Added value of positron emission tomography imaging in the surgical treatment of colorectal liver metastases


Objective [F-18]-Fluorodeoxyglucose-positron emission tomography (FDG-PET) is used increasingly in the work-up to surgery for patients with potentially resectable colorectal liver metastases. This study evaluates the clinical effectiveness, impact on health care resources and cost-effectiveness of adding FDG-PET to the diagnostic algorithm alongside a randomized clinical trial from a health care perspective.

Methods In a randomized clinical trial, the net monetary benefit (NMB) of FDG-PET added to conventional diagnostic work-up (CWU) was determined in patients with colorectal liver metastases. Seventy-five patients were included in each arm. Change in clinical management, futile laparotomies, preoperative findings and all relevant health care consumption were prospectively documented during 3 years. To assess health-related quality of life European Quality of Life-5 Dimensions was administered at the time of randomization, 3 and 6 weeks postoperatively, and every 3 months postoperatively for 3 years. Quality-adjusted life years (QALYs) were calculated based on European Quality of Life-5 Dimensions outcomes.

Results In adding FDG-PET, diagnostic performance increased and futile laparotomies were reduced by 38%. Both health-related quality of life and QALYs showed no significant difference between the CWU and PET groups. For CWU and PET groups costs were €92,836 and €81,776, respectively, accumulated in 3 years after randomization.

Introduction Survival of patients with colorectal cancer strongly depends on the development of distant metastases. Approximately 50–60% of the patients with primary colorectal tumour develop hepatic metastases [1,2]. These metastases are either identified at the time of diagnosis of the primary tumour (synchronous metastases) or during follow-up (metachronous metastases). Hepatic resection is the only curative therapy for a selected group of patients with colorectal liver metastases. For surgical treatment stringent selection criteria apply because no survival benefit is found if residual disease remains after hepatectomy. After resection, 5-year survival rates have been reported between 35 and 45% [3–6], compared with less than 15% in patients with unresectable colorectal liver metastases [7,8]. In up to 50% of the patients who undergo curative liver resection, recurrent disease is detected in the first year after surgery [9–12]. These figures strongly suggest that both before and during surgery minimal residual metastatic disease remains undetected. Therefore, a critical evaluation of preoperative and intraoperative diagnostic modalities is necessary with regard to both efficacy and cost-effectiveness.

Positron emission tomography (PET) with [F-18]-fluorodeoxyglucose (FDG) is a sensitive diagnostic tool for displaying both primary colorectal cancer and its metastases [13,14]. As a promising and upcoming modality for the evaluation of recurrent colorectal disease, FDG-PET imaging is based on metabolic changes instead of anatomical and structural changes, as is the case for computed tomography (CT) or magnetic resonance imaging (MRI). As such,
FDG-PET has the potential to show tumour activity ahead of CT or magnetic resonance imaging [15]. Furthermore, data suggest that additional FDG-PET results in a change in clinical management of up to 30% during the diagnostic work-up of patients eligible for surgical treatment of metastatic colorectal cancer [16–18].

The added value of the integration of FDG-PET in the diagnostic algorithm for surgical treatment of colorectal liver metastases has yet to be determined, despite increased sensitivity and specificity of FDG-PET as reported in the literature. Several studies have examined the cost-effectiveness of the FDG-PET in a number of diagnostic scenarios through modelling [19–21], whereas others have emphasized the cost-effectiveness of hepatic resection [22,23]. These studies attempt to show that FDG-PET is cost-effective in colorectal metastatic disease when added to conventional work-up (CWU), although strong data have yet to emerge. To increase the level of evidence of the value of adding FDG-PET to the CWU, a randomized controlled study [24] was conducted. Alongside this study the costs from a healthcare perspective of the CWU versus CWU supplemented by FDG-PET was evaluated.

Materials and methods

Study design

Patients were enrolled in a randomized, controlled, multicentre trial between May 2002 and February 2006. Eligible patients were required to have a history of histologically documented colorectal cancer treated with adequate surgical resection (tumour-free resection margins), suspicion of up to four potentially resectable colorectal liver metastases without evidence of extrahepatic metastatic disease (with exception of a maximum of two resectable lung metastases) on contrast-enhanced CT scan of the abdomen, pelvis and chest and no signs of recurrent or second colorectal carcinoma.

Randomization was done at the central trial office. When randomized to the conventional arm the patients were scheduled for hepatic surgery without any further diagnostic procedures. When randomized to the experimental arm, additional whole-body FDG-PET scanning was performed, generally within 2 weeks, and the results of both CT imaging and FDG-PET scan were again reported at a multidisciplinary oncology meeting. A joint assessment of all available data was made to review the clinical information and diagnostic imaging on a case-by-case basis.

For more details about study protocol, eligibility and standardized follow-up we refer to a recent paper by this group [24]. After hepatic resection all patients were followed prospectively at regular predetermined intervals of 3 months for the first 3 years after intervention. The study was approved by the institutional ethics review boards of all participating centres and all patients provided written informed consent.

Outcome measures and data collection

Clinical outcomes

Case record forms were used to collect data during the trial period. The standardized forms were used to record preoperative and demographic data, all diagnostic and perioperative information, essential costs during hospital stay and all follow-up data, including additional diagnostics for 3 years after hepatic resection.

Futile laparotomies were defined as any laparotomy that did not result in complete removal of the tumour, hepatic or extra-hepatic, which revealed benign disease at laparotomy or at histopathology, or that did not result in a disease-free survival period longer than 6 months. Disease-free and overall survival was assessed from the day of randomization.

Health-state values

Overall health-related quality of life (HRQoL) appraisals in different states of health can be expressed as a single metric number, often referred to as values or utilities. These health-state values are combined with survival data to compute quality-adjusted life-years (QALYs). One QALY equals 1 year in full health.

HRQoL during the first 3 years of follow-up was measured with the EQ-5D (European Quality of Life-5 Dimensions) [25–27]. Health status on the EQ-5D was described according to five attributes: mobility, self-care, usual activities, pain/discomfort and anxiety/depression; each of these attributes consists of three levels: no problems, some problems or severe problems. The HRQoL of both the groups was evaluated at time of randomization, 3 and 6 weeks after surgery, and subsequently every 3 months for 3 years.

Treatment costs

All direct health-care-related costs during the first 3 years after inclusion were prospectively documented in the case record forms, including inpatient and outpatient costs, such as the surgical technique used, the amount of blood loss, total operation time, chemotherapy, secondary operations, complications (e.g. ileus, wound infection, biloma, pneumonia), outpatient visits, diagnostic tests. For prices, the actual, true unit costs were used. Unit costs were built up on personnel, material and capacity cost. For standard units of care, like intensive care unit days, in hospital days, out-patient visits etc., standard cost prices as stated in the Dutch manual for costing research by Oosterbrink et al. [28] indexed using the consumer price index to 2007 prices, were used (www.statline.nl). Overhead was defined as 35% of the accumulated direct costs in line with the Dutch guidelines for costing research (Oostenbrink et al. [28]).

Excluding chemotherapy costs seems obvious when dealing with a diagnostic-driven trial, but given the influence on outcomes in disease-free and overall survival, both
evaluations were taken into account. This underlines the fact that considering only the diagnostic channel provides intermediate outcomes with potentially biased results and potentially different policy decisions about adding FDG-PET technology to the benefit package. Whether and under what conditions FDG-PET is an efficient modality for patients adhering to the inclusion criteria of this trial are therefore described in a transparent way presenting both scenarios, the diagnostic channel only and the diagnostic channel plus the clinical pathway up to 3 years. Costs and effectiveness have been discounted at a 3% rate for the duration of the study, that is, 3 years. The perspective of accumulated costs in this study is that of the Dutch national health system.

**Efficiency analysis**

Cost–utility analysis was conducted from a health-care perspective. The efficiency of FDG-PET compared with CWU is expressed as the net monetary benefit (NMB). The NMB is an outcome presented in money terms that subtracts the net cost from the net effect between FDG-PET and CWU. The net costs are measured as the difference in costs between FDG-PET and CWU. The net effect is measured as the difference in QALYs between FDG-PET and CWU multiplied by a monetary value given to a QALY. The decision rule states that when the NMB is positive, FDG-PET is cost-effective compared with CWU. For the Netherlands the range for ‘willingness to pay for a QALY’ is based on a report by the Council for Public Health and Health Care (Raad voor Volksgezondheid en Zorg), an advisory board to the Ministry of health, titled ‘Duurzame en Zinnige zorg, Volksgezondheid en Zorg), an advisory board to the Ministry of health, titled ‘Duurzame en Zinnige zorg, 2006’ and ranges from €10 000–€80 000 depending on the burden of illness [29]. Furthermore, a bootstrap procedure was performed to provide an estimate of the uncertainty surrounding the incremental cost-effectiveness ratio (ICER) [30]. Bootstrapping is a statistical procedure that, for example, 1000 times resamples an actual measured point of additional costs and additional QALY with replacement from the original sample. This method reduces the chance that outliers significantly influence the cost–utility analysis with extra health care consumption in the nominator and QALYs gained in the denominator. The cost-effectiveness plane visualizes the results of the bootstrap procedure with effects on the x-axis and costs on the y-axis. This is presented in a cost-effectiveness plane in which FDG-PET can be placed, where every point represents an estimated ICER. An ICER is the ratio of the change in costs of a therapeutic intervention (compared with the alternative) to the change in effects of the intervention.

**Sensitivity analysis**

To show the influence of a more intensive use, that is, a higher occupancy rate, of the PET, possibly resulting in a lower unit cost for each scan, a sensitivity analysis varying the occupancy rate of the PET scanner was done.

**Statistical analysis**

For the difference between groups in costs and clinical outcomes, Fisher’s exact test (two-sided) was used for categorical data and t-tests for continuous data. Differences in survival were analyzed with the log-rank test. All statistical analyses were done with the SPSS statistical software (version 16.0, SPSS Inc., Chicago, Illinois, USA) and all data analyses in Excel 2007 (Microsoft, Redmond, Washington, USA).

**Results**

**Clinical outcomes**

There were no baseline differences between both diagnostic work-up modalities (Table 1). Additional FDG-PET findings resulted in the cancellation of planned resection of the suspected liver metastases in five patients. Follow-up of these cases showed that FDG-PET correctly predicted benign disease in two patients and unresectable extrahepatic disease in three patients. Therefore, a total of 75 patients in the conventional arm without FDG-PET and 70 patients in the experimental arm with FDG-PET underwent laparotomy (Table 2). During laparotomy, 17 (23%) patients in the conventional arm and seven (9%) patients in the experimental arm either showed significant additional metastatic disease precluding any further curative surgical treatment or showed benign disease, both of which has led to futile laparotomy ($P = 0.043$).

In addition, follow-up showed disease recurrence within 6 months after surgical treatment in 16 and 13 patients in the conventional and experimental groups, respectively. As a result a significantly larger proportion of patients underwent futile laparotomy in the control arm without FDG-PET (45%) than in the experimental arm with

<table>
<thead>
<tr>
<th>Table 1 Demographic data</th>
<th>Control arm ($n=75$)</th>
<th>Experimental arm ($n=70$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years (range)</td>
<td>62.9 (37.9–79.9)</td>
<td>61.8 (32.8–78.1)</td>
</tr>
<tr>
<td>Sex (female:males)</td>
<td>19:56</td>
<td>25:45</td>
</tr>
<tr>
<td>Primary tumour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pN0</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>pN $\geq 1$</td>
<td>41</td>
<td>40</td>
</tr>
<tr>
<td>Disease-free survival</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt;12$ months</td>
<td>29</td>
<td>32</td>
</tr>
<tr>
<td>$\geq 12$ months</td>
<td>46</td>
<td>38</td>
</tr>
<tr>
<td>Number of hepatic tumours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq 1$</td>
<td>41</td>
<td>40</td>
</tr>
<tr>
<td>$&gt;1$</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>Size of greatest hepatic tumours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt;50$ mm</td>
<td>60</td>
<td>54</td>
</tr>
<tr>
<td>$&gt;50$ mm</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>CEA preoperatively</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt;200$ ng/ml</td>
<td>75</td>
<td>70</td>
</tr>
<tr>
<td>$\geq 200$ ng/ml</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

CEA, carcinoembryonic antigen.

No significant statistical differences were found among the groups.

pN0, no nodal metastases in primary tumour, according to the Tumor, Node, Metastasis (TNM) classification.

pN $\geq 1$, one or more nodal metastases in primary tumour, according to the TNM classification.

*A* as preoperatively predicted on computed tomography, at the time of randomization.
FDG-PET (28%; \(P = 0.042\)). The relative risk reduction was 38% [95% confidence interval (CI): 4–60%]. Futile laparotomy was not related to demographic differences or prognostic factors, as reported elsewhere [24].

### Health-state values and quality-adjusted life-years

The calculated health-state values were based on 73 patients in the CWU group (two patients did not return any EQ-5D form) and 75 patients in the FDG-PET group. During the follow-up in the first 3 years the health-state values were not different for the two study groups (Fig. 1). As 3-year overall survival seemed to be almost identical for the two groups (65.8% for CWU and 61.3% for PET group), QALYs accumulated over the 3-year follow-up period were not significantly different. For the CWU group 1.78 QALYs were generated, whereas for the PET group, QALYs of 1.67 were generated (mean difference 0.10, 95% CI: −0.19–0.39).

### Cost analysis

The different costs over several diagnostic and treatment options are presented in Table 3. The cost for surgery were standardized to the cost for a 3-h OR (operation room) session, without radio-frequency ablation (which would raise costs with 1950 euros). Overall, after comparing the diagnostic strategies, the average total costs were lower in the PET group (Table 2). Three years after randomization the mean costs for the CWU and PET groups were € 92 836 (range: €7516–290 308) and €81 776 (range: €6087–€341 012), respectively. When the costs for chemotherapy were not taken into account, the mean costs for the CWU group were €15 874 (range: €5974–€34 143) versus €18 664 (range: €1984–€87 930) for the PET group.

### Economic analysis

The NMB (with 95% CIs) shows a decline (Fig. 2). This decline can be explained by the fact that as the willingness to pay for a QALY increases, more weight is added to the slightly (insignificant) negative QALY effect in the FDG-PET group.

Results of the bootstrap procedure are shown in the cost-effectiveness plane (Fig. 3), in which most outcomes lie beneath the \(x\)-axis, meaning that additional FDG-PET is more likely to save money than to lose money over a period of 3 years (given the indecisive effect). In a second scenario when the bootstrap procedure was performed over the data generated without chemotherapy, it is shown that the inclusion of FDG-PET results in a less favourable and more costly strategy (Fig. 4). Every point in this cost-effectiveness plane represents an estimated ICER derived from the bootstrap procedure expressed as cost for each QALY gained.

### Sensitivity analysis

Sensitivity analysis varying the occupancy rate of the PET scanner (Table 4) showed that 1200 PET scans per scanner per year (with 260 days for clinical use, thus 4.6 PETs per day) would result in a unit cost of €987 per PET. If the PET scanner could be used at a higher occupancy rate these costs would decrease. For instance, with 10 scans a day, the unit costs would become €759. However, this strategy does not result in either a statistically significant advantage in costs or in NMB for the PET group.
Discussion
This study shows that the addition of FDG-PET to CWU in the diagnostic work-up for patients with potentially resectable colorectal liver metastases results in a better selection of patients for surgical intervention, and thus in avoiding futile laparotomies. When considering a follow-up period of 3 years and including all health care costs accumulated in these years, the addition of FDG-PET remained favourable, and resulted in an expected NMB of up to €11 060, depending on the willingness to pay for a QALY gained. However, conclusions about these findings should be made with caution as statistical significance could not be established. This leads to the conclusion that the introduction of FDG-PET might be a potentially cost-efficient diagnostic modality in the diagnostic work-up of these patients.

Our findings are in line with earlier nonrandomized studies on the use of FDG-PET in patients with colorectal liver metastases [13,15,17,18]. Bipat et al. [13] showed that FDG-PET had higher sensitivity for the detection of liver metastases compared with other imaging modalities. Identical observations were made for the detection of extrahepatic disease [18].

In this study, 14 patients (18.7%) in the control arm underwent futile laparotomy because of liver disease.

Table 3  Actual cost of diagnostic, therapy and hospital stay

<table>
<thead>
<tr>
<th></th>
<th>CT scan</th>
<th>MRI</th>
<th>Ultrasound</th>
<th>Chest X-ray</th>
<th>PET scan</th>
<th>Colonoscopy</th>
<th>Outpatient*</th>
<th>Normal care*</th>
<th>Intensive care*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs of diagnostics, hospital days, intensive care treatment, outpatients visits (£)</td>
<td></td>
<td></td>
<td></td>
<td>1.78</td>
<td>663.26</td>
<td>28.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity</td>
<td>67.55</td>
<td>143.41</td>
<td>3.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personnel</td>
<td>56.51</td>
<td>60.10</td>
<td>41.82</td>
<td>16.49</td>
<td>67.98</td>
<td>100.52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overhead</td>
<td>43.42</td>
<td>71.23</td>
<td>15.80</td>
<td>6.40</td>
<td>255.93</td>
<td>45.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>167.48</td>
<td>274.74</td>
<td>60.95</td>
<td>24.67</td>
<td>987.17</td>
<td>174.48</td>
<td>107.28</td>
<td>490.12</td>
<td>1692.91</td>
</tr>
</tbody>
</table>

Costs operation (£)  

<table>
<thead>
<tr>
<th>Operation</th>
<th>855.00</th>
<th>1500.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional costs: RFA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Costs chemotherapy (£)  

<table>
<thead>
<tr>
<th>Scheme</th>
<th>5FU/LV</th>
<th>Oxaliplatin 5FU/LV</th>
<th>Oxaliplatin 5FU/LV</th>
<th>Irinotecan</th>
<th>Oxaliplatin capacetabine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per therapy</td>
<td>1393.12</td>
<td>13 267.52</td>
<td>13 087.48</td>
<td>1820.45</td>
<td>5828.59</td>
</tr>
<tr>
<td>Per week</td>
<td>1393.12</td>
<td>1895.36</td>
<td>1869.64</td>
<td>606.82</td>
<td>1942.86</td>
</tr>
<tr>
<td>Overhead</td>
<td>487.60</td>
<td>663.38</td>
<td>654.37</td>
<td>312.38</td>
<td>680.11</td>
</tr>
<tr>
<td>Total/week</td>
<td>1880.72</td>
<td>2558.74</td>
<td>2524.01</td>
<td>819.20</td>
<td>2622.87</td>
</tr>
</tbody>
</table>

5FU, 5-fluorouracil; CT, computed tomography; LV, leucovorin; MRI, magnetic resonance imaging; OR, operation room; PET, positron emission tomography, RFA, radiofrequency ablation.

*Based on Oostenbrink et al. [28].
being too extensive at laparotomy or the detection of extrahepatic disease (data not shown). The outcome is comparable with a European Organization for Research and Treatment of Cancer study investigating the role of adjuvant perioperative chemotherapy in patients undergoing hepatic resection for colorectal liver metastases. In this study, 16.4% of the randomized patients in the control arm (without chemotherapy) did not undergo planned hepatic resection [31]. This study emphasizes the fact that chemotherapy is a significant cost driver.

This cost-effectiveness analysis was based on a conservative approach meaning that not all beneficial effects were included in the analysis. Additional FDG-PET findings were divided in extrahepatic disease, evidence of benign disease, and additional liver disease. In seven patients, FDG-PET detected additional extrahepatic disease initially missed on CT scan (lung or mediastinal metastases in five patients and extensive abdominal lymph node metastases in two cases). In five patients, FDG-PET did not show uptake in focal liver lesions, indicating benign liver disease; additional diagnostics confirmed this. In 18 patients, FDG-PET showed additional liver findings discordant with CT. In three of these patients, FDG-PET predicted extensive central liver involvement judged resectable on CT. In all three patients resection was judged impossible at laparotomy and further surgical treatment was cancelled. In this last group, FDG-PET correctly predicted futile laparotomies in 10 patients who were operated on, because of discordance compared with CT (the leading diagnostic in our study). Those 10 operations could have increased the nonchemotherapy costs of the PET group.

The results of this study showed that futile laparotomies were avoided; therefore, the main economic benefits of adding FDG-PET to the diagnostic work-up would be logistical, as hospital performance and quality of care improves because of more effective planning of OR capacity and unnecessary operations are avoided. This ‘quality benefit’ is not represented in the outcome of our cost–utility analysis. Overall, this could imply that hospital performance and management improves because unnecessary operations are avoided and the use of OR capacity is optimized.

The results of sensitivity analyses indicate that the costs of PET could benefit from a better use of occupancy and suggest that PET scanning should be concentrated in high-volume centres that are better able to exploit the economies of scale. When the capacity of PET scanners is optimally exploited, this results in a more efficient use of hardware and trained personnel, providing a solid argument for the centralization of this modality.

In contrast to other studies, no improvement in the HRQoL, because of the decrease in futile laparotomies, could be observed in this study. A possible explanation for this is the fact that 77% of the treated patients in the CWU group and the 84% patients in the PET group showed almost identical HRQoL outcomes which could be because of the timing of the start of chemotherapy. Patients for whom laparotomy is avoided because of the addition of FDG-PET are scheduled early for chemotherapy. This could lead to a relatively early decrease in the HRQoL compared with the patients who receive chemotherapy later because of disease recurrence after an initial hepatic resection, as reported earlier [24].

The main limitations of this paper is that the randomized clinical trial was carried out in the dawning of the hybrid PET/CT scanner; therefore, this technology was not taken into account for cost calculation. When data on this hybrid scanning become available, the added value of the application of PET/CT scanners should become more evident. The next challenge will be to develop imaging methods for the better assessment of liver involvement, for example, with hybrid PET/MRI, thus further exploiting innovations in the hybrid scanner technology.

In conclusion, the addition of FDG-PET to the diagnostic work-up in patients with potentially resectable colorectal liver metastases results in better patient selection without leading to additional costs and with the potential for savings.
Acknowledgements
All authors have read and approved this manuscript, and attest that there is no financial or other relationships that could be construed as a conflict of interest.


References