Esophageal cancer is rising in incidence and has a poor prognosis. Positron Emission Tomography (PET) is increasingly being investigated as a tool to more discriminately manage these patients. Several studies have indicated benefits in the use of PET for staging and assessment of treatment response while others have provided contradicting results. There are many possible factors that might contribute to these results, including variability in the manner of PET administration and interpretation, timing, and study design. PET acquired after chemoradiation or chemotherapy may give important prognostic information that can guide additional management decisions. Studies have had substantial variability in the timing and manner of assessing PET for this purpose, and additional study is needed.

Esophageal cancer is rising in incidence with an estimated 16470 new cases in the United States in 2008, and 14280 deaths, (146,726 and 124,728 respectively worldwide in 2006) (1). Unfortunately this disease has a poor prognosis with low long term survival. This may be due partly to late detection of disease with tumors frequently remaining undetected until they are locally advanced or metastatic. Esophageal cancer staging is intended to group patients with similar prognosis for appropriate therapy. The accuracy of staging is contingent on the sensitivity and specificity of the tools available to the physician, as is ongoing management based on response to prior therapy. Positron Emission Tomography (PET) is one such tool that has increased in usage over the last several years. Despite variability in the manner of administration and interpretation between institutions, investigators have sensed great promise with PET. There have been several recent reports investigating its potential impact on patient management for esophageal cancer, as researchers attempt to find out the best way to apply this unique imaging modality.

PET scans reveal metabolically active tissue regardless of whether the activity is from malignancy, inflammation, or other causes. This, along with the limited spatial resolution inherent with this modality, limits the interpretation of PET in oncologic management generally. Regardless of PET’s limitations, it has improved the accuracy of staging and its value in post-therapy evaluation is recognized but not yet fully defined. PET is now typically added to clinical assessment, diagnostic CT, endoscopic gastroduodenoscopy, and endoscopic ultrasonography for staging workup. There have been a number of recent studies suggesting new beneficial uses of the modality, but the findings have been somewhat mixed and are difficult to collectively summarize into a coherent, well-supported guideline.

**PET utility in staging**

Patients with locally advanced disease are often treated with neoadjuvant chemoradiation followed by surgery. Several meta-analyses have shown a benefit in local recurrence, complete resection, and survival with trimodality therapy compared with surgery alone (2-4). However, the addition of neoadjuvant therapy limits initial staging due to the absence of histopathological information. This raises the potential value of additional information that can be used for clinical staging, such as through PET.

Esophageal cancer uses the AJCC TNM staging convention to represent primary, nodal, and metastatic disease respectively. The T stage depends on the invasiveness of the primary tumor and is well-appreciated with endoscopic ultrasound. PET scans may have value in determining the size and location of the primary malignancy, and thereby may be used to assist in radiation treatment planning target delineation, but these do not influence the T stage (5,6). There are other limitations to PET in regard to primary tumor evaluation as well. Although most esophageal malignancies are hypermetabolic and manifest on PET, lesions less than 1 cm may be too small to be detected. Also, the spatial resolution of PET is inadequate to contribute to the T stage by sug-
gesting a degree of invasion with any certainty even when it is positive.

PET may improve the accuracy of the N stage by distinguishing metabolically active lymph nodes from enlarged benign nodes. However, the low resolution of PET imaging makes it difficult to distinguish loco-regional lymph nodes from direct primary tumor extension; and metabolically active nodes may reflect sarcoidosis, granulomatous disease, reactive nodes, or other non-malignant conditions. Using PET for N staging also shares the T stage limitation of failing to identify microscopic disease or gross disease less than 1 cm (7).

The area in which PET has the greatest potential utility in esophageal cancer staging is in the assessment of distant metastases, the M stage. PET/CT may detect metastatic disease at unusual sites that may otherwise have been overlooked, and has thereby been shown to improve staging and prevent inappropriate surgery for patients with metastatic disease (8,9). However, considering the unnecessary investigation of false positive findings after suspected metastases are detected on PET, some have concluded that PET offers limited additional value over other staging modalities and may not have a justifiable role (7).

PET utility in assessing treatment response

Patients with persistent disease after neoadjuvant therapy and prior to surgery have a poorer outcome (10,11). Accurately assessing each patient’s response to therapy then becomes of critical importance in making additional management decisions. A PET scan may be helpful in accurately determining patient response to treatment to inform ongoing management and facilitate choosing appropriate subsequent therapy.

There have been mixed reports on this topic with some finding particular utility in PET for this purpose, and some finding no value. Perhaps the key issue in all of these studies is the relatively small cohorts of patients that have been available for comparison, leading to somewhat erratic formulas for the meaningful use of PET. The way that PET has been used and evaluated has varied so widely that distilling the literature into a unified algorithm with clinical meaning, if such is even possible, is a difficult task.

A reduction in SUVmean or SUVmax between pre- and post-treatment PET scans was a predictor of pathologic response in some series, but the cutoff point varied widely between the studies (e.g. 10% to 80%) and has often been chosen tailored to a retrospective data set rather than prospectively evaluated (9,10,12-15). Although many studies of PET in esophageal cancer designate themselves as prospectively designed and executed, the specific analysis of SUV cutoffs or percentage decreases have nearly uniformly been retrospectively determined.

In other studies, persistent uptake within the primary tumor site on a single post-treatment PET correlated with residual viable tumor and poor survival (11,16-18). However, the specific SUVmax value used in these series as a cutoff varied from 2.5 to 5.5, and unfortunately other recent studies similarly designed have concluded that a single post-therapy PET scan is not adequate in determining response within the primary tumor (19-21).

The findings of a number of key studies investigating PET usage in the management of esophageal cancer patients have been summarized in Table 1. The variability within this group's study design is striking, and this helps explain why the outcomes also have such variability. Only the Lordick study had more than 100 patients; the rest ranged from 10 to 83. The timing of post-therapy PET ranged from one week post initiation to 10 weeks post completion. The SUV values presumed to have the potential of predictive value ranged from 2.5 to 5.5; the percentage changes in SUV from 10% to 80% and have often been chosen tailored to a retrospective data set rather than prospectively evaluated (9,10,12-15). Understandably, the table does not capture the specific clinical questions being addressed, and oversimplifies the individually tailored by pothe sis-driven study designs. However, when considered together, the

---

**Table 1**

<table>
<thead>
<tr>
<th>Study</th>
<th>No. pts</th>
<th>Therapy</th>
<th>Timing</th>
<th>SUV reduction</th>
<th>Sens / Spec</th>
<th>Aden/SCC</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brucher (2001)³⁰</td>
<td>27</td>
<td>CRT</td>
<td>3-4 wks pc</td>
<td>52%</td>
<td>100% / 55%</td>
<td>0/27</td>
<td>Prospective</td>
</tr>
<tr>
<td>Weber (2001)³¹</td>
<td>40</td>
<td>C</td>
<td>2 wks pi</td>
<td>35%</td>
<td>93% / 95%</td>
<td>40/0</td>
<td>Prospective</td>
</tr>
<tr>
<td>Kato (2002)³⁰</td>
<td>10</td>
<td>CRT</td>
<td>2 wks pc</td>
<td>50%</td>
<td>-</td>
<td>0/10</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Flamen (2002)³²</td>
<td>36</td>
<td>CRT</td>
<td>4-6 wks pc</td>
<td>80%</td>
<td>71% / 82%</td>
<td>9/27</td>
<td>Prospective</td>
</tr>
<tr>
<td>Downey (2003)³²</td>
<td>39</td>
<td>CRT</td>
<td>pc</td>
<td>60%</td>
<td>-</td>
<td>26/13</td>
<td>Prospective</td>
</tr>
<tr>
<td>Wieder (2004)³⁰</td>
<td>38</td>
<td>CRT</td>
<td>2 wks pi</td>
<td>30%</td>
<td>93% / 88%</td>
<td>0/38</td>
<td>Prospective</td>
</tr>
<tr>
<td>Song (2005)³³</td>
<td>32</td>
<td>CRT</td>
<td>4 wks pc</td>
<td>Any</td>
<td>16-27% / 95-98%</td>
<td>0/32</td>
<td>Prospective</td>
</tr>
<tr>
<td>Gillham (2006)³⁴</td>
<td>32</td>
<td>CRT</td>
<td>1 wk pi</td>
<td>36%</td>
<td>-</td>
<td>27/5</td>
<td>Prospective</td>
</tr>
<tr>
<td>Ott (2006)³⁵</td>
<td>65</td>
<td>C</td>
<td>2 wks pi</td>
<td>35%</td>
<td>80% / 78%</td>
<td>65/0</td>
<td>Prospective</td>
</tr>
<tr>
<td>Westerterp (2006)³⁶</td>
<td>26</td>
<td>CRT+H</td>
<td>2 wks pi</td>
<td>31%</td>
<td>75% / 75%</td>
<td>20/6</td>
<td>Prospective</td>
</tr>
<tr>
<td>Lordick (2007)³⁷</td>
<td>119</td>
<td>C</td>
<td>2 wks pi</td>
<td>35%</td>
<td>-</td>
<td>110/0</td>
<td>Prospective</td>
</tr>
<tr>
<td>Mamede (2007)³⁸</td>
<td>25</td>
<td>CRT</td>
<td>1.5 wks pc</td>
<td>32%</td>
<td>75% / 63%</td>
<td>22/3</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Port (2007)³⁹</td>
<td>62</td>
<td>CRT</td>
<td>2-3 wks pc</td>
<td>50%</td>
<td>78% / 53%</td>
<td>51/11</td>
<td>Retrospective</td>
</tr>
</tbody>
</table>
study designs suggest an overall erratic approach to the issue, and whether or not this is an accurate assessment of the planning and execution of these studies, the fact that no clear clinical conclusions can be drawn is hardly surprising. While some have concluded that PET is not adequate in determining response within the primary tumor after neoadjuvant therapy, perhaps a more precise conclusion is that no algorithm for the usage of PET has been found to have the merit to be considered standard and warrant additional rigorous investigation.

There are several issues that may contribute to the disparate findings among these studies. Some studies examined only adenocarcinoma patient response while others were exclusively squamous cell carcinoma. Most were mixed. This may explain the relatively large difference in SUVmax cutoff values used to assess treatment response. Additionally, negative findings often remain unpublished and could be under-represented in the published literature. Retrospective studies are also widely understood to suffer from bias, and that seems particularly relevant in a group of studies with similar conclusions but widely disparate objective data.

Another possible reason for the range of findings in studies that address PET as a tool to assess clinical response is the changing technical format of PET administration. Earlier studies routinely obtained PET without CT using a separate transmission scan for attenuation correction. PET/CT uses CT data to perform attenuation correction and the difference in time acquisition results in mismatches. This may be corrected using respiration-averaged CT, but because independent PET was used for many of the earlier studies while PET/CT has been used most frequently for recent studies may explain some of the disparity in findings. There are also disparities between treatment centers in FDG dose and attenuation correction procedures (33). These differences in administration may also explain why there is such disparity in the average SUV values obtained within a population at different centers. Song noted that their cohort’s mean SUV prior to treatment was 5.6 ± 3.6, which is quite different from the 9.3 ± 2.8 reported by Wieder (23).

A potential limitation of post therapy PET is the esophagitis and ulceration that is induced by chemoradiation during treatment and which manifests as increased uptake on PET. Reactive uptake in non-malignant tissues increases three or more weeks after treatment, but tumor tissue uptake may not yet have diminished within the first week or two after treatment. The timing of PET is important to minimize the potential masking of high uptake in actual persistent disease (20,32). There seems to be little agreement in the literature on the precise time-point at which post-therapeutic PET is most useful.

PET has also been used as an assessment of treatment response after brief chemotherapy and prior to the full course of chemoradiation. This holds advantages for the group of patients who have a poor response to chemoradiation because surgical outcome is poorer after trimodality therapy than it would have been if surgery had not been delayed for neoadjuvant therapy. In a 2001 study, Weber et al. showed that PET/CT after two cycles of chemotherapy predicts pathologic response to neoadjuvant therapy and long-term outcome with a sensitivity of 93% and a specificity of 95% (14). In a 2004 study, Wieder et al. supported this idea that changes in tumor metabolic activity after 2 weeks of neoadjuvant therapy significantly correlate with tumor response and patient survival. Then in a 2008, Lordick et al. further supported the idea with a report of the MUNICON trial in which the utility of PET was prospectively assessed when used as an earlier assessment of neoadjuvant treatment response. Patients were divided into responder and non-responder groups after administering two weeks of preliminary chemotherapy. Non-responders were allowed to proceed directly
to surgery without additional neoadjuvant therapy while responders received the full course of chemoradiation. The results suggested the feasibility of a PET-guided treatment algorithm for esophageal cancer (29).

An ongoing phase II study by the CALGB 80302 is partly seeking to address the potential utility of PET after induction therapy. This study is designed to determine the pathologic complete response rate in patients with surgically resectable esophageal cancer treated pre-operatively with induction chemotherapy with weekly cisplatin and irinotecan followed by concurrent cisplatin/irinotecan and radiation therapy. Endpoints to be analyzed in relation to PET include histopathological response, clinical response, overall survival, and disease free survival. This will include a larger cohort of prospective patients and should be helpful in clarifying some of the current ambiguity.

Discussion

PET is useful in esophageal cancer for staging and evaluation of treatment response. However, this is only true when PET is carefully interpreted with awareness of its limitations. An awareness of the scientific basis for PET will allow physicians to interpret the results within the patient’s overall clinical history, including timing of PET acquisition prior to biopsies and other procedures that confound results. Both SUV magnitude and variability seem to be site-specific, suggesting that criteria for PET usage may best be defined by individual sites. Specific prognostic information and appropriate treatment management in response to PET evaluation will become better defined as additional studies, particularly prospective trials, are published in the future.

References


