

---

---

# Predicting Response to Neoadjuvant Chemoradiotherapy in Esophageal Cancer with Textural Features Derived from Pretreatment <sup>18</sup>F-FDG PET/CT Imaging

Roelof J. Beuinga<sup>1-3</sup>, Jan B. Hulshoff<sup>1</sup>, Lisanne V. van Dijk<sup>4</sup>, Christina T. Muijs<sup>4</sup>, Johannes G.M. Burgerhof<sup>5</sup>, Gursah Kats-Ugurlu<sup>6</sup>, Riemer H.J.A. Slart<sup>2,3</sup>, Cornelis H. Slump<sup>7</sup>, Véronique E.M. Mul<sup>4</sup>, and John Th.M. Plukker<sup>1</sup>

<sup>1</sup>Department of Surgery (Division of Surgical Oncology), University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; <sup>2</sup>Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; <sup>3</sup>Department of Biomedical Photonic Imaging, University of Twente, Enschede, The Netherlands; <sup>4</sup>Department of Radiation Oncology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; <sup>5</sup>Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; <sup>6</sup>Department of Pathology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; and <sup>7</sup>MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, The Netherlands

---

Adequate prediction of tumor response to neoadjuvant chemoradiotherapy (nCRT) in esophageal cancer (EC) patients is important in a more personalized treatment. The current best clinical method to predict pathologic complete response is SUV<sub>max</sub> in <sup>18</sup>F-FDG PET/CT imaging. To improve the prediction of response, we constructed a model to predict complete response to nCRT in EC based on pretreatment clinical parameters and <sup>18</sup>F-FDG PET/CT-derived textural features. **Methods:** From a prospectively maintained single-institution database, we reviewed 97 consecutive patients with locally advanced EC and a pretreatment <sup>18</sup>F-FDG PET/CT scan between 2009 and 2015. All patients were treated with nCRT (carboplatin/paclitaxel/41.4 Gy) followed by esophagectomy. We analyzed clinical, geometric, and pretreatment textural features extracted from both <sup>18</sup>F-FDG PET and CT. The current most accurate prediction model with SUV<sub>max</sub> as a predictor variable was compared with 6 different response prediction models constructed using least absolute shrinkage and selection operator regularized logistic regression. Internal validation was performed to estimate the model's performances. Pathologic response was defined as complete versus incomplete response (Mandard tumor regression grade system 1 vs. 2–5). **Results:** Pathologic examination revealed 19 (19.6%) complete and 78 (80.4%) incomplete responders. Least absolute shrinkage and selection operator regularization selected the clinical parameters: histologic type and clinical T stage, the <sup>18</sup>F-FDG PET-derived textural feature long run low gray level emphasis, and the CT-derived textural feature run percentage. Introducing these variables to a logistic regression analysis showed areas under the receiver-operating-characteristic curve (AUCs) of 0.78 compared with 0.58 in the SUV<sub>max</sub> model. The discrimination slopes were 0.17 compared with 0.01, respectively. After internal validation, the AUCs decreased to 0.74 and 0.54, respectively. **Conclusion:** The predictive values of the constructed models were superior to the standard method (SUV<sub>max</sub>). These results can be considered as an initial step in predicting tumor re-

sponse to nCRT in locally advanced EC. Further research in refining the predictive value of these models is needed to justify omission of surgery.

**Key Words:** textural analysis; response prediction; esophageal cancer; <sup>18</sup>F-FDG PET/CT

**J Nucl Med 2017; 58:723–729**

DOI: 10.2967/jnumed.116.180299

---

**E**sophageal cancer (EC) is one of the most aggressive tumors, with early recurrences even after radical surgery. The standard treatment in locally advanced (T1/N1–3/M0 and T2–4a/N0–3/M0) resectable EC is neoadjuvant chemoradiotherapy (nCRT) followed by a radical esophagectomy. In the Dutch CROSS (Chemo-Radiotherapy for Oesophageal cancer followed by Surgery Study) trial, nCRT improved the 5-y overall survival rate from 34% to 47% (1). Not all patients benefit from nCRT; 29% of the patients in the CROSS trial had a complete response, 52% had a partial response, and even 18% had no tumor response (1). For complete responders, surgical intervention might not be beneficial and a wait-and-see policy might suffice. Hence, adequate response prediction is important in developing personalized treatment in EC. Moreover, accurate response prediction may be relevant in patient counseling in future clinical trial strategies based on personalized treatment. So far, response prediction showed only promising results with functional imaging of tumor viability with <sup>18</sup>F-FDG PET and recently with diffusion-weighted MRI (2–4). Traditional image-derived indices used in PET rely on quantification of lesion SUVs and overall tumor volume, which have been shown to be important factors for patient outcome and treatment response (5,6). Although useful, these parameters omit available information related to the spatial distribution and specific features regarding intratumor radiotracer accumulation. This may limit the possibility to further characterize the biologic behavior of the tumor, based on hypoxia-induced heterogeneity and genomic instability. Intratumoral heterogeneity is correlated with aggressive tumor behavior and a decreased response due to expression of specific receptors

---

Received Jul. 11, 2016; revision accepted Sep. 19, 2016.

For correspondence or reprints contact: John Th.M. Plukker, Department of Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

E-mail: j.t.m.plukker@umcg.nl

Published online Oct. 13, 2016.

COPYRIGHT © 2017 by the Society of Nuclear Medicine and Molecular Imaging.

with high cellular proliferation and angiogenesis (7–9). Hence, even small tumor biopsies lack complete molecular characterization due to spatial heterogeneity. A novel approach is to quantify spatial heterogeneity of metabolism and tissue density characterized by  $^{18}\text{F}$ -FDG uptake and Hounsfield units with textural features. The concept of textural analysis is based on the spatial arrangement of voxels in a predefined volume of interest (VOI). This spatial intratumoral heterogeneity can be depicted from different spatial interrelationships on  $^{18}\text{F}$ -FDG PET/CT scans. Therefore,  $^{18}\text{F}$ -FDG PET/CT textural features have been proposed to be valuable in response prediction (10–15). The aim of this study was to develop a model to predict complete response to nCRT in locally advanced EC based on pretreatment clinical predictors and  $^{18}\text{F}$ -FDG PET/CT-derived textural features.

## MATERIALS AND METHODS

### Patients

In this retrospective study, potentially curatively resectable EC patients were consecutively selected who underwent nCRT followed by esophagectomy between December 2009 and March 2016. Patients with fewer than 4 courses of chemotherapy, with missing  $^{18}\text{F}$ -FDG PET/CT or with incomplete medical records, were excluded, yielding a total of 97 patients. In line with the rules of the Dutch National Health Sciences, our Institutional Review Board approved this retrospective study, and the requirement to obtain informed consent was waived.

Data were obtained from a prospectively maintained single-institution database including patient characteristics, tumor- and treatment-related data, and follow-up data. All patients were clinically staged with esophagoscopy and biopsy, endoscopic ultrasonography with fine-

needle aspiration if indicated, and whole-body integrated  $^{18}\text{F}$ -FDG PET/CT. Patients were staged according to the seventh edition of the TNM system maintained by the American Joint Committee on Cancer (16) and discussed in the hospital's multidisciplinary esophageal tumor board.

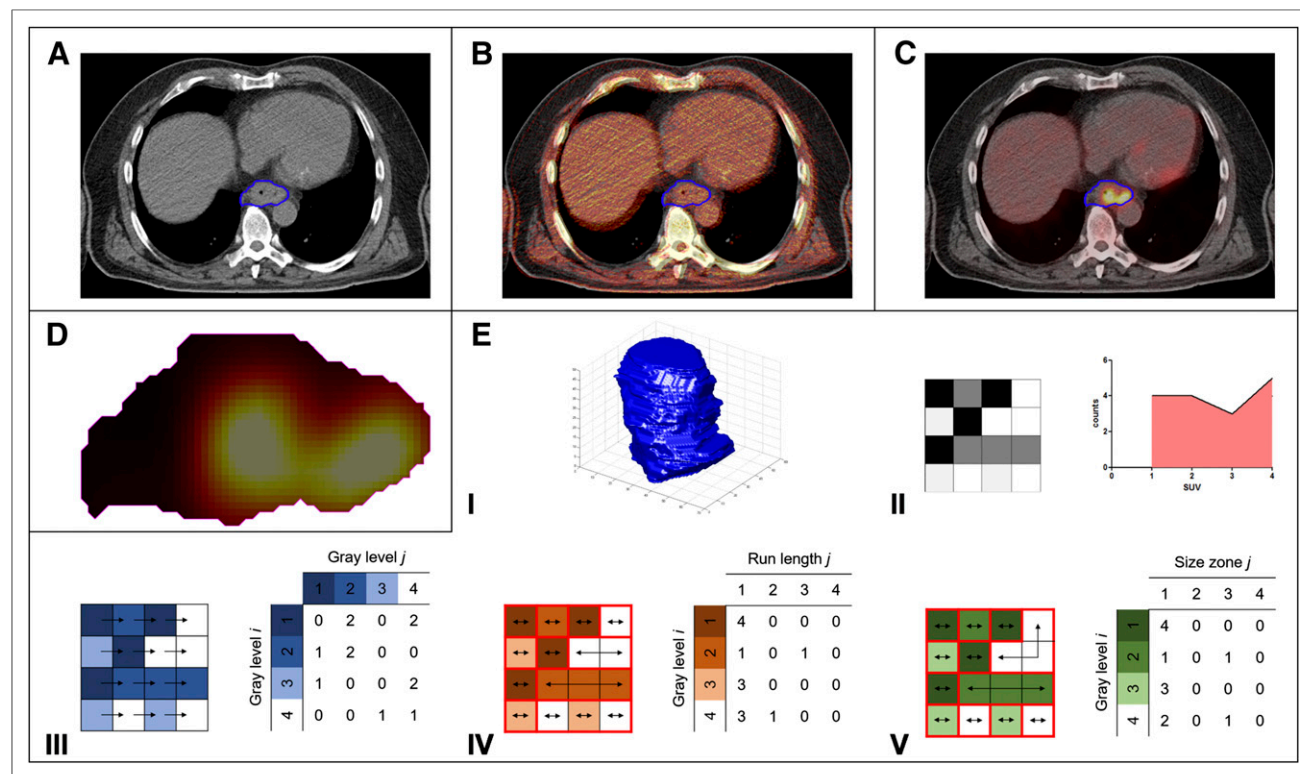
### Imaging

PET/CT imaging was performed with an integrated  $^{18}\text{F}$ -FDG PET/CT system (Biograph mCT 4–64 PET/CT; Siemens). Patients fasted for at least 6 h before PET/CT, with no restrictions on drinking water. Serum glucose levels were measured before  $^{18}\text{F}$ -FDG administration with a weight-based dose of 3 MBq/kg. Sixty minutes after tracer injection, patients were scanned in treatment position. An inspiration breath-hold low-dose CT for attenuation correction was performed, and PET acquisitions were obtained in the caudal–cranial direction with a field of view of  $500 \times 500 \times 500$  mm, 3-dimensional setting, 2–3 min per bed position, matrices of  $512 \times 512$  (0.98  $\times$  0.98 mm pixel size), and 2-mm slice thickness. Image data were reconstructed according to guidelines of the European Association of Nuclear Medicine (17).

Radiotherapy treatment planning including target volume delineation and CT texture analysis was performed on a 16- or 64-multidetector row spiral CT machine (Somatom Sensation 16 or 64; Siemens Medical Systems). A CT thorax/abdomen scan was obtained in the cranial–caudal direction with matrices of  $512 \times 512$  (0.98  $\times$  0.98 mm pixel size) and a 3-mm slice thickness.

### Treatment and Pathology

On the basis of the experiences and the good results of the CROSS study, in which our institute had participated, our multidisciplinary tumor board decided to continue nCRT according to the CROSS schedule. This treatment consisted of weekly intravenously administered paclitaxel



**FIGURE 1.** (A) Manual delineation of gross tumor volume on planning CT. (B) Coregistration of low-dose CT (lava colormap) to radiotherapy planning CT (gray scale colormap). (C) Overlay of PET image (lava colormap) onto radiotherapy planning CT (gray scale colormap). (D) Cropping of PET VOI. (E) Feature extraction: (I) assessment of tumor shape by means of geometry features, (II) global assessment of tonal distribution by means of first-order textural features, (III) assessment of pairwise arrangement of voxels by means of gray level co-occurrence matrix, (IV) assessment of alignment of voxels with the same intensity by means of gray level run-length matrix, (V) assessment of characteristics of homogeneous zones by means of gray level size-zone matrix.

**TABLE 1**  
Patient Characteristics

| Characteristic                           | n  | %    |
|--|----|------|
| <b>Sex</b>                               |    |      |
| Male                                     | 82 | 84.5 |
| Female                                   | 15 | 15.5 |
| <b>Age</b>                               |    |      |
| <70 y                                    | 78 | 80.4 |
| ≥70 y                                    | 19 | 19.6 |
| <b>Histology</b>                         |    |      |
| AC                                       | 88 | 90.7 |
| SCC                                      | 9  | 9.3  |
| <b>Tumor grade</b>                       |    |      |
| Missing                                  | 6  | 6.2  |
| G1                                       | 49 | 50.5 |
| G3                                       | 42 | 43.3 |
| <b>EUS tumor length</b>                  |    |      |
| <5 cm                                    | 37 | 40.2 |
| ≥5 cm                                    | 58 | 59.8 |
| <b>Localization</b>                      |    |      |
| Mid                                      | 4  | 4.1  |
| Distal                                   | 62 | 63.9 |
| GEJ                                      | 31 | 32.0 |
| <b>nCRT cycles</b>                       |    |      |
| 4  | 18 | 18.6 |
| 5  | 79 | 81.4 |
| <b>Circumferential resection margin*</b> |    |      |
| R0                                       | 90 | 92.8 |
| R1                                       | 7  | 7.2  |
| <b>Proximal resection margin*</b>        |    |      |
| R0                                       | 96 | 99.0 |
| R1                                       | 1  | 1.0  |
| <b>Mandard TRG</b>                       |    |      |
| 1  | 19 | 19.6 |
| 2  | 23 | 23.7 |
| 3  | 37 | 38.1 |
| 4  | 15 | 15.5 |
| 5  | 3  | 3.1  |

\*Tumor-free (R0) resection margin defined according to criteria of College of American Pathologists as > 0 mm.

AC = adenocarcinoma; SCC = squamous cell carcinoma; EUS = endoscopic ultrasonography; GEJ = gastroesophageal junction; TRG = tumor regression grade.

(50 mg/m<sup>2</sup>) and carboplatin (areas under the receiver-operating-characteristic curve [AUC], 2 mg·min·mL<sup>-1</sup>) during 5 wk with concurrent external radiotherapy (41.4 Gy in 23 fractions, 5 d per wk) (1). Transthoracic esophagectomy with 2-field lymphadenectomy was performed within 6–8 wk after completion of nCRT. The resected specimens were examined according to a standard protocol (18). Resection margins were defined according to the definitions of the College of American Pathologists as microscopic tumor-free (R0: > 0 mm) or

tumor-positive (R1). Pathologic response was assessed by 2 expert gastrointestinal pathologists according to the Mandard tumor regression grade (TRG) (19), ranging from complete response (TRG 1) without viable tumor cells left, to partial response (TRG 2–4) with viable tumor cells left, to no response at all (TRG 5).

#### VOI

Textural analysis was performed on a VOI incorporating the gross tumor volume for radiation treatment planning. Tumor delineation was performed manually with consensus between 3 experienced radiation oncologists on axial planes of the radiotherapy planning CT, to enclose 3-dimensional coverage of the entire tumor. Involved lymph nodes were not included into the VOI, because these lesions are too small (<10 cm<sup>3</sup>) for reliable textural analysis (20). The gross tumor volume was rigidly registered to the <sup>18</sup>F-FDG PET/CT data series (RTx Workstation 1.0; Mirada Medical). Erroneous registrations were manually adjusted after consensus of the collaborating investigators.

#### Tonal Discretization

<sup>18</sup>F-FDG PET/CT imaging data and VOI delineations were loaded into Matlab 2014b (MathWorks; an interactive image processing environment) for processing and analyses. The SUV, for semiquantitative analysis of metabolism, was corrected for individual variations in serum glucose level and was discretized to reduce the continuous scale to a finite set of values and to reduce noise throughout the entire study in increments of 0.5 g/mL according to Doane's optimal bin width (21). Similarly, the Hounsfield unit scale for quantitative analysis of tumor density was discretized in increments of 30 Hounsfield units for textural analysis.

**TABLE 2**  
Staging According to Seventh Edition of TNM System Maintained by American Joint Committee on Cancer (16)

| Characteristic   | n  | %    |
|------------------|----|------|
| <b>cT stage</b>  |    |      |
| cT1              | 2  | 2.1  |
| cT2              | 16 | 16.5 |
| cT3              | 74 | 76.3 |
| cT4a             | 5  | 5.2  |
| <b>cN stage</b>  |    |      |
| cN0              | 21 | 21.7 |
| cN1              | 45 | 46.4 |
| cN2              | 28 | 28.9 |
| cN3              | 3  | 3.1  |
| <b>ypT stage</b> |    |      |
| ypT0             | 19 | 19.6 |
| ypT1             | 13 | 13.4 |
| ypT2             | 13 | 13.4 |
| ypT3             | 52 | 53.6 |
| <b>ypN stage</b> |    |      |
| ypN0             | 63 | 65.0 |
| ypN1             | 18 | 18.6 |
| ypN2             | 10 | 10.3 |
| ypN3             | 6  | 6.2  |

cT/N = clinical tumor/nodal stage; ypT/N = pathologic tumor/nodal stage after nCRT.

**TABLE 3**

Estimated Regression Coefficients of Prediction Models for Pathologic Complete Response Without Optimism Correction

| Variable           | Model 1     |      |      | Model 2–3   |      |      | Model 4     |      |      | Model 5     |      |      | Model 6     |      |      |
|--------------------|-------------|------|------|-------------|------|------|-------------|------|------|-------------|------|------|-------------|------|------|
|                    | Coefficient | SE   | P    | Coefficient | SE   | P    | Coefficient | SE   | P    | Coefficient | SE   | P    | Coefficient | SE   | P    |
| Intercept          | -0.88       | 0.50 | 0.08 | -0.42       | 0.50 | 0.39 | -0.86       | 0.65 | 0.19 | -1.15       | 0.72 | 0.11 | -1.83       | 0.91 | 0.04 |
| SUV <sub>max</sub> | -1.72       | 1.45 | 0.24 |             |      |      |             |      |      |             |      |      |             |      |      |
| Histology          |             |      |      |             |      |      |             |      |      |             |      |      |             |      |      |
| AC                 |             |      |      | 1.00        |      |      | 1.00        |      |      | 1.00        |      |      | 1.00        |      |      |
| SCC                |             |      |      | -1.70       | 0.61 | 0.01 | -1.47       | 0.65 | 0.02 | -1.89       | 0.64 | 0.00 | -1.63       | 0.67 | 0.02 |
| cT stage           |             |      |      |             |      |      |             |      |      |             |      |      |             |      |      |
| cT1 & cT2          |             |      |      | 1.00        |      |      | 1.00        |      |      | 1.00        |      |      | 1.00        |      |      |
| cT3 & cT4a         |             |      |      | 2.03        | 0.78 | 0.01 | 1.98        | 0.80 | 0.01 | 2.27        | 0.81 | 0.00 | 2.23        | 0.83 | 0.01 |
| LRLGLe-PET         |             |      |      |             |      |      | 0.56        | 0.55 | 0.31 |             |      |      | 0.71        | 0.59 | 0.23 |
| RP-CT              |             |      |      |             |      |      |             |      |      | 0.01        | 0.01 | 0.15 | 0.02        | 0.01 | 0.10 |

**Candidate Predictors**

For each patient, a total of 88 parameters were evaluated, including 7 clinical parameters; 16 geometry features; the glycolytic volume based on tumor volume and SUV<sub>mean</sub>; and 19 first-order, 24 second-order, and 22 higher order textural features extracted from <sup>18</sup>F-FDG PET and CT (supplemental materials [available at <http://jnm.snmjournals.org>]). First-order textural features are statistics based on the gray level distribution of the image but do not consider relative positions of gray levels. Second- and higher order textural features do consider relative positions of gray levels and therefore allow quantification of heterogeneity. For various spatial interrelationships, frequency distributions (Fig. 1) were obtained—that is, the gray level co-occurrence (spatial dependence) matrix for pairwise arrangement of voxels (extracted with a pixel-to-pixel distance equal to 1) (22), the gray level run-length matrix for alignment of voxels with the same intensity (23), and the gray level size-zone matrix for characteristics of homogeneous zones (24). Directional voxel analysis was performed in 3 dimensions with a connectivity of 26 voxels and analysis in 13 angular directions. All second- and higher order textural features are weighted averages of these matrices to express the relative importance of their properties. All extracted textural features were normalized to the range [0,1].

**Statistical Analysis**

Statistical analysis was performed with R 3.2.2 open-source software using the *glmnet* package (version 2.0–2) and the *rms* package (version 4.4–0), available from the Comprehensive R Archive Network (<http://www.r-project.org>).

Because textural feature values may be subject to interobserver variability in the delineation of the tumor, the original delineations were uniformly eroded by ball-shaped structuring elements with radii of 1 and 2 voxels. For each delineation, textural features were extracted and the stability of each feature was evaluated with the intraclass correlation. Only stable features (intraclass correlation > 0.7) were considered for further analysis. Predictors were then selected by a univariable logistic model with a response variable labeling complete (Mandard TRG 1) and incomplete response (Mandard TRG 2–5). All potential predictors that met the Akaike information criterion (AIC) were considered significant. To discourage overfitting, the AIC is based on rewarding goodness of fit and penalizing the complexity of the model. The AIC requires  $\chi^2 > 2 \cdot df$ —that is, when considering a predictor with one degree of freedom df, this implies an  $\alpha = P(\chi^2 \geq 2) = 0.157$  (25).

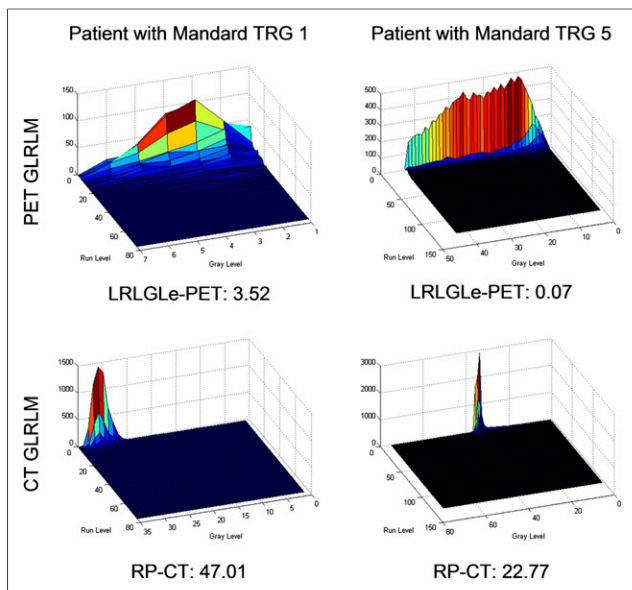
Significant predictors were used to construct 6 multivariable logistic regression models for comparison with current most accurate prediction model with SUV<sub>max</sub> as predictor variable (model 1). These models were constructed by introducing clinical parameters (model 2); clinical parameters and geometry features (model 3); clinical parameters, geometry features, and PET textural features (model 4); clinical parameters, geometry features, and CT textural features (model 5); and clinical parameters, geometry features, and PET/CT textural features (model 6) to a least absolute shrinkage and selection operator, a technique for L1-norm regularization. By increasing the shrinkage parameter  $\lambda$ , the regularization shrinks the estimated coefficients and excludes variables when they become zero. The  $\lambda$ -value that minimized

**TABLE 4**

Estimates of Model Performance for the Prediction Models

| Model | Goodness of fit |       |                | Discrimination |      | Calibration |       |      | Validation          |          |
|-------|-----------------|-------|----------------|----------------|------|-------------|-------|------|---------------------|----------|
|       | -2LLH           | AIC   | R <sup>2</sup> | AUC            | DS   | Intercept   | Slope | HLp  | R <sup>2</sup> boot | AUC boot |
| 1     | 94.46           | 98.46 | 0.02           | 0.58           | 0.01 | 2.09        | 2.46  | 0.75 | 0.00                | 0.54     |
| 2–3   | 81.80           | 87.80 | 0.22           | 0.71           | 0.14 | -0.04       | 0.94  | 1.00 | 0.17                | 0.70     |
| 4     | 80.59           | 88.59 | 0.23           | 0.71           | 0.15 | -0.15       | 0.87  | 0.45 | 0.17                | 0.69     |
| 5     | 79.82           | 87.82 | 0.24           | 0.79           | 0.16 | -0.14       | 0.86  | 0.42 | 0.18                | 0.76     |
| 6     | 78.03           | 88.03 | 0.27           | 0.78           | 0.17 | -0.22       | 0.81  | 0.46 | 0.18                | 0.74     |

-2LLH = -2 loglikelihood; R<sup>2</sup> = Nagelkerke R<sup>2</sup>; DS = discrimination slope; HLp = Hosmer–Lemeshow P value; boot = internal validated with bootstrapping.



**FIGURE 2.** Example of values of selected textural features and their corresponding frequency distributions for complete and nonresponder. GLRLM = gray level run-length matrix.

the 10-fold cross-validated mean squared error was repeatedly determined with 100 repetitions. The optimal  $\lambda$ -value was robustly determined by averaging over these obtained  $\lambda$ -values. The selected variables were fitted to the data with a logistic regression.

The model's calibration was evaluated using visual inspection of calibration plots and the Hosmer–Lemeshow test. The model's performance was quantified in terms of discrimination with the AUC and the discrimination slope. The goodness of fit was evaluated with the  $-2$  loglikelihood and the Nagelkerke  $R^2$ . The model was internally validated by a bootstrap approach with 2,000 repetitions. Bootstrapping allowed for obtaining the optimism-corrected measures for model performance and for shrinkage of the estimated regression coefficients using the optimism-corrected slope.

## RESULTS

### Patients and Treatment

Patients characteristics are shown in Tables 1 and 2. Seventy-nine patients (81.4%) received the complete nCRT regimen (all patients received the full radiotherapy dose). Resection with curative intent was performed within a mean time of 56 (SD, 14) d after completion of nCRT. R0 resection was achieved in 90 (92.8%) patients and R1 resection in 7 (7.2%) patients, all with positive circumferential resection margins and one with a positive proximal resection margin. Pathologic findings revealed complete response in 19 patients (19.6%) and incomplete response in 78 patients (80.4%).

### Model Development

For the preselection, 144 of the 147 (97.3%) variables were found to be robust for contour variations. These variables were introduced to univariable logistic regression analysis, resulting in 24 significant variables predictive for response, including 4 clinical parameters; 0 geometry features; 1 first-order, 8 second-order, and 5 higher order PET textural features; and 1 first-order, 1 second-order, and 4 higher order CT textural features. All constructed prediction models performed significantly better than model 1 (based on  $SUV_{max}$ ). The introduction of only significant clinical parameters to the least absolute shrinkage and selection operator regularization

process resulted in the selection of histologic type and clinical T-stage (model 2). These variables were selected in each subsequently constructed model. Compared with model 1, the AUC improved from 0.58 to 0.71, the discrimination slope improved from 0.01 to 0.14, and the AIC decreased ( $\Delta AIC = 10.66$ ). For model 3, no additional variables were selected compared with model 2, because no geometry features were significant at the univariable logistic regression analysis. For model 4, the PET textural feature long run low gray level emphasis (LRLGLE-PET) was selected. Adding this variable did slightly improve the discrimination and the likelihood compared with models 2 and 3, but resulted in a higher AIC ( $\Delta AIC = -0.79$ ). After internal validation, the AUC was equal to 0.69. For model 5, the CT textural feature run percentage (RP-CT) was selected. Although the AIC was almost equal compared with models 2 and 3 ( $\Delta AIC = -0.02$ ), adding this variable improved the discrimination slope to 0.16 and the AUC remarkably to 0.79. This also persisted after internal validation (AUC = 0.76). Finally, entering all variables to the modeling process resulted in the selection of all above-mentioned variables (model 6). Model 6 had the best goodness of fit, but not the lowest AIC ( $\Delta AIC = -0.23, 0.56, \text{ and } -0.21$  compared with models 2–3, 4, and 5, respectively). The AUC slightly decreased to 0.78, whereas the discrimination slope was increased to 0.17. After internal validation, the AUC decreased to 0.74. The model regression coefficients and the corresponding model performance measures are shown in Tables 3 and 4, respectively. Figure 2 gives the values of the selected textural features and their corresponding frequency distributions for a complete and a nonresponder. For the selected textural features, the range of values to reproduce the normalization process and the found intraclass correlations for quantifying contouring robustness are given in the supplemental materials.

## DISCUSSION

An adequate method to predict pathologic complete response after nCRT has not yet been defined in EC patients. In personalized treatment, accurate response prediction will lead to a paradigm shift with omission of surgical treatment in complete responders or prevention of unnecessary nCRT in nonresponders. Response evaluation of nCRT is commonly based on tumor metabolic response measured by  $SUV_{max}$  with  $^{18}F$ -FDG PET, but with a low sensitivity and specificity of 67% and 68%, respectively (26). The current study is, to our knowledge, the first in predicting complete response with  $^{18}F$ -FDG PET/CT-derived textural features in a homogeneous group of EC patients treated according to the CROSS regimen. We demonstrated that all constructed prediction models showed significant improvement compared with predictions based on  $SUV_{max}$  alone and may therefore be considered as an initial step in predicting response.

In this study, the most predictive textural features were LRLGLE-PET and RP-CT. LRLGLE-PET depends on long runs (coarse texture) with low gray levels and was higher (i.e., low and homogeneous  $^{18}F$ -FDG uptake) for complete responders and lower (i.e., high and heterogeneous  $^{18}F$ -FDG uptake) for incomplete responders, possibly due to tumor hypoxia and necrosis. RP-CT measures the homogeneity of runs (fine texture) and was higher in complete responders. In univariable logistic regression, high LRLGLE-PET and RP-CT values were associated with squamous cell carcinoma ( $P = 0.12$  and  $0.13$ , respectively), confirming the higher complete response rates in squamous cell carcinoma (1).

The clinical value of  $SUV_{max}$  was limited, possibly because it is extracted from a single voxel and does not characterize the total

**TABLE 5**  
Current Literature Describing <sup>18</sup>F-FDG PET Texture Analysis in Response Prediction in EC

| Study                  | n   | Type    | nCRT   | Timing PET/CT      | Outcome                | Reported entered variables  |
|------------------------|-----|---------|--|--------------------|------------------------|---|
| Tixier et al. (13)     | 41  | AC, SCC | 60 Gy + cisplatin or carboplatin/ fluorouracil                               | Pre-nCRT           | CR vs. PR vs. NR (32)  | Pre-local homogeneity, pre-local entropy, pre-coarseness, pre-intensity variability, pre-size-zone features   |
| Hatt et al. (11)       | 50  | AC, SCC | 60 Gy + cisplatin/ fluorouracil  | Pre-nCRT           | CR + PR vs. NR (32)    | Pre-MATV, pre-entropy, pre-homogeneity, pre-dissimilarity, pre-intensity variability, pre-zone percentage   |
| Tan et al. (12)        | 20  | AC, SCC | 50.4 Gy + cisplatin/ fluorouracil  | Pre- and post-nCRT | TRG 1 + 2 vs. 3–5 (19) | $\Delta$ SUV <sub>max</sub> , SUV <sub>max</sub> ratio, $\Delta$ SUV <sub>mean</sub> , pre-skewness, post-inertia, post-correlation, post-cluster prominence  |
| Zhang et al. (15)      | 20  | AC, SCC | 50.4 Gy + cisplatin or carboplatin   | Pre- and post-nCRT | TRG 1 + 2 vs. 3–5 (19) | Post-orientation, tumor involves GEJ, $\Delta$ inertia, post-energy, post-entropy, $\Delta$ skewness  |
| Van Rossum et al. (14) | 217 | AC      | 45 or 50.4 Gy + fluoropyrimidine with either a platinum compound or a taxane | Pre- and post-nCRT | TRG 1 vs. 2–4 (33)     | EUS tumor length, cT stage, induction chemotherapy, post-nCRT endoscopic biopsy, subjective PET assessment, post-nCRT TLG, pre-cluster shade, $\Delta$ run percentage, $\Delta$ GLCM entropy, post-nCRT roundness |
| Current study          | 97  | AC, SCC | 41.4 Gy + carboplatin/ paclitaxel  | Pre-nCRT           | TRG 1 vs. 2–5 (19)     | Histology, cT stage, pre-LRLGLe-PET, pre-RP-CT  |

AC = adenocarcinoma; SCC = squamous cell carcinoma; CR = complete response; PR = partial response; NR = no response; MATV = metabolically active tumor volume; GEJ = gastroesophageal junction; EUS = endoscopic ultrasonography; cT = clinical T stage; TLG = total lesion glycolysis; GLCM = gray level co-occurrence matrix.

<sup>18</sup>F-FDG uptake. This causes a high dependency on the quality of the PET images (including noise) and the voxel size, which induces a low reproducibility.

Several studies focused on response prediction in EC using <sup>18</sup>F-FDG PET/CT-derived textural features (Table 5). Van Rossum et al. concluded that <sup>18</sup>F-FDG PET-derived textural features provide statistical value (14), but this does not translate into a clinically relevant benefit, which is in line with our findings. Van Rossum et al. performed only <sup>18</sup>F-FDG PET textural analysis, whereas this study demonstrates the additional value of CT textural analysis. Other studies demonstrated promising findings but are hampered by several limitations, including small patient cohorts with heterogeneous treatment schedules, lack of multivariable analyses, and a substantial chance of model overfitting due to the lack of optimism correction (11–13,15).

A limitation of this study is the absence of external validation, which is essential for implementation into clinical practice. Moreover, the following factors that affect textural analysis should be considered. Changing the bin width influences the quantization noise and has a crucial effect on textural features (27). Although only an indication, we used Doane's optimal bin width to discretize the SUVs and Hounsfield units (21). Moreover, respiratory gated PET/CT acquisitions could be considered to reduce respiration-induced smearing and contrast degradation (28).

The constructed prediction model may serve as a basic model, which can be extended with new features for usage for other applications. The current constructed model might be helpful toward a safe decision in postponing a burdensome surgical procedure in patients with a doubtful adequate physical condition after nCRT. In patients treated with definitive chemoradiotherapy, detection of nonresponders might allow additional treatments when available, whereas in complete responders an adjusted follow-up might be justified, to identify candidates for salvage surgery.

Up to now, the authors do not consider the predictive value of the constructed model high enough to justify the omission of surgery after nCRT in EC. A potential approach to improve the constructed basic prediction models could be (a) adding interim- or posttreatment textural analysis. Studies investigating both pre- and posttreatment textural analysis mainly reported posttreatment textural features to be associated with response (12,14,15). We performed a posttreatment textural analysis in patients with a posttreatment PET/CT scan (n = 20) and found 21 significant textural features for response in univariable regression analysis. However, posttreatment textural analysis suffers from radiation-induced esophagitis, which complicates delineation of the primary tumor and creates difficulties with tumor delineation in complete responders. (b) Texture could be characterized with more specific PET tracers such as <sup>18</sup>F-fluoroerythronitroimidazole (quantifying hypoxia) (29) or <sup>18</sup>F-fluorothymidine (targeting

cellular proliferation) (30) or by other functional imaging modalities including the apparent diffusion coefficient in diffusion-weighted MRI (4). (c) Moreover, biologic markers have shown to be potential molecular markers in individualizing EC treatment and may be incorporated to improve prediction models (31).

## CONCLUSION

The constructed models are a valuable initial step in predicting response to nCRT in locally advanced EC. Adding the  $^{18}\text{F}$ -FDG PET-derived textural feature LRLGLE-PET and RP-CT to a model with the clinical parameter histologic type and clinical T-stage is potentially predictive and was more accurate than response prediction based on  $\text{SUV}_{\text{max}}$ . These models may serve as basic models in determining clinical complete responders and can be extended with new features for usage for other applications.

## DISCLOSURE

No potential conflict of interest relevant to this article was reported.

## REFERENCES

- van Hagen P, Hulshof MC, van Lanschot JJ, et al. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med*. 2012;366:2074–2084.
- Chen YM, Pan XF, Tong LJ, Shi YP, Chen T. Can  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography predict responses to neoadjuvant therapy in oesophageal cancer patients? A meta-analysis. *Nucl Med Commun*. 2011;32:1005–1010.
- Ott K, Weber WA, Lordick F, et al. Metabolic imaging predicts response, survival, and recurrence in adenocarcinomas of the esophagogastric junction. *J Clin Oncol*. 2006;24:4692–4698.
- van Rossum PS, van Lier AL, 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.
- Boggs DH, Hanna A, Burrows W, Horiba N, Suntharalingam M. Primary gross tumor volume is an important prognostic factor in locally advanced esophageal cancer patients treated with trimodality therapy. *J Gastrointest Cancer*. 2015;46:131–137.
- Pan L, Gu P, Huang G, Xue H, Wu S. Prognostic significance of SUV on PET/CT in patients with esophageal cancer: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol*. 2009;21:1008–1015.
- Diaz-Cano SJ. Tumor heterogeneity: mechanisms and bases for a reliable application of molecular marker design. *Int J Mol Sci*. 2012;13:1951–2011.
- Harris AL. Hypoxia: a key regulatory factor in tumour growth. *Nat Rev Cancer*. 2002;2:38–47.
- Rajendran JG, Schwartz DL, O'Sullivan J, et al. Tumor hypoxia imaging with [ $^{18}\text{F}$ ] fluoromisonidazole positron emission tomography in head and neck cancer. *Clin Cancer Res*. 2006;12:5435–5441.
- Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in  $^{18}\text{F}$ -FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging*. 2013;40:133–140.
- Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour  $^{18}\text{F}$ -FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging*. 2013;40:1662–1671.
- Tan S, Kligerman S, Chen W, et al. Spatial-temporal [ $^{18}\text{F}$ ]FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2013;85:1375–1382.
- Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline  $^{18}\text{F}$ -FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med*. 2011;52:369–378.
- van Rossum PS, Fried DV, Zhang L, et al. The incremental value of subjective and quantitative assessment of  $^{18}\text{F}$ -FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J Nucl Med*. 2016;57:691–700.
- Zhang H, Tan S, Chen W, et al. Modeling pathologic response of esophageal cancer to chemoradiation therapy using spatial-temporal  $^{18}\text{F}$ -FDG PET features, clinical parameters, and demographics. *Int J Radiat Oncol Biol Phys*. 2014;88:195–203.
- Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th ed. of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17:1471–1474.
- Boellaard R, O'Doherty MJ, Weber WA, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging*. 2010;37:181–200.
- Muijs C, Smit J, Karrenbeld A, et al. Residual tumor after neoadjuvant chemoradiation outside the radiation therapy target volume: a new prognostic factor for survival in esophageal cancer. *Int J Radiat Oncol Biol Phys*. 2014;88:845–852.
- Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma: clinicopathologic correlations. *Cancer*. 1994;73:2680–2686.
- Hatt M, Majdoub M, Vallieres M, et al.  $^{18}\text{F}$ -FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med*. 2015;56:38–44.
- Doane DP. Aesthetic frequency classifications. *Am Stat*. 1976;30:181–183.
- Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;3:610–621.
- Galloway MM. Texture analysis using gray level run lengths. *Comput Graph Image Process*. 1975;4:172–179.
- Thibault G, Angulo J, Meyer F. Advanced statistical matrices for texture characterization: application to cell classification. *IEEE Trans Biomed Eng*. 2014;61:630–637.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19:716–723.
- Kwee RM. Prediction of tumor response to neoadjuvant therapy in patients with esophageal cancer with use of  $^{18}\text{F}$  FDG PET: a systematic review. *Radiology*. 2010;254:707–717.
- Leijenaar RT, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep*. 2015;5:11075.
- Sindoni A, Minutoli F, Pontoriero A, Iati G, Baldari S, Pergolizzi S. Usefulness of four dimensional (4D) PET/CT imaging in the evaluation of thoracic lesions and in radiotherapy planning: Review of the literature. *Lung Cancer*. 2016;96:78–86.
- Yue J, Yang Y, Cabrera AR, et al. Measuring tumor hypoxia with  $^{18}\text{F}$ -FETNIM PET in esophageal squamous cell carcinoma: a pilot clinical study. *Dis Esophagus*. 2012;25:54–61.
- Yue J, Chen L, Cabrera AR, et al. Measuring tumor cell proliferation with  $^{18}\text{F}$ -FLT PET during radiotherapy of esophageal squamous cell carcinoma: a pilot clinical study. *J Nucl Med*. 2010;51:528–534.
- Honing J, Pavlov KV, Mul VE, et al. CD44, SHH and SOX2 as novel biomarkers in esophageal cancer patients treated with neoadjuvant chemoradiotherapy. *Radiother Oncol*. 2015;117:152–158.
- Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*. 2000;92:205–216.
- 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.