Limited Predictive Value of FDG-PET for Response Assessment in the Preoperative Treatment of Esophageal Cancer: Results of a Prospective Multi-Center Trial (SAKK 75/02)

Bernd Klaesera,b, Egbert Nitzschec, Jan C. Schullerd, Dieter Köberle, Lucas Widmerf, Sabine Balmer-Majong, Thomas Hanhy, Corinne Cescato-Wengeri, Peter Brauchlidi, Michael Zündk, Bernhard C. Pestalozzi, Clemens Casparm, Susanne Albrecht, Roger von Mooos, Thomas Ruhstaller
e

a Department of Nuclear Medicine, Inselpital Bern, b Department of Nuclear Medicine, Kantonsspital St. Gallen, c Department of Nuclear Medicine, Kantonsspital Aarau, d Coordinating Center SAKK, Bern, e Department of Oncology and Hematology, Kantonsspital St. Gallen, f Department of Oncology, Triemlispital Zurich, g Department of Radiation Oncology, University Hospital of Geneva, h Department of Nuclear Medicine, University Hospital Zurich, i Department of Medical Oncology, Claraspital Basel, j Department of Surgery, Kantonsspital St. Gallen, k Department of Oncology, University Hospital Zurich, l Department of Oncology and Hematology, Kantonsspital Baden, m Department of Nuclear Medicine, University Hospital of Geneva, n Department of Medical Oncology and Hematology, Kantonsspital Graubünden, Chur, Switzerland

Key Words
Esophageal cancer · Positron emission tomography (PET) · Predictive marker · Response assessment

Summary
Background: Only responding patients benefit from preoperative therapy for locally advanced esophageal carcinoma. Early detection of non-responders may avoid futile treatment and delayed surgery.

Patients and Methods: In a multi-center phase II trial, patients with resectable, locally advanced esophageal carcinoma were treated with 2 cycles of induction chemotherapy followed by chemoradiotherapy (CRT) and surgery. Positron emission tomography with 2-[fluorine-18]fluoro-2-deoxy-o-glucose (FDG-PET) was performed at baseline and after induction chemotherapy. The metabolic response was correlated with tumor regression grade (TRG). A decrease in FDG tumor uptake of less than 40% was prospectively hypothesized as a predictor for histopathological non-response (TRG > 2) after CRT. Results: 45 patients were included. The median decrease in FDG tumor uptake after chemoradiotherapy correlated well with TRG after completion of CRT (p = 0.021). For an individual patient, less than 40% decrease in FDG tumor uptake after induction chemotherapy predicted histopathological non-response after completion of CRT, with a sensitivity of 68% and a specificity of 52% (positive predictive value 58%, negative predictive value 63%). Conclusions: Metabolic response correlated with histopathology after preoperative therapy. However, FDG-PET did not predict non-response after induction chemotherapy with sufficient clinical accuracy to justify withdrawal of subsequent CRT and selection of patients to proceed directly to surgery.

Schlüsselwörter
Ösophaguskarzinom · Positronenmissionstomografie (PET) · Marker, prädiktiver · Response-Vorhersage

Zusammenfassung
**Introduction**

Esophageal carcinoma is often diagnosed in the locally advanced stage and is associated with a poor long-term outcome of only 20–30% survival at 2 years. Chemoradiotherapy (CRT) followed by surgery has been integrated into standard treatment. However, only patients who respond to neoadjuvant therapy and achieve an R0 resection at surgery have a substantial long-term survival, while non-responders may not benefit [1] but experience adverse effects or even tumor progression. In this context, early differentiation of non-responders from responders is desirable to prevent non-responders from receiving inefficient chemotherapy and delayed surgery. Positron emission tomography with 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG-PET) visualizes the enhanced glucose metabolism of viable esophageal tumor tissue [2, 3]. A number of studies have addressed the feasibility of measuring changes in tumor FDG uptake to monitor response and to predict the outcome of neoadjuvant treatment.

Promising results for early response assessment after neoadjuvant chemotherapy without concomitant radiotherapy were published for adenocarcinomas of the esophagogastric junction [4]. Metabolic response was defined by a retrospectively calculated cut-off of at least 35% decrease in tumor FDG uptake after 14 days of chemotherapy. Metabolic response correlated significantly with clinical and histopathological response, disease-free survival and overall survival (OS). A subsequent study by the same group prospectively tested this cut-off and confirmed that metabolic non-responders had a low histopathological response rate of only 5% and a poor prognosis compared with responders [5].

In contrast, single-center studies of CRT for adenocarcinoma and/or squamous cell carcinoma (SCC), designed to assess the correlation between a decrease in tumor FDG uptake and histopathological response [6–15] and survival [6, 7, 9, 13, 15], have produced conflicting results. A group from Munich, Germany, evaluated PET for the response assessment in the course of CRT of SCC. They found PET, after completion of CRT and early in the course of CRT, to have similar predictive values for histopathological response [9, 15].

Subsequently, the same group prospectively tested a predefined cut-off of 30% decrease in tumor standardized uptake values (SUV) after 2 weeks of CRT. However, the negative predictive value (NPV) for histopathological response remained clearly below 80% and, consequently, did not meet the requirement for a clinically relevant predictive test [16]. Radiation-induced esophagitis, which can mask treatment-induced reduction of SUV, may confound metabolic response evaluation after CRT [6, 8, 10, 11, 15].

This study was designed to quantify metabolic tumor response without interference from radiation-induced esophagitis, by determining whether changes in tumor FDG uptake after induction chemotherapy, but prior to CRT, can predict histopathological non-response after completion of neoadjuvant treatment. A cut-off of 40% SUV decrease for the differentiation of non-responders from responders was prospectively defined, based on data from previous studies [4, 9]. The primary objective was to predict histopathological non-response. Secondary objectives were to correlate metabolic response with event-free survival (EFS) and OS and to determine whether metabolic response may be a useful prognostic parameter.

**Patients and Methods**

**Patients**

PET imaging was performed as part of a prospective multi-center phase II trial investigating a cisplatin- and taxane-based regimen of 2 cycles of induction chemotherapy, followed by CRT and surgery (Swiss Group for Clinical Cancer Research, SAKK 75/02, NCT00072033) [17]. The protocol was approved by the local ethics committees of the participating centers. The study population consisted of previously untreated patients with histologically confirmed locally advanced but resectable SCC and adenocarcinoma of the thoracic esophagus or esophagogastric junction classified as clinical stage T3 N0, T1–3 N+ or T4 Nx, according to the AJCC Cancer Staging Manual, 6th edition [18]. Written informed consent was obtained for all patients. Contrast-enhanced computed tomography (CT) of the thorax and abdomen and endoluminal ultrasound of the esophagus were performed in all patients before PET or PET-CT. PET or PET-CT staging was strongly recommended in the study protocol, but was not mandatory because PET scanners were not available in all regions of Switzerland. Patients with distant metastases detected during pre-treatment evaluation were excluded.

**PET Imaging**

PET or PET-CT was performed at up to 4 weeks before initiation of induction chemotherapy and after completion of 2 cycles of induction chemotherapy in week 5, before the start of CRT (fig. 1). Both PET scans were performed at the same center on the same PET or PET-CT machine and under identical conditions for each patient. All PET scanners fulfilled the quality requirements defined by the Swiss Society of Nuclear Medicine and had a spatial resolution of 6 mm or less.

Patients fasted for at least 6 h before an intravenous injection of 5 MBq FDG/kg bodyweight. The blood glucose level was recorded in all patients. Patients were examined according to the local acquisition protocols at each center, and acquisition parameters were kept constant for both PET scans with regard to the time point of acquisition after tracer injection. Maximum SUV (SUV_{max}) of the primary tumor was calculated to semi-quantify FDG tumor uptake. Percentage changes of SUV_{max} between baseline PET and PET after induction chemotherapy were calculated to quantify metabolic response. PET data were evaluated at each site and reviewed centrally.

**Preoperative Treatment**

The preoperative treatment regimen consisted of induction chemotherapy with intravenous cisplatin 75 mg/m² and docetaxel 75 mg/m² on days 1 and 21, followed by radiotherapy (total dose 45 Gy) and concurrent chemotherapy comprising intravenous cisplatin 25 mg/m² and docetaxel 20 mg/m² weekly for 5 weeks. Surgery was scheduled 3 to 8 weeks after CRT (fig. 1). Patients with evidence of newly detected stage M1 and/or inoperable T4 disease were not eligible for surgery.

**Criteria for Response**

Histopathological response was based on pathological findings after esophagectomy. Specimens were examined according to standardized procedures in local pathology laboratories, and all specimens were centrally reviewed at the University of Basel by an experienced pathologist.

**Predictive Value of FDG-PET in the Preoperative Treatment of Esophageal Cancer**

Onkologie 2009;52:724–730

725
The histopathological response was classified according to the Mandard classification of tumor regression grade (TRG) [19]. TRG 1 was defined as complete regression, TRG 2 as presence of rare residual cancer cells scattered throughout the fibrosis, and TRG > 2 as an increased number of residual cancer cells. Patients with complete regression (TRG 1) or near-complete regression (TRG 2) were classified as responders, while patients with partial and minimal response and no change (TRG 3-5) were classified as non-responders, as previously validated [15].

A cut-off of 40% decrease in tumor FDG uptake between initial and second FDG-PET scan was prospectively hypothesized to discriminate non-responders from responders, based on previously published optimal cut-off values for response assessment after chemotherapy and CRT of adenocarcinoma and SCC [4, 9]. More than 40% decrease in FDG uptake from baseline to the second chemotherapy cycle was considered as a metabolic response to chemotherapy.

Statistical Methods
The sensitivity, positive predictive value (PPV), specificity, and negative predictive value (NPV) of metabolic non-response (i.e. less than 40% decrease in FDG uptake) predicting TRG > 2 was calculated. Survival endpoints and 95% confidence intervals (CIs) were calculated using the Kaplan-Meier method. The survival times of strata (metabolic responders vs. non-responders, adenocarcinoma vs. SCC, histopathological responders vs. non-responders) were compared with the log-rank test. Overall survival was calculated as the time from registration until death (event) or last follow-up (censored). EFS was calculated as time from registration until death or progression (event) or last follow-up (censored). The reverse Kaplan-Meier method was used to calculate the median follow-up time. The difference of means or medians was tested with the t-test or Wilcoxon’s test. Calculations and plots were performed with SAS 9.1 and S-Plus 7.0.

Results

Patient and Tumor Characteristics
Of 66 patients enrolled in the SAKK 75/02 trial, 57 underwent surgical tumor resection and assessment of histopathological response. From this population, 45 patients were included in the PET study. The remaining 12 patients were not included for the following reasons: baseline PET scans not performed (9 patients; 6 because study centers did not participate in the study), 1 patient refused, and 2 scans were not correctly scheduled. In 2 patients, PET uptake was not measurable due to a patient refusal. The median age was 61 years for patients with adenocarcinoma (range 48–71 years) and 59 years for those with SCC (range 44–70 years). The patient characteristics are listed in table 1.

Change of Tumor FDG Uptake after Two Cycles of Chemotherapy
The median relative change in FDG uptake after induction chemotherapy compared with baseline uptake was ~53% for patients with histological complete or near-complete response (TRG 1/2) and ~31% for non-responders with an increased number of residual cancer cells (TRG > 2). This difference was statistically significant (Wilcoxon’s test: p = 0.021; fig. 2).

Prediction of Histopathological Non-Response after CRT
The prospectively defined cut-off value of less than 40% decrease in FDG uptake after 2 cycles of induction chemotherapy did not reliably predict pathological non-response after completion of CRT (fig. 3). The overall sensitivity and specificity for prediction of non-response were 68 and 52%, respectively, resulting in a PPV of 58% and an NPV of 63%. The sensitivity, specificity, PPV and NPV for non-response were 60, 50, 64 and 45% for adenocarcinoma and 86, 54, 50 and 88% for SCC, respectively. Moreover, no significant differences in the prediction of non-response according to SUV decrease were found between patients presenting with adenocarcinoma or SCC.
Predictive Value of FDG-PET in the Preoperative Treatment of Esophageal Cancer

Figure 4 shows retrospectively calculated PPV for different cut-off values for tumor SUV decrease after chemotherapy. Of note, only an increase in FDG uptake would provide a markedly better PPV than the predefined cut-off of less than 40% decrease. However, this accounted for only 5 patients. In addition, 11 patients classified as metabolic non-responders after induction chemotherapy (5 adenocarcinoma, 6 SCC, patients originating from different participating centers) had histological complete or near-complete response after completion of CRT.

**Table 1. Patient and tumor characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Adenocarcinoma (n = 25)</th>
<th>Squamous cell carcinoma (n = 20)</th>
<th>All (n = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>Median</td>
<td>61</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>48–71</td>
<td>44–70</td>
</tr>
<tr>
<td>Median follow-up, months</td>
<td></td>
<td>27</td>
<td>30</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>17</td>
<td>40</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Clinical stage at diagnosis</td>
<td></td>
<td>T1 N1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2 N1</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T3 N0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T3 N1</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T3 Nx</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T4 N1</td>
<td>0</td>
</tr>
<tr>
<td>WHO tumor grading</td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not done</td>
<td>2</td>
</tr>
<tr>
<td>Dysphagia grade at diagnosis (NCI-CTC grading v2.0)</td>
<td></td>
<td>0/1</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Tumor regression grade</td>
<td></td>
<td>TRG 1 (pCR)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TRG 2</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TRG 3-5</td>
<td>15</td>
</tr>
</tbody>
</table>

WHO, World Health Organization; NCI-CTC, National Cancer Institute common toxicity criteria; TRG, tumor regression grade; pCR, pathological complete response.

**Prognostic Significance of Metabolic Response to Treatment**

Median follow-up was 28 months (adenocarcinoma 27 months, SCC 30 months). Median OS was 32.4 months (95% CI lower limit 25.7 months). Metabolic responders according to the predefined criterion of at least 40% decrease in tumor FDG uptake had a median OS of 36.5 months (95% CI 8.0–36.5 months) compared with 31.0 months (95% CI lower limit 21.4 months) for non-responders. Median EFS was 22.8 months (95% CI lower limit 7.3 months) for metabolic responders and 21.7 months (95% CI 9.8–28.3 months) for non-responders. Differences in OS and EFS failed to reach significance (log-rank test).

**Discussion**

To date, the utility of FDG-PET for early response assessment has been shown only after neoadjuvant chemotherapy alone. Promising results using retrospectively calculated or prospectively defined cut-off values have been published based on extensive work at a single center [4, 5, 20–22], and prospective multi-center trials are planned [23]. However, a role for FDG-PET in the response assessment after combined CRT has not been established and published single-center studies have shown conflicting results. While some studies suggested a possible correlation between the decrease in FDG tumor uptake and histopathological response [7–9, 15], these results
were not supported for adenocarcinoma or SCC by other, subsequent trials [10–14, 16]. Radiation-induced esophagitis correlates metabolically with FDG-PET in up to 60% of patients after radiotherapy and may explain obvious differences in response assessment after chemotherapy and CRT [6, 8, 10, 11, 15].

The patients in the current study were treated with preoperative induction chemotherapy followed by CRT. It is critical to identify non-responders to neoadjuvant treatment so that they can be scheduled for immediate surgery. No optimal cut-off values for response assessment in this setting were available. Therefore, based on previously published studies with optimal cut-off values for response assessment after chemotherapy alone [40%], adenocarcinoma [4]) and after completion of combined CRT (52%, SCC [9]), it was hypothesized that a prospectively defined cut-off of less than 40% decrease in tumor FDG would identify non-responding adenocarcinoma and/or SCC after completion of induction chemotherapy.

For the whole group, the median change of FDG tumor uptake after 2 cycles of induction chemotherapy correlated significantly with the TRG after completion of CRT (p = 0.021). These results were in accordance with those published for response assessment after chemotherapy alone [4, 5] but without chemoradiation before surgery. After a median follow-up of 28 months there was a trend towards prolonged OS for metabolic responders vs. non-responders after induction chemotherapy. No differences in EFS were found between metabolic responders and non-responders. Therefore, these results did not confirm previously published single-center studies describing a significant correlation between metabolic response and EFS and/or OS [4, 6, 7, 9, 20].

The hypothesis that FDG-PET with a predefined cut-off of less than 40% decrease in tumor SUV after induction chemotherapy may predict histological non-response and treatment failure was not proven. The PPV for prediction of non-response of 58% was not sufficiently accurate to justify withdrawal of CRT in metabolic non-responders and to select patients to proceed directly to surgery. A retrospective calculation of an optimal cut-off did not improve the results (fig. 4) because only an increase in SUV reliably pointed towards non-response. These results were less favorable than those of other studies of adenocarcinoma in which the second PET was performed 2 weeks after initiation of chemotherapy alone (PPV for non-response, 95–100%) [4, 5]. Lordick et al. [20] noted no histological responders among metabolic non-responders (less than 35% SUV decrease; response assessment 2 weeks after initiation of chemotherapy), while in the present study 11 of 26 metabolic non-responders (5 adenocarcinoma, 6 SCC, less than 40% SUV decrease) had histological complete or near-complete response after completion of CRT (specificity 52%; fig. 3). Apparently, subsequent CRT can transform some metabolic non-responders after chemotherapy into histopathological responders. These 11 metabolic non-responders after induction chemotherapy with subsequent histopathological response (TRG 1/2) after completion of CRT had median EFS similar to all patients with histopathological response. It can be concluded that histopathological response was a better predictor of outcome than metabolic response in this study. Brücher et al. reported promising results for SCC evaluated 3 weeks after completion of CRT (PPV for non-response 100%, with a retrospectively defined cut-off of 52%) [9]. However, in a follow-up study by the same group, using a prospectively
defined cut-off of 30%, the PPV for non-response (64%) did not meet the requirement for a clinically relevant predictive test [16] and was in the same range as data from Song et al. [8] using a second PET performed 8 weeks after CRT.

To the authors’ knowledge, this is the first study prospectively testing the predictive and prognostic value of FDG-PET in esophageal cancer in a nationwide multi-center context. A strong point of this study is the assessment of non-response as the objective of clinical relevance using a prospectively defined cut-off of tumor SUV decrease during the course of treatment. A new approach using a second PET after induction chemotherapy and before the onset of combined CRT was evaluated to avoid problems with semi-quantification of tumor activity, i.e. radiation-induced esophagitis [10]. A limitation of this study is the small number of patients and maybe the combined evaluation of adenocarcinoma and SCC. Recently, separate trials for adenocarcinoma and SCC have been proposed because the thresholds for measuring PET response may vary [23]. However, all patients in the study received identical treatment and no significant differences in the prediction of non-response according to SUV decrease were found between adenocarcinoma and SCC. Using a multi-center design means that the mode of acquisition and the instrumentation were standardized to a lesser degree than would be possible in a single-center study. Nevertheless, only relative changes of FDG uptake in the pre- and post-treatment scans were evaluated, and all acquisition parameters were kept uniform in individual patients. Under such conditions, no significant effects of methodological variations, i.e. acquisition parameters, timing of the acquisition after FDG injection, reconstruction algorithm and method of SUV measurement, have been described to date [13, 23–25], indicating that the results of this trial can be regarded as reliable. Consensus recommendations for the use of FDG-PET as an indicator of therapeutic response in multi-center trials, which have been published in the meantime [26], had already been essentially respected in the design of this trial.

In conclusion, metabolic response after induction chemotherapy correlated well with histopathological response after completion of CRT. However, response assessment with FDG-PET using semi-quantitative uptake measurement was not able to predict histological non-response after neoadjuvant treatment with sufficient clinical accuracy to select patients to proceed directly to surgery. It remains to be determined whether a more reliable response prediction may be possible by employing recent technical developments including high-resolution PET-CT, partial volume and recovery correction, and kinetic modeling.

Acknowledgements
We thank Sandra Thierstein, trial coordinator of the SAKK Bern, for her tremendous administrative support for this trial, and Pierluigi Ballabeni, Ph.D., for assistance in designing this trial. The Swiss Group for Clinical Cancer Research (SAKK) was the sponsor of the trial and was financially supported by the Federal Government of Switzerland and Sanofi-Aventis Switzerland.

Conflict of Interest Statement
All authors: None declared.

References