Early prediction of response to first-line chemotherapy by sequential $[^{18}\text{F}]$-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with advanced colorectal cancer


1Department of Oncology and Pathology, Karolinska Institute, Stockholm; 2Department of Oncology, Radiology and Clinical Immunology, Uppsala University, Uppsala; 3Department of Medical Sciences, Nuclear Medicine, University Hospital, Uppsala and 4Department of Nuclear Medicine, Karolinska Institute, Stockholm, Sweden

Received 18 June 2008; revised 9 November 2008; accepted 11 November 2008

Background: To evaluate $[^{18}\text{F}]$-2-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET), for early evaluation of response to palliative chemotherapy and for prediction of long-term outcome, in patients with metastatic colorectal cancer (mCRC).

Patients and methods: In a randomized trial, patients with mCRC received irinotecan-based combination chemotherapy. FDG–PET was carried out before treatment and after two cycles in 51 patients at two centers. Visual changes in tumor FDG uptake and changes measured semi-automatically, as standard uptake values (SUVs), were compared with radiological response after four and eight cycles.

Results: The mean baseline SUV for all tumor lesions per patient was higher in nonresponders than in responders (mean 7.4 versus 5.6, $P = 0.02$). There was a strong correlation between metabolic response (changes in SUV) and objective response ($r = 0.57$, $P = 0.00001$), with a sensitivity of 77% and a specificity of 76%. There was no significant correlation between metabolic response and time to progression ($P = 0.5$) or overall survival ($P = 0.1$).

Conclusions: Although metabolic response assessed by FDG–PET reflects radiological tumor volume changes, the sensitivity and specificity are too low to support the routine use of PET in mCRC. Furthermore, PET failed to reflect long-term outcome and can, thus, not be used as surrogate end point for hard endpoint benefit.

Key words: colorectal cancer, FDG–PET, palliative chemotherapy, prediction, response evaluation

Introduction

Palliative treatment of metastatic colorectal cancer (mCRC) with chemotherapy combinations in first-line settings yields response rates in ~50%. Treatment evaluation is mostly based on imaging by computerized tomography (CT) after three to four cycles of chemotherapy. Earlier evaluation of response would allow for avoidance of unnecessary toxicity and costs for nonresponding patients and a change to a potentially more active treatment. Imaging of tumor metabolism with positron emission tomography (PET) using $[^{18}\text{F}]$-2-fluoro-2-deoxy-D-glucose (FDG) can visualize the enhanced glucose utilization in tumor tissue. Studies indicate that reduction of tumor glucose uptake after chemotherapy correlates with tumor regression in different tumor types [1–6]. Changes in glucose metabolism may also allow prediction of subsequent response before the reduction of tumor size [1–5, 7, 8].

In mCRC patients, the experience is limited to small studies in patients with liver metastases [9–12]. One report showed that FDG–PET 4–5 weeks after start of chemotherapy predicted the effect with a sensitivity of 100% and specificity of 90% on a per-lesion basis in 18 patients [12]. In another study in 50 heterogeneously treated patients, the FDG–PET changes after 2 months predicted progression-free survival and overall survival (OS) [10]. The aim of this prospective study was to assess the value of FDG–PET, for early evaluation of response and for prediction of long-term outcome, in patients with mCRC treated with first-line combination chemotherapy.

Patients and methods

Patients

In a randomized multicenter trial [13], patients with mCRC received irinotecan with either the Nordic bolus 5-fluorouracil (5-FU) and folinic acid schedule (FLIRI) or the de Gramont schedule (Lv5FU2-CPT11).
Patients at two centers, Stockholm and Uppsala, Sweden, were asked to participate in this substudy from January 2002 to March 2004. Altogether, 71 patients were randomized and 57 patients accepted participation. Six patients were not assessable, one due to myocardial infarction and five due to logistic problems to get PET or CT examinations on time. Demographic characteristics and response to chemotherapy in the 51 assessable patients are similar to those in the entire study [13]. Median age was 59 (range 42–75) years and 35% were females. Liver (n = 48), lungs (n = 13) and lymph nodes (n = 11) were the predominant metastatic sites. The trial protocol, including the PET substudy, was approved by the regional ethical committees.

**chemotherapy**

FLIRI consists of irinotecan 180 mg/m² day 1, followed by 5-FU 500 mg/m² bolus i.v. injection and 40 min later leucovorin (Lv) 60 mg/m² bolus i.v. for two consecutive days. Lv5FU2-IRI consists of irinotecan 180 mg/m² day 1, 5-FU 400 mg/m² bolus injection and Lv 200 mg/m² 2-h infusion for two consecutive days and 5-FU 600 mg/m² as a 22-h infusion days 1–2. All treatments were repeated every 14 days. Patients were generally treated until disease progression. After disease progression, second- and third-line chemotherapy was at the discretion of the physician in charge of the patient.

**PET imaging**

PET was made 1–14 days before start of treatment and immediately before the third cycle. Examinations were made in an ECAT EXACT 31 or HR + PET camera (CTI, Knoxville, TN) after at least 4-h fasting and testing that the patient was normoglycemic. Series of consecutive 10-min scans including the trunk and the neck was initiated 60 min after i.v. administration of 400 MBq FDG. Attenuation correction was carried out using the transmission-based attenuation maps obtained from measurements of 511 keV photons at each 10-cm wide bed position emitted from three symmetrically located 68Ge/68Ga rods around the patient. The transmission data were added into the attenuation correction routine in the software of the camera. Subtraction of scattered events in the images was carried out in a similar way. Images were reconstructed using the ordinary iterative OSEM reconstruction algorithm of the manufacturer.

**image analysis and assessment of response**

PET evaluations were carried out by a nuclear medicine physician (UG) blinded for the clinical and radiological evaluation. Each scan was read visually in iterative and filtered back projected reconstructions. For the PET evaluations, the MultiModality software by Hermes Medical Solutions (Stockholm, Sweden). The pre-treatment study was automatically aligned with the post-treatment study by the Automatic Registration Tool. Up to five index lesions were chosen in the liver and/or the lungs in the pre-treatment study. The SUVs were calculated in regions of interest (ROI) drawn manually around the target lesions. For each lesion, the hot spot consisting of four up to six pixels in the transversal slice containing the highest uptake was determined by narrowing the cutoff values when applying the automatic ROI tool. This procedure was repeated in the post-treatment study. In lesions where no increase of uptake could be detected visually, the ROI from the pre-treatment study was mirrored, and the procedure carried out accordingly. Studies on the reproducibility of the FDG signal in malignant tumors indicate that PET imaging can reliably measure changes by >20% of the baseline value. A change of SUV by >25% during treatment is accepted as a surrogate for tumor response, according to the European Organization for Research and Treatment of Cancer guidelines [14]. Metabolic complete response (PET-CR) was defined as no difference in SUV in the area of the former lesion compared with the background, metabolic partial response (PET-PR) as a decrease in SUV by ≥25% and metabolic progressive disease (PET-PD) as when the SUV increased by ≥25% or new lesions were found. Changes in SUV in between PR and PD were scored as metabolic stable disease (PET-SD).

A qualitative, visual PET response assessment, before and independent from the SUV calculations and radiological evaluation, was done (UG). In half of the cases, a similar blinded evaluation was done by another nuclear medicine physician (HI) with complete agreement in all cases.

**radiological response evaluation and follow-up**

CT scans of chest and abdomen were carried out at baseline and repeated after every four cycles for assessment of radiological (objective) tumor response. Tumor response was evaluated according to RECIST [15] by an experienced radiologist (AS) who was blinded for clinical outcome, radiological response evaluation and PET imaging. The best overall response achieved with the first-line chemotherapy regimen was determined and correlated with changes in metabolic response. OS and time to tumor progression (TTP) defined as first observation of radiological tumor progression whether under active treatment or not, and time to failure on strategy (TFS) [16] was calculated from the date of randomization. Even if the strategy in the trial was not to make a break in the treatment, some responding patients had this. Since breaks could influence the relations between PET response and TTP, we included TFS for comparison.

**statistical analyses**

The primary end point was the correlation between metabolic response and radiological response and secondary end points the correlations to TTP, TFS and OS. Linear regression and Spearman’s rank correlation coefficient (rho) were used to describe the correlations. Median OS, TTP and TFS were estimated according to the Kaplan–Meier method. Survival times in patients with and without a metabolic or an objective response were compared with the log-rank test. The Landmark method [17], to compensate for a longer guarantee time of responders, was not needed since all assessable patients had a follow-up CT investigation after 2 months and the median time to when the best overall response category was reached was 2 months (range 2–4 months). Most objective responses were seen at the first CT evaluation, although additional decrease in tumor size was frequently also seen at subsequent evaluations.

**results**

changes of metabolic activity during chemotherapy and radiological tumor response

Of the 51 assessable patients, 27 received FLIRI and 28 Lv5FU2-IRI. The overall response rate according to CT was 43% (22 of 51 patients). Since we could not detect any differences in the relations between PET response and outcomes between the treatment groups, they were analyzed together. The SUV evaluation was carried out on median five tumor lesions per patient (1–8). The majority of tumor lesions were located in the liver [involved organs; liver (n = 48), lungs (n = 13) and lymph nodes (n = 11)]. Metastasis in lungs and lymph nodes were generally smaller compared with the liver metastasis, so PET evaluations were carried out mostly on liver metastasis. PET evaluation of liver metastasis was carried out on 47 patients, of lung metastasis on eight patients and of lymph node metastasis on seven patients. The mean SUV hot spot at baseline varied between 0.6 and 13.8 (mean 6.7). In the group of radiological responders, the mean baseline SUV was 5.6 (4.4–6.8 ± 95%
PET examination revealed unexpected extrahepatic disease. Discussion for metastatic surgery, but were disqualified when the three had no objective response. Additional patients were Four patients had surgery without metabolic response, of which these had metabolic response and six also radiological response. The 1-year survival rates for PET responders and PET nonresponders were 96% for objective responders and 60% for nonresponders. These rates nonresponders were 78% and 57%, respectively. These rates The correlation between objective response and changes in SUV was higher when the mean value of the hot spot of all evaluated lesions were used \( r = 0.57, P = 0.00001 \) than when the hot spot from the hottest tumor lesion was used \( r = 0.43, P = 0.0014 \). The visual qualitative PET response assessment and the semiquantitative SUV evaluation were strongly correlated \( r = 0.79, P < 10^{-11} \).

patient follow-up and survival

Median follow-up time was 19 months. During this period, 35 patients died. Median TTP was 9.7 months, median TFS 9.9 months and median OS 21.9 months. For patients with metabolic response, median TTP was 10.8 months and TFS 12.3 months, whereas TTP and TFS were 8.8 months (log-rank \( P = 0.7 \)) for patients without metabolic response. For patients with radiological response, median TTP was 15.1 months and TFS 17.5 months, whereas TTP was 7.2 months (log-rank \( P = 0.01 \)) and TFS 7.2 months (log-rank \( P = 0.02 \)) for patients without radiological response. Median OS for PET responders was 24.9 months, for PET nonresponders 21.5 months \( (P = 0.11) \), for objective responders 28.2 months and for objective nonresponders 16.1 months \( (P = 0.0002) \) (Figure 1). The 1-year survival rates for PET responders and PET nonresponders were 78% and 57%, respectively. These rates were 96% for objective responders and 60% for nonresponders. Eleven patients had surgery for liver metastases. Seven of these had metabolic response and six also radiological response. Four patients had surgery without metabolic response, of which three had no objective response. Additional patients were discussed for metastatic surgery, but were disqualified when the PET examination revealed unexpected extrahepatic disease.

discussion

Early metabolic response evaluation by FDG–PET, after two cycles of chemotherapy for patients with mCRC, showed positive and negative predictive values of 71% and 81%, respectively, for standard radiological response. In contrast to radiologic response, metabolic response did not reflect survival indicating that factors additional to the immediate inhibition of tumor cell metabolism influence long-term outcome.

<table>
<thead>
<tr>
<th>Objective response</th>
<th>PET response</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>PR</td>
<td>SD</td>
</tr>
<tr>
<td>CR</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>PR</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>SD</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>51</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Correlation between metabolic response in \(^{18}F\)-2-fluoro-2-deoxy-D-glucose positron emission tomography and subsequent best overall response according to RECIST

Spearman rank order correlation \( P \text{ level } = 0.00007, \text{ sensitivity } = 0.77 \) and specificity = 0.76.

PET, positron emission tomography; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

Figure 1. Kaplan–Meier curves comparing the proportion of surviving patients over time (months) for (A) positron emission tomography (PET) responders \( (n = 24) \) and PET nonresponders \( (n = 27) \) \( (P = 0.11) \), and (B) for objective responders \( (n = 22) \) and objective nonresponders \( (n = 29) \) \( (P = 0.0002) \).
liver metastasis [9, 11, 18]. In the study [10], the patients were evaluated with FDG–PET after 2 months of treatment with various chemotherapy schedules and only about half of the 50 patients received first-line treatment. The clinical benefit of FDG–PET in response evaluation has otherwise more support in breast, esophageal and lung cancer [1, 3–8, 19, 20]. The strong correlation between metabolic response after 4 weeks and radiological response after 2–4 months of treatment, seen here, also indicates that PET imaging could be used to predict response and clinical outcome of chemotherapy in mCRC. The present data, seen in a homogeneously treated population participating in a prospective randomized study, however, do not fully support this conclusion. Thus, metabolic response was less efficient than radiologic assessment in discrimination of long-term outcome and the overall performance of PET as a test for objective response cannot be considered sufficient for replacement of radiological assessment. PET failed to early identify the four patients with PD as best objective response. In this study, SD and PD categories were combined as nonresponders in the analyses. With regard to clinical decision making, SD is included with responders, i.e. therapy continues as long as there is no progression. Only one patient had PD by PET evaluation, and this represents a clear limitation of the technology in this setting. It also limits the support to withdraw patients from treatment based only on an early PET evaluation. The PET technique for mCRC patients is thus not of sufficient clinical value in the first-line setting since ~50% of the patients have an objective response and an additional 30%–40% show SD. Its value in second and third lines, where fewer individuals have a response, may be greater.

Although baseline SUV predicted the probability of an objective response, but not the probability of a metabolic response, the correlation with objective response was rather weak and likely not clinically relevant. This information can be obtained much easier using simple clinical parameters [21, 22]. In patients with potentially resectable liver metastatic CRC, it is tempting to propose an early PET examination for staging and, as investigated here, for an early response evaluation [23, 24]. To support the first part of the proposal, a study in candidates for resection of colorectal liver metastases showed that 20% of the preoperative clinical management decisions, based on spiral CT of the thorax and abdomen, changed after FDG–PET [25]. When intrahepatic or extrahepatic unresectable disease was detected, the patients were spared unwarranted liver resections. On the other hand, the detection rate of liver metastasis on a lesion basis was generally better for CT than for FDG–PET (80% versus 65%). In lesions ≤1.5 cm in diameter, only 14% were detected by FDG–PET, compared with 80% by CT.

In summary, although metabolic response assessed by FDG–PET reflects radiological tumor volume changes, the sensitivity and specificity are too low to support the routine use of PET according to this protocol in first-line chemotherapy for mCRC. Furthermore, PET failed to reflect long-term outcome and can, thus, not be used as a surrogate end point for hard endpoint benefit. The use of PET as a staging tool with impact on patient treatment seems to have stronger support. However, PET performance will probably improve by the ongoing rapid technical development. Thus, more studies are needed to define the clinical role of PET imaging for treatment assessment in mCRC.

references


