Cost-effectiveness of CT and PET-CT for determining the need for adjuvant neck dissection in locally advanced head and neck cancer

D. J. Sher1,2*, R. B. Tishler1, D. Annino3 & R. S. Punglia1,2

1Department of Radiation Oncology, Dana-Farber Cancer Institute and Brigham and Women’s Hospital, Boston; 2Center for Outcomes and Policy Research, Dana-Farber Cancer Institute, Boston, MA, USA and 3Division of Otolaryngology, Brigham and Women’s Hospital, Boston, MA, USA

Received 20 March 2009; accepted 14 July 2009

Background: Patients with node-positive head and neck squamous cell carcinomas (HNC) have a significant risk of residual disease (RD) in the neck after treatment, despite optimal chemoradiotherapy (CRT). Adjuvant neck dissection (ND) after CRT has been considered standard treatment, but its morbidity has led investigators to consider using post-CRT imaging to determine the need for surgery. We analyzed the cost-effectiveness of computed tomography (CT) and positron emission tomography–computed tomography (PET-CT) as predictors of the need for ND compared with ND for all patients.

Materials and methods: We developed a Markov model to describe health states in the 5 years after CRT for HNC in a 50-year-old man. We compared three strategies: dissect all patients, dissect patients with RD on CT, and dissect patients with RD on PET-CT. Probabilistic sensitivity analyses were carried out to model uncertainty in PET-CT performance, up-front and salvage dissection costs, and patient utilities.

Results: ND only for patients with RD on PET-CT was the dominant strategy over a wide range of realistic and exaggerated assumptions. Probabilistic sensitivity analyses confirmed that the PET-CT strategy was almost certainly cost-effective at a societal willingness-to-pay threshold of $500 000/quality-adjusted life year.

Conclusion: Adjuvant ND reserved for patients with RD on PET-CT is the dominant and cost-effective strategy.

Key words: cost-effectiveness analysis, neck dissection, radiotherapy, PET-CT

introduction

Approximately 42 000 Americans are diagnosed with head and neck cancer each year, of whom roughly 50% will have involved lymph nodes [1]. Despite significant improvements in chemoradiotherapy (CRT), a significant proportion of patients have residual disease (RD) in their neck at the conclusion of therapy, and therefore, adjuvant neck dissection (ND) after CRT has been the standard of care. Modified radical neck dissection (MRND) entails removal of all ipsilateral cervical lymph nodes, from the inferior border of the mandible to the clavicle, with preservation of at least the internal jugular vein, sternocleidomastoid, or spinal accessory nerve. In addition to the cost of the procedure itself, post-CRT ND can lead to increased pain [2] and functional disability [3, 4]. The goals of the procedure were to eliminate the risk of subsequent nodal recurrences (NRs), which are difficult to salvage, as well as to provide prognostic information.

Since routine adjuvant ND is clearly associated with additional morbidity, multiple investigators have attempted to define criteria for sparing patients this additional surgery, including response based on clinical exam, computed tomography (CT), and positron emission tomography (PET). Despite the promise of these approaches, nonreproducible test sensitivities and specificities have limited broad acceptance of these strategies, in part because the optimal timing of these tests has not been determined [5]. As we look forward to establishing guidelines for post-CRT neck management, it is important to note that no group has yet considered the cost-effectiveness of methods for determining when to proceed with adjuvant ND.

Although CT and positron emission tomography–computed tomography (PET-CT) are associated with nontrivial costs, the benefits of avoiding the cost and morbidity of an ND may render them cost-effective strategies. In order to test this hypothesis, we developed a model to analyze the cost-effectiveness of using post-CRT CT and PET-CT to determine the necessity for adjuvant ND. Our results were tested against a wide range of reasonable disease and treatment assumptions.
materials and methods

decision model

We designed a Markov model to simulate the clinical history of a 50-year-old man with node-positive stage IVA (i.e. T1-3 N2 M0) squamous cell carcinoma of the oropharynx (Figure 1). Markov simulation allows hypothetical cohorts of patients to transition between different health states over fixed increments of time [6].

In the model, patients began in the well state [no evidence of disease (NED)], having received definitive CRT. These initial states were titled NED-Dissection or NED-No Dissection depending on whether the patients underwent post-CRT ND (Figure 1). Patients then remained in the NED states or proceeded to the disease states, including local recurrence (LR), NR, and distant metastasis. LRs were defined as disease reappearing in the original primary location. In reality, many patients experience simultaneous LR and NR, but these were separated in the model to better describe the benefits of ND. Possible terminal states into which a patient could transition were death from head and neck cancer and death from other causes.

Three strategies were evaluated. One strategy entailed carrying out NDs on all patients, without evaluating response to CRT. The second strategy involved carrying out NDs only for patients with RD on CT, and the third option was only carrying out NDs on patients with RD on PET-CT. A 5-year time horizon was used in the model since the vast majority of clinical trials present 5-year survival data and most locoregional and distant recurrences occur within the first 3 years [7]. The model was created and analyzed using Data TreeAge Pro (TreeAge Software, Inc., Williamstown, MA).

model assumptions and data

Assumptions and data sources for the decision model are shown in Table 1. All probabilities were extracted from the published literature and calibrated for consistency with Radiation Therapy Oncology Group (RTOG) 99-14, a phase II trial of concurrent chemotherapy and concomitant boost radiotherapy that has formed the basis of the current RTOG randomized trial in head and neck oncology [7, 8]. Randomized trials have all shown the superiority of CRT compared with radiotherapy alone, but the exact survival rates are highly variable; therefore, sensitivity analyses tested a wide range of disease and treatment assumptions.

Costs accrued in each health state were taken from the published literature and publicly available Medicare payment schedules. Costs are expressed in 2006 US$. Unfortunately, there are no studies evaluating the costs of palliative treatments in head and neck cancer. For chemotherapy costs for patients with metastases, we used recently published Surveillance, Epidemiology and End Results–Medicare data estimating the costs of palliative chemotherapy in lung cancer [9] since the drugs used are comparable (platinum- and taxane-based compounds). For patients who ultimately required a salvage ND, the costs associated with a radical neck dissection (RND) were used [10]. The best treatment for patients who have a local failure is controversial. It appears that patients who can undergo a surgical resection have the best outcomes, while the addition of CRT to surgery has a more questionable benefit [11]. We therefore assumed that all patients who failed locally would undergo a surgical procedure without adjuvant therapy; this cost was extrapolated from Medicare diagnosis-related group 129 (major head and neck procedure with comorbidities) [12]. All patients who died from other causes.

Table 1. Probabilities, utilities, and costs used in this study

<table>
<thead>
<tr>
<th>Event</th>
<th>Baseline value</th>
<th>Range studied</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NED ⇒ nodal recurrence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD, dissection</td>
<td>20% (5 years)</td>
<td>10–35%</td>
<td>[14–16]</td>
</tr>
<tr>
<td>RD, no dissection</td>
<td>95% (2 years)</td>
<td>80–99%</td>
<td></td>
</tr>
<tr>
<td>CR, dissection</td>
<td>3% (5 years)</td>
<td>0–10%</td>
<td></td>
</tr>
<tr>
<td>CR, no dissection</td>
<td>6% (5 years)</td>
<td>2–15%</td>
<td></td>
</tr>
<tr>
<td>NED ⇒ local recurrence</td>
<td>30% (5 years)</td>
<td>15–45%</td>
<td>[7]</td>
</tr>
<tr>
<td>NED ⇒ metastasis</td>
<td>12% (5 years)</td>
<td>5–20%</td>
<td></td>
</tr>
<tr>
<td>Nodal recurrence ⇒ death</td>
<td>75% (3 years)</td>
<td>50–90%</td>
<td>[17, 18]</td>
</tr>
<tr>
<td>Local recurrence ⇒ death</td>
<td>85% (2 years)</td>
<td>70–95%</td>
<td>[11]</td>
</tr>
<tr>
<td>Distant metastasis ⇒ death</td>
<td>90% (2 years)</td>
<td>85–100%</td>
<td>[19]</td>
</tr>
<tr>
<td>Imaging performance (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT sensitivity</td>
<td>72</td>
<td>70–90</td>
<td>[20]</td>
</tr>
<tr>
<td>CT specificity</td>
<td>79</td>
<td>60–80</td>
<td></td>
</tr>
<tr>
<td>PET-CT sensitivity</td>
<td>94</td>
<td>80–99</td>
<td>[21]</td>
</tr>
<tr>
<td>PET-CT specificity</td>
<td>82</td>
<td>70–90</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NED-No Dissection</td>
<td>0.913</td>
<td>0.8–0.95</td>
<td>[10]</td>
</tr>
<tr>
<td>NED-Dissection</td>
<td>0.875</td>
<td>0.7–0.9</td>
<td></td>
</tr>
<tr>
<td>Nodal recurrence</td>
<td>0.675</td>
<td>0.6–0.8</td>
<td></td>
</tr>
<tr>
<td>Local recurrence</td>
<td>0.270</td>
<td>0.2–0.6</td>
<td>[22]</td>
</tr>
<tr>
<td>Distant metastasis</td>
<td>0.520</td>
<td>0.4–0.8</td>
<td>[23]</td>
</tr>
<tr>
<td>Costs ($)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate nodal dissection</td>
<td>7731.00</td>
<td>6000–10 000</td>
<td>[10]</td>
</tr>
<tr>
<td>Salvage nodal dissection</td>
<td>9202.05</td>
<td>7000–30 000</td>
<td></td>
</tr>
<tr>
<td>Salvage surgery for local recurrence</td>
<td>9599.74</td>
<td>8000–12 000</td>
<td>[12]</td>
</tr>
<tr>
<td>Chemotherapy for metastasis</td>
<td>35 629.42</td>
<td>20 000–50 000</td>
<td>[9]</td>
</tr>
<tr>
<td>Hospice care</td>
<td>21 930.79</td>
<td>5000–30 000</td>
<td>[13]</td>
</tr>
<tr>
<td>CT</td>
<td>367.69</td>
<td>300–600</td>
<td>[24]</td>
</tr>
<tr>
<td>PET-CT</td>
<td>1359.16</td>
<td>1000–1800</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Markov model. The utility for the NED health states varies according to whether the patient had a dissection, and the utility for the salvage states varies according to whether the patient had a nodal or local recurrence. NED, no evidence of disease; HNC, head and neck squamous cell carcinoma.
incurred a hospice cost in the last 3 months of life, as determined by Emanuel et al. [13]. All costs are expressed in 2006 US$

Only one paper has assessed utilities for various health states in head and neck cancer; these were elicited from physicians for the health states after MRND, RND, radiation alone, and radiation plus those two health states [10]. We used the utility after RND, 0.675 (on a scale of 0–1, 0 = being dead and 1 = optimal health) as the utility for the salvage ND health state. For unclear reasons, the elicited utility in this study after radiotherapy alone (0.875) was lower than MRND plus radiotherapy (0.913), which is a patently inconsistent preference. Thus, for the base case, we assumed that this utility difference represented the decrement for two progressively worse health states, taking the utility for NED-No Dissection at 0.913 and the utility for NED-Dissection at 0.875. In sensitivity analyses, this difference was tested from 0 (i.e. the dissection did not impact the utility) to 0.3 (i.e. the NED-Dissection state was <0.6, a large decrement). The utility for the metastatic state, 0.52, was taken from a preference elicitation from metastatic esophageal cancer patients, who would presumably have reasonably similar health limitations [23]. Finally, there are no data on utilities for patients who have undergone salvage surgery for an LR. We therefore used a utility value taken from a preference study in which participants were asked to evaluate a health state with the following characteristics: you cannot eat any solid foods, eat a lot less, frequent problems carrying out daily activities, and two or more of the following symptoms (pain, dyspnea, vomiting and regurgitation, weak/sore muscles, loss of taste, and halitosis) [22]. The average utility for this health state was 0.27, which may seem low, but since patients will have often been left with permanent tracheostomy and gastrostomy tubes, it is not an unreasonable estimate.

discounting
Both costs and QALYs were discounted at an annual rate of 3%.

results
model validity
This model predicted an expected 5-year overall survival of 59% for this cohort of patients. These results are similar to the recently reported 4-year overall survival of 53% for patients enrolled on RTOG 99-14, a phase II trial of accelerated fractionation with concurrent chemotherapy [7]. Furthermore, this model predicted a 5-year disease-free survival of 49%, which is quite similar to the 4-year disease-free survival of
48.5%. The predicted 5-year locoregional-only and distant-only recurrence rates of 31% and 8% are similar to this trial’s respective 4-year probabilities of 36% and 11%. These outcome similarities are further evidence that this model used reasonable assumptions and parameters to emulate the natural disease process.

**One-way sensitivity analyses**

Sensitivity analyses allow the modeler to vary model parameters and test whether adjusting these parameters changes the final conclusions. Results of a series of sensitivity analyses are presented in Table 2. With a few exceptions, the PET-CT strategy dominated the other two strategies. This finding means that over a wide variety of realistic assumptions, carrying out PET-CT and then nodal dissection only for those patients with RD is both more effective and less costly. In the few analyses in which PET-CT did not dominate the other two strategies, either PET-CT was more effective and more costly, but the ICER to use PET-CT was trivially small, or PET-CT was less effective and less costly, but the ICER to use the more effective strategy was prohibitively very large. These results strongly support the conclusion that PET-CT is the most cost-effective approach in this clinical scenario. In contrast, although CT was less expensive than PET-CT, its lower sensitivity and specificity led to more unnecessary dissections and fewer necessary dissections, thus increasing total cost and lowering efficacy in comparison to PET-CT.

**Probabilistic sensitivity analyses**

Probabilistic sensitivity analyses study the likelihood that the PET-CT strategy is cost-effective over a wide range of assumptions and societal WTP thresholds. This type of analysis provides a robust estimate of the probability of cost-effectiveness while varying a number of assumptions thousands of times.

Three probabilistic sensitivity analyses are shown in Figure 2. Despite widely varying the PET-CT sensitivity and specificity from ~88% to 98% and 76% to 86%, respectively, PET-CT was always cost-effective, even when the societal WTP was $500 000/QALY, well above any established threshold (Figure 2a). Even when the costs of up-front dissection were studied as low as $4000 and salvage dissections to as high as $30 000, PET-CT was still cost-effective for the vast majority of the iterations (Figure 2b). Finally, even when the utilities for the NED-Dissection and NED-No Dissection health states were identical—which is extremely unlikely to be the case—and the utility for the salvage dissection health state went down to 0.6, PET-CT was always cost-effective (Figure 2c). These results further confirmed the cost-effectiveness of the PET-CT strategy over a wide range of realistic disease, treatment, and quality-of-life assumptions.

**Figure 2.** Acceptability curves. These curves represent the proportion of trials drawn on a hypothetical distribution of varied parameters that result in an incremental cost-effectiveness ratio for posttreatment positron emission tomography–computed tomography (PET-CT) that is less than the societal willingness to pay. Panel (A): Varied parameters are the PET-CT sensitivity and specificity. Panel (B): Varied parameters are costs of immediate and salvage nodal dissections. Dashed line indicates willingness to pay of $50 000/QALY. Panel (C): Varied parameters are the utility for NED-Dissection and salvage dissection health states.
discussion

We have shown that PET-CT followed by ND for RD was not only cost-effective but that it was less expensive and associated with more quality-adjusted life years than either dissecting all patients or using CT to define RD. A strategy that is more effective and less costly than competing strategies is called dominant and strongly supports its implementation. This novel result was consistent across a broad range of reasonable assumptions and robust throughout several probabilistic sensitivity analyses.

Despite the widespread use of PET and PET-CT in oncology [25], there are relatively few studies that have investigated its cost-effectiveness. Previous studies have generally found that functional imaging is cost-effective [10, 26–28]. In part, these results are a function of reduced morbidity from invasive tests or therapies that can be avoided. Adjuvant ND is an excellent example of this benefit as the increased up-front imaging cost and small risk of a false-negative result are vastly outweighed by the cost and quality-of-life benefits of avoiding surgery.

Our study has several strengths that enhance the validity of our findings. First, the underlying Markov model was calibrated for and validated with RTOG 99-14, which studied the now-standard arm of the current phase III RTOG trial, RTOG 05-22. Extrapolating consistent local, nodal, and distant control data from randomized trials and large phase II data can be challenging since these values are highly variable. However, we widely varied the assumptions for all three end points, and the results were identical. In addition, although utilities are patient dependent, probabilistic sensitivity analysis showed that PET-CT is almost certainly cost-effective over 10 000 different combinations of relevant utility values.

Of course, any model-based investigation has inherent limitations. We based our assumptions on published data, but these data are not derived from a prospective cost-effectiveness trial; in that sense, our results may be considered retrospective. Although we can test the validity of these assumptions through a variety of sensitivity analyses, ultimately, the real disease process may not behave in the predicted manner. For instance, utilities elicited in a clinical trial from healthy individuals may not reflect the preferences of patients actually afflicted with the disease, nor are those preferences necessarily consistent between patients. That said, we tested each utility estimate in deterministic sensitivity analyses and key variables in probabilistic sensitivity analyses, and the results were consistent over an extremely wide range of possible values.

Presumably, the morbidity from nodal dissections will improve as surgical techniques evolve, but it is unclear whether this will augment or limit the utility for PET-CT. On one hand, the short- and long-term complications from nodal dissections may become progressively better over time, thus reducing the harm of more routine adjuvant surgery. On the other hand, the emergence of selective ND as a viable alternative to MRND may allow the surgeon to focus on higher risk nodal basins, highlighting the importance (and cost-effectiveness) of sensitive preoperative imaging [29]. In either case, our results strongly favored PET-CT even when the utility for the post-dissection health state was ‘identical’ to the no-dissection health state.

In summary, we have shown that after CRT for N2–N3 head and neck squamous cell cancer, determining the necessity for adjuvant ND using PET-CT is the dominant, cost-effective clinical strategy, resulting in superior efficacy with reduced cost. Although our conclusion was based on a decision model, the result is robust over a wide range of reasonable disease, imaging, and patient assumptions and is therefore applicable to clinical practice today. In this current era of cost-containment, clinicians should feel comfortable using post-CRT PET-CT as the defining test to determine necessity for ND.

disclosure

This work is completely original and has not been presented before.

references


