell carcinoma	1 (1)	1 (6)	0 (0)		1 (2)	0 (0)	
Tumor regression grade				NA			NA
TRG 1 (pCR)	18 (26)	18 (100)	NA		18 (42)	NA	
TRG 2	25 (36)	NA	25 (49)		25 (58)	NA	
TRG 3	20 (29)	NA	20 (39)		NA	20 (77)	
TRG 4	6 (9)	NA	6 (12)		NA	6 (23)	
Neoadjuvant chemoradiotherapy regimen				0.902			0.357
Carboplatin/paclitaxel + 41.4Gy	43 (45)	11 (61)	32 (63)		25 (58)	18 (69)	

Characteristics	All	pCR (n=18)	non-pCR (n=51)	p-value	GR (n=43)	non-GR (n=26)	p-value
	n (%)	n (%)	n (%)		n (%)	n (%)	
5-fluorouracil-based + 50.4 Gy	26 (38)	7 (39)	19 (37)		18 (42)	8 (31)	
Time interval nCRT to surgery (median, IQR)	8 (7 – 10)	8 (7 – 9)	8 (7 – 10)	0.269	8 (7 – 10)	8 (6 – 9)	0.630

BMI body mass index at diagnosis; *IQR* interquartile range; *NA* not applicable; *pCR* pathologic complete response; *TRG* tumor regression grade; *WHO* World Health Organization

*Clinical T and N status are based on AJCC TNM 7th edition.

Note. P-values in this table were based on χ^2 or Fisher's exact test for categorical variables and the Mann-Whitney U test for continuous variables.

Table 2.

Relative changes in ^{18}F-FDG PET/CT and DW-MRI parameters between pathologic complete response (TRG 1) and non-pathologic complete response (TRG 2–4), as well as between good response (TRG 1–2) and poor response (TRG 3–4).

		Median (IQR)		p-value*	Benjamini-Hochberg adjusted p-value	c-statistic
		pCR (TRG 1) (n=18)	non-pCR (TRG 2–4) (n=51)			
During treatment	ΔSUV_{mean} (%)	−27 (−38, −6)	−19 (−33, −3)	*0.246*	*0.264*	0.59
	ΔSUV_{max} (%)	−29 (−46, −8)	−23 (−35, −2)	*0.408*	*0.408*	0.57
	ΔTLG_{during} (%)	−54 (−64, −23)	−20 (−50, 10)	***0.020***	*0.052*	0.69
	ΔADC_{during}(%)	28 (15, 39)	11 (4, 17)	***0.001***	***0.008***	0.77
Post treatment	ΔSUV_{mean} (%)	−63 (−68, −49)	−42 (−58, −16)	***0.003***	***0.016***	0.74
	ΔSUV_{max} (%)	−68 (−78, −58)	−54 (−71, −31)	***0.026***	*0.052*	0.68
	ΔTLG_{post} (%)	−86 (−93, −81)	−65 (−88, −32)	***0.006***	***0.024***	0.72
	ΔADC_{post} (%)	34 (13, 46)	20 (10, 38)	*0.139*	*0.201*	0.63
		Median (IQR)		**p-value***	**Benjamini-Hochberg adjusted p-value**	**c-statistic**
		GR (TRG 1–2) (n=43)	**non-GR (TRG 3–4) (n=26)**			
During treatment	ΔSUV_{mean} (%)	−26 (−42, −5)	−19 (−27, −2)	*0.054*	*0.086*	0.64
	ΔSUV_{max} (%)	−31 (−44, −14)	−12 (−29, 2)	***0.016***	*0.051*	0.67
	ΔTLG_{during} (%)	−37 (−62, −9)	−17 (−38, 20)	***0.024***	*0.052*	0.66
	ΔADC_{during} (%)	20 (13, 33)	7.3 (0, 11)	<***0.001***	***0.008***	0.84
Post treatment	ΔSUV_{mean} (%)	−51 (−66, −34)	−42 (−56, −11)	***0.042***	*0.075*	0.65
	ΔSUV_{max} (%)	−63 (−76, −44)	−53 (−74, −26)	*0.187*	*0.230*	0.60
	ΔTLG_{post} (%)	−81 (−89, −54)	−64 (−88, −35)	*0.248*	*0.264*	0.59
	ΔADC_{post} (%)	24 (13, 44)	21 (6, 30)	*0.151*	*0.201*	0.61

*p-value calculated based on Mann-Whitney U test

ADC apparent diffusion coefficient; *c-statistic* concordance statistic; *GR* good response; *IQR* interquartile range; *pCR* pathologic complete response; *SD* standard deviation; *SUV* standardized uptake value; *TLG* tumor lesion glycolysis; *TRG* tumor regression grade

Table 3.

Ridge regression analyses demonstrating the complementary value of ^{18}F-FDG PET/CT and DW-MRI parameters with pathologic complete response (TRG 1) and good response (TRG 1–2) as outcome variables.

pCR (TRG 1)			GR (TRG 1 +2)		
Predictors	**Odds ratio**	**Optimism-corrected c-statistic (95% bootstrapped CI)**	**Predictors**	**Odds ratio**	**Optimism corrected c-statistic (95% bootstrapped CI)**
		Model 1: Histology			
Squamous cell carcinoma†	8.23	**0.66 (0.53 – 0.78)***	Squamous cell carcinoma†	7.81	**0.60 (0.52 – 0.67)***
		Model 2: DW-MRI parameter and histology			
ΔADC_{during} (%)	1.04	**0.81 (0.70 – 0.93)***	ΔADC_{during} (%)	1.08	**0.83 (0.73 – 0.92)***
Squamous cell carcinoma†	4.56		Squamous cell carcinoma†	2.95	
		Model 3: ^{18}F-FDG PET/CT parameter and histology			
$\Delta SUV_{mean,post}$ (%)	0.98	**0.79 (0.67 – 0.90)§**	$\Delta SUV_{max,during}$ (%)	0.99	**0.74 (0.60 – 0.88)¶**
Squamous cell carcinoma†	3.90		Squamous cell carcinoma†	3.12	
		Model 4: DW-MRI and ^{18}F-FDG PET/CT parameters and histology			
ΔADC_{during} (%)	1.03	**0.83 (0.74 – 0.94)§**	ΔADC_{during} (%)	1.06	**0.82 (0.75 – 0.94)¶**
$ASUV_{mean,post}$ (%)	0.98		$\Delta SUV_{max,during}$ (%)	0.99	
Squamous cell carcinoma†	3.19		Squamous cell carcinoma†	3.14	

ADC apparent diffusion coefficient; *CI* confidence interval; *c-statistic* concordance statistic; *GR* good response (TRG 1–2); *pCR* pathologic complete response; *SUV* standardized uptake value.

*Analyses based on 68 patients of whom 17 had pCR (TRG 1) and 42 had GR (TRG 1–2). The applied exclusion criterion was histology other than adenocarcinoma or squamous cell carcinoma (n=1).

§Analyses based on 64 patients of whom 16 had pCR (TRG 1). Exclusion criteria were no post-treatment ^{18}F-FDG PET/CT available (n=3), no uptake on post-treatment ^{18}F-FDG PET/CT (n=1) and histology other than adenocarcinoma or squamous cell carcinoma (n=1).

¶Analyses based on 67 patients of whom 41 had GR (TRG 1–2). Exclusion criteria were no ^{18}F-FDG PET/CT available during treatment (n=1) and histology other than adenocarcinoma or squamous cell carcinoma (n=1).

†Adenocarcinoma was used as reference category.

Note. For the analyses of good response only ^{18}F-FDG PET/CT and DW-MRI parameters acquired during neoadjuvant chemoradiotherapy were included in the analyses considering the relevance of the research question. Regression coefficients and intercepts of all models are included in the supplementary file.

Table 4.

Multivariable Cox regression analyses for overall survival (OS) and disease free survival (DFS) based on the model with DW-MRI and ^{18}F-FDG PET/CT parameters and histology.

	Predictors	Hazard ratio (95% CI)
Overall survival *	ΔADC_{during} (%)	0.98 (0.95 – 1.01)
	$\Delta SUV_{mean,post}$ (%)	0.99 (0.97 – 1.01)
	Squamous cell carcinoma †	0.67 (0.21 – 2.16)
Disease free survival §	ΔADC_{during} (%)	0.98 (0.95 – 1.01)
	$\Delta SUV_{mean,post}$ (%)	1.00 (0.98 – 1.01)
	Squamous cell carcinoma †	0.65 (0.17 – 2.44)

*Analyses based on 63 patients of whom 27 deceased during follow-up. The applied exclusion criteria were no post-treatment ^{18}F-FDG PET/CT available (n=3), no uptake on post-treatment ^{18}F-FDG PET/CT (n=1), histology other than adenocarcinoma or squamous cell carcinoma (n=1) and incomplete follow-up data (n=1).

§Analyses based on 57 patients of whom 24 had disease recurrence during follow-up. The applied exclusion criteria were no post-treatment ^{18}F-FDG PET/CT available (n=3), no uptake on post-treatment ^{18}F-FDG PET/CT (n=1), histology other than adenocarcinoma or squamous cell carcinoma (n=1), incomplete follow-up data (n=5) and censored before earliest event (n=1).

†Adenocarcinoma was used as reference category.

ADC apparent diffusion coefficient; *CI* confidence interval; *SUV* standardized uptake value; *TLG* tumor lesion glycolysis; *TRG* tumor regression grade