PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer

Citation for published version:

Digital Object Identifier (DOI):
10.1056/NEJMoa1514493

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
The New England Journal of Medicine

Publisher Rights Statement:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer


ABSTRACT

BACKGROUND
The role of image-guided surveillance as compared with planned neck dissection in the treatment of patients with squamous-cell carcinoma of the head and neck who have advanced nodal disease (stage N2 or N3) and who have received chemoradiotherapy for primary treatment is a matter of debate.

METHODS
In this prospective, randomized, controlled trial, we assessed the noninferiority of positron-emission tomography–computed tomography (PET-CT)–guided surveillance (performed 12 weeks after the end of chemoradiotherapy, with neck dissection performed only if PET-CT showed an incomplete or equivocal response) to planned neck dissection in patients with stage N2 or N3 disease. The primary end point was overall survival.

RESULTS
From 2007 through 2012, we recruited 564 patients (282 patients in the planned-surgery group and 282 patients in the surveillance group) from 37 centers in the United Kingdom. Among these patients, 17% had nodal stage N2a disease and 61% had stage N2b disease. A total of 84% of the patients had oropharyngeal cancer, and 75% had tumor specimens that stained positive for the p16 protein, an indicator that human papillomavirus had a role in the causation of the cancer. The median follow-up was 36 months. PET-CT–guided surveillance resulted in fewer neck dissections than did planned dissection surgery (54 vs. 221); rates of surgical complications were similar in the two groups (42% and 38%, respectively). The 2-year overall survival rate was 84.9% (95% confidence interval [CI], 80.7 to 89.1) in the surveillance group and 81.5% (95% CI, 76.9 to 86.3) in the planned-surgery group. The hazard ratio for death slightly favored PET-CT–guided surveillance and indicated noninferiority (upper boundary of the 95% CI for the hazard ratio, <1.50; P = 0.004). There was no significant difference between the groups with respect to p16 expression. Quality of life was similar in the two groups. PET-CT–guided surveillance, as compared with neck dissection, resulted in savings of £1,492 (approximately $2,190 in U.S. dollars) per person over the duration of the trial.

CONCLUSIONS
Survival was similar among patients who underwent PET-CT–guided surveillance and those who underwent planned neck dissection, but surveillance resulted in considerably fewer operations and it was more cost-effective. (Funded by the National Institute for Health Research Health Technology Assessment Programme and Cancer Research UK; PET-NECK Current Controlled Trials number, ISRCTN13735240.)
Chemoradiotherapy has become a mainstay of primary treatment in patients with squamous-cell carcinoma of the head and neck. However, there are wide variations in the management of advanced nodal disease (stage N2 or N3) in these patients because of the lack of prospective, randomized, controlled trials.  

Retrospective studies showed persistent disease on histopathological examination of nodes in up to 40% of patients who underwent neck dissection after chemoradiotherapy and some evidence of a significant survival advantage associated with planned neck dissection. However, owing to improvements in cross-sectional imaging, consistently low rates of recurrence (<10%) have been reported among the 30 to 45% of patients who have been found to have a complete response on imaging after chemoradiotherapy. Thus, the adoption of image-guided, response-based approaches has increased, albeit without level I evidence.

Unequivocal data are lacking, and a significant percentage of clinicians (35 to 48%) in some countries still perform planned neck dissection, either before or after chemoradiotherapy, in these patients. This procedure is associated with an attendant risk of clinically significant complications.

Combined morphologic and functional imaging with the use of combined 18F-fluorodeoxyglucose (FDG) positron-emission tomography and computed tomography (PET-CT) can identify both structural and metabolic abnormalities in tumors. Meta-analyses of mainly small, single-center PET-CT studies involving patients with squamous-cell carcinoma of the head and neck who have received chemoradiotherapy have shown high negative predictive values of 94.5 to 96.0%. Stratification of these patients for neck dissection with the use of PET-CT after chemoradiotherapy may therefore result in fewer neck dissections and a reduced incidence of complications. However, data from prospective, randomized, multicenter trials to support routine adoption of this approach are lacking. We therefore performed a prospective, randomized, controlled trial to compare the clinical usefulness and health economic outcomes of planned neck dissection versus PET-CT-guided surveillance in patients with nodal stage N2 or N3, metastasis stage M0 disease.

**Methods**

**Trial Conduct**

The first and last author and the trial management group designed the study. The Warwick Clinical Trials Unit gathered the data. Tissue collection and staining for the p16 protein were performed at the University of Birmingham. The authors vouch for the accuracy and completeness of the data and analysis and for adherence to the study protocol, which is provided with the full text of this article at NEJM.org. No one who is not an author contributed to the writing of the manuscript. There was no commercial support for the study.

**Patients**

Eligible patients were at least 18 years of age and had a histologically confirmed diagnosis of squamous-cell carcinoma of the oropharynx, hypopharynx, larynx, oral cavity, or an unknown primary site in the head or neck, with clinical and radiologic (CT or magnetic resonance imaging [MRI]) stage N2 or N3 nodal metastases. Patients had to be suitable candidates for chemoradiotherapy with curative intent and could not have contraindications to neck dissection. The study was approved by the Oxfordshire Multi-centre Research Ethics Committee. All patients provided written informed consent.

**Trial Design**

In this unblinded, multicenter, randomized, controlled, noninferiority trial, eligible patients were randomly assigned to undergo either a planned neck dissection (planned-surgery control group) or PET-CT 12 weeks after completion of chemoradiotherapy (surveillance group). Before randomization, each participating center had to specify on a per-patient basis whether planned neck dissection would be performed within 4 weeks or within 4 to 8 weeks after completion of chemoradiotherapy. In addition, before randomization, clinicians selected chemoradiotherapy regimens from a list of the approved study regimens (Table S1 in the Supplementary Appendix, available at NEJM.org).

Patients underwent central randomization in a 1:1 ratio. Trial-group assignments were balanced with the use of a minimization algorithm according to center, timing of neck dissection (before or after chemoradiotherapy), chemothera-
apy schedule, disease site, tumor stage (T1 or T2 vs. T3 or T4), and nodal stage (N2a or N2b vs. N2c or N3).

ASSESSMENTS AND OUTCOMES

Imaging assessments of the patients’ response to therapy were performed 12 weeks after the last radiotherapy fraction was delivered. Patients assigned to the planned-surgery group were evaluated by means of CT or MRI; those assigned to the surveillance group were evaluated by means of PET-CT.

All PET-CT findings were interpreted locally by PET-CT specialty radiologists and nuclear-medicine physicians. PET-CT scans were assessed qualitatively. Results of PET-CT that showed intense FDG uptake at 12 weeks after chemoradiotherapy, with or without enlarged lymph nodes in the neck, were classified as incomplete nodal responses. Mild or no FDG uptake in enlarged nodes or mild FDG uptake in normal-sized nodes was considered to be an equivocal response. All other PET-CT scans were considered to show complete responses.

Patients who had an incomplete or equivocal response in lymph nodes in the neck and who had a complete response in the primary site underwent neck dissection within 4 weeks after PET-CT. Both modified radical and selective neck dissections were permitted in both groups, with the decision made by the treating surgeon.

Follow-up consisted of clinical examination and imaging for at least 24 months after randomization. Suspected recurrences were assessed by means of biopsy and pathological inspection. Complications of surgery were recorded for a period of 30 days after the operation. Serious adverse events were reported up to 3 months after the last treatment (either chemoradiotherapy or surgery).

Patients completed the European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire—Core 30 general (EORTC QLQ-C30, version 3) and head and neck–specific (EORTC QLQ-H&N35) questionnaires, the M.D. Anderson Dysphagia Inventory, and the EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D) at baseline before randomization, 2 weeks after completion of chemoradiotherapy, and then at 6, 12, and 24 months after randomization. Health economics questionnaires and resource utilization data were collected in a subgroup of patients and centers at the same time points. EQ-5D questionnaires were used to derive patient quality-adjusted life-years (QALYs), which were combined with cost data to assess cost-effectiveness.

LABORATORY STUDIES AND QUALITY ASSURANCE

All treating hospitals were required to be approved as head and neck treatment centers by the U.K. Department of Health. Formalin-fixed, paraffin-embedded tumor samples were tested centrally for p16 expression by means of immunohistochemical analysis with the use of proprietary reagents (CINtec Histology kit, Roche). Testing to detect p16 expression was recorded as positive if more than 70% of the malignant cells showed strong diffuse nuclear and cytoplasmic staining. All PET-CT scans, 10% of other radiologic investigations, and 10% of histologic specimens were reviewed centrally by experienced specialists (details are provided in the Supplementary Appendix).

STATISTICAL ANALYSIS

The power calculations were based on an assumption of a 2-year overall survival rate of 75%. We calculated that with a sample size of 560 patients, the study would have 90% power to show the noninferiority of PET-CT surveillance to planned neck dissection, at a 5% one-sided significance level. Noninferiority was defined as an overall survival rate that was no more than 10 percentage points below the estimated 75% 2-year overall survival rate among patients in the planned-surgery group (i.e., hazard ratio for death, <1.50); this calculation allowed for a 3% loss to follow-up.

The analysis was performed on an intention-to-treat basis. Statistical analyses were performed with the use of SAS software, version 9.3. A prespecified early stopping guideline was applied, with two interim analyses of overall survival (details are provided in the Supplementary Appendix).

The clinical primary end point of overall survival was measured from randomization to the date of death from any cause. Data on survival among patients who were lost to follow-up or who had incomplete follow-up were censored at the date when they were last known to be alive.
Kaplan–Meier survival curves were plotted, and a forest plot was generated to examine the effect across risk factors. To test noninferiority, the hazard ratio was estimated with the use of a Cox proportional-hazards model stratified according to intended timing of planned neck dissection (before or after chemoradiotherapy), with the trial group as the only covariate. Noninferiority was considered to have been demonstrated if the 95th percentile of the estimated hazard ratio was less than the inferiority limit (hazard ratio, 1.50).

Time to recurrence was measured from the date of completion of chemoradiotherapy. Deaths were classified as being due to head and neck cancer or to other causes. The standard scoring methods for the quality-of-life questionnaires were applied. All scores were transformed to scales in which higher scores indicated better health. On the EORTC questionnaires and the M.D. Anderson Dysphagia Inventory, a difference of 10 points in the score was considered to be clinically relevant.

**Results**

**Patient Characteristics and Treatment**

A total of 564 patients (282 in each trial group) were recruited from October 2007 through August 2012 from 37 head and neck treatment centers (43 hospitals) in the United Kingdom. Two patients in the surveillance group were subsequently found to be ineligible immediately after randomization; they were included in the intention-to-treat analysis.

There were no notable imbalances in baseline characteristics between the two trial groups (Table 1, and Table S2 in the Supplementary Appendix). The mean age was 58 years. The majority of patients were male (82%) and had oropharyngeal cancer (84%). A total of 79% of the patients had N2a or N2b nodal disease, and 74% were either current or past smokers.

Tissue samples obtained from 446 patients (79%) were tested for p16 expression; all characteristics (i.e., age, tumor site, tumor stage, smoking history, performance status, and neck surgery before or after chemoradiotherapy) were well matched between the tested and nontested cohorts. Of the patients tested, 335 (75%) had tumor specimens that stained positive for the p16 protein, an indicator that human papillomavirus had a role in the causation of the cancer, with a nonsignificant preponderance in the planned-surgery group as compared with the surveillance group (78% vs. 73%).

Nonsurgical treatments (chemoradiotherapy and radiotherapy) were very well balanced between the groups (Tables S2 and S3 in the Supplementary Appendix). A total of 536 patients (95%) received concomitant platinum chemotherapy.

**Planned-Surgery Group**

There was little crossover from the planned-surgery group to the surveillance group; only 8 patients in the planned-surgery group (3%) underwent PET-CT after chemoradiotherapy. Of the 282 patients in the planned-surgery group, 61 (22%) did not undergo neck dissection; 8 of the 77 (10%) who had been scheduled to undergo neck dissection before chemoradiotherapy and 53 of the 205 (26%) who had been scheduled to undergo neck dissection after chemoradiotherapy. Of the patients who did not undergo planned neck dissection after chemoradiother-

---

**Table 1. Baseline Characteristics of the Patients.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Surveillance Group (N = 282)</th>
<th>Planned-Surgery Group (N = 282)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age – yr</td>
<td>57.6±7.5</td>
<td>58.2±8.1</td>
</tr>
<tr>
<td>Male sex — no. (%)</td>
<td>223 (79.1)</td>
<td>237 (84.0)</td>
</tr>
<tr>
<td>Tumor site — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oropharynx</td>
<td>240 (85.1)</td>
<td>226 (83.7)</td>
</tr>
<tr>
<td>Larynx</td>
<td>18 (6.4)</td>
<td>19 (6.7)</td>
</tr>
<tr>
<td>Hypopharynx</td>
<td>15 (5.3)</td>
<td>14 (5.0)</td>
</tr>
<tr>
<td>Tumor stage — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 or T2</td>
<td>162 (57.4)</td>
<td>160 (56.7)</td>
</tr>
<tr>
<td>T3 or T4</td>
<td>116 (41.1)</td>
<td>116 (41.1)</td>
</tr>
<tr>
<td>Nodal stage — no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2a or N2b</td>
<td>221 (78.4)</td>
<td>222 (78.7)</td>
</tr>
<tr>
<td>N2c</td>
<td>52 (18.4)</td>
<td>52 (18.4)</td>
</tr>
<tr>
<td>N3</td>
<td>9 (3.2)</td>
<td>8 (2.8)</td>
</tr>
<tr>
<td>HPV status — no./total no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p16-positive</td>
<td>164/226 (72.6)</td>
<td>171/220 (77.7)</td>
</tr>
<tr>
<td>p16-negative</td>
<td>62/226 (27.4)</td>
<td>49/220 (22.3)</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. There were no statistically significant differences between the two groups at baseline. Additional details are provided in the Table S2 in the Supplementary Appendix. HPV denotes human papillomavirus.
apy, half declined surgery and the other half did not undergo surgery owing to clinical reasons such as progression of disease, inoperability of recurrent disease, or insufficient medical fitness for surgery (Fig. 1, and Table S4 in the Supplementary Appendix).

Of the patients who underwent neck dissection before chemoradiotherapy, 57% underwent modified radical dissection, and the rest underwent selective node dissection. In contrast, of the patients who underwent planned neck dissection after chemoradiotherapy, only 38% underwent modified radical neck dissection (Table S5 in the Supplementary Appendix).

**SURVEILLANCE GROUP**

Of the 282 patients who were randomly assigned to the surveillance group, 270 (96%) underwent PET-CT scanning according to the protocol at 12 weeks (median, 11.1 weeks; interquartile range, 10.4 to 12.4). There was a 92% concordance between local and central PET-CT reviewers with respect to assessment of the primary site and a 97% concordance with respect to assessment of nodal disease.

Local PET-CT reports indicated complete imaging responses in both the primary site and the neck nodes in 185 of the 270 patients who underwent PET-CT (69%). Four of these 185 patients (2%) underwent neck dissection rather than surveillance. Complete responses in the primary site with incomplete or equivocal responses in the neck nodes were seen in 47 of 270 patients (17%). Of these 47 patients, 36 (77%) underwent neck dissection. The reasons that patients did not undergo a neck dissection are listed in Table S6 in the Supplementary Appendix.

A total of 15 patients (6%) had an incomplete response in the primary site but a complete response in the neck nodes; these patients did not undergo neck dissection. An additional 19 patients (7%) had incomplete responses in both the primary site and the neck nodes; 12 of these 19 patients (63%) underwent neck dissection. In 4 patients (1%), PET-CT after chemoradiotherapy showed new lesions that indicated disease progression with distant metastases or new primary sites in the lung. Two patients underwent neck dissection without undergoing PET-CT. Of the 9 patients with N3 disease, 2 did not undergo PET-CT because of early disease progression or death, 5 had a complete response, and 2 had persistent nodal disease and underwent neck dissection. Overall, 54 neck dissections were performed in the surveillance group, as compared with 221 in the planned-surgery group.

**OUTCOMES AND EFFICACY**

Patients were followed for up to 5 years, with a median follow-up of 36 months; 520 patients (92%) were followed for at least 2 years. Overall, 122 patients died (60 in the surveillance group and 62 in the planned-surgery group). The 2-year overall survival rate was 84.9% (95% confidence interval [CI], 80.7 to 89.1) in the surveillance group and 81.5% (95% CI, 76.9 to 86.3) in the planned-surgery group. The hazard ratio for death with surveillance as compared with planned surgery was 0.92 (95% CI, 0.65 to 1.32); this outcome slightly favored the surveillance group and met the prespecified definition of noninferiority (upper boundary of the 95% CI for the hazard ratio, <1.50; P=0.004). The hazard ratio of 0.92 excluded an unfavorable difference of more than 4 percentage points (at a one-sided alpha level of 0.05) between the two groups. The two-sided P value for the difference between treatment strategies was 0.66.

The results were similar after adjustment for treatment center, tumor stage, nodal stage, primary tumor site, chemotherapy and radiotherapy schedules, sex, and age at randomization. In addition, the results were similar after further adjustment for HPV status on the basis of available samples (Table S7 in the Supplementary Appendix).

Disease-specific mortality and mortality from other causes did not differ significantly between the two groups (P=0.80 and 0.41, respectively, according to Gray’s test for differences).19 Status with respect to p16 expression was highly prognostic of overall survival in both groups (Fig. 2), a finding that was consistent with results of previous studies. There was no significant difference in overall survival between the planned-surgery and surveillance groups. However, the median survival was longer in the surveillance group, with 2-year survival rates of 84.9% (95% CI, 80.7 to 89.1) in the surveillance group and 81.5% (95% CI, 76.9 to 86.3) in the planned-surgery group. The hazard ratio for death with surveillance as compared with planned surgery was 0.92 (95% CI, 0.65 to 1.32); this outcome slightly favored the surveillance group and met the prespecified definition of noninferiority (upper boundary of the 95% CI for the hazard ratio, <1.50; P=0.004). The hazard ratio of 0.92 excluded an unfavorable difference of more than 4 percentage points (at a one-sided alpha level of 0.05) between the two groups. The two-sided P value for the difference between treatment strategies was 0.66.

The results were similar after adjustment for treatment center, tumor stage, nodal stage, primary tumor site, chemotherapy and radiotherapy schedules, sex, and age at randomization. In addition, the results were similar after further adjustment for HPV status on the basis of available samples (Table S7 in the Supplementary Appendix).

Disease-specific mortality and mortality from other causes did not differ significantly between the two groups (P=0.80 and 0.41, respectively, according to Gray’s test for differences).19 Status with respect to p16 expression was highly prognostic of overall survival in both groups (Fig. 2), a finding that was consistent with results of previous studies. There was no significant difference in overall survival between the planned-
564 Eligible patients underwent randomization

282 Were assigned to surveillance group
276 Received chemoradiotherapy
2 Proceeded directly to neck dissection
1 Had disease progression
3 Died

270 Underwent PET-CT scanning
4 Had incomplete imaging results

185 Had a complete response in primary tumor and nodes
181 Did not undergo neck dissection
4 Underwent neck dissection

47 Had a complete response in primary tumor only
36 Underwent neck dissection
11 Did not undergo neck dissection
7 Did not undergo neck dissection

19 Did not have a complete response in either primary tumor or nodes
12 Underwent neck dissection
7 Did not undergo neck dissection

276 Received chemoradiotherapy
282 Patients assigned to surveillance group were followed up
259 Were alive at 1 yr
24 Died
11 Were lost to follow-up
224 Were alive at 2 yr
13 Died
101 Did not undergo further follow-up
110 Were alive at 3 yr
5 Died
22 Did not undergo further follow-up
33 Were alive at 4 yr
27 Did not undergo further follow-up
6 Were alive at 5 yr

282 Were assigned to planned-surgery group

77 Planned to undergo neck dissection before chemoradiotherapy

205 Planned to undergo neck dissection after chemoradiotherapy
69 Underwent neck dissection
199 Received chemoradiotherapy
63 Received chemoradiotherapy after neck dissection
152 Underwent neck dissection after chemoradiotherapy

282 Patients assigned to planned surgery were followed up
243 Were alive at 1 yr
21 Died
18 Withdraw
204 Were alive at 2 yr
4 Died
82 Did not undergo further follow-up
118 Were alive at 3 yr
32 Were alive at 4 yr
1 Died
23 Did not undergo further follow-up
82 Were alive at 5 yr
1 Died
10 Withdraw

surgery and surveillance groups among patients with p16-positive tumors (hazard ratio, 0.74; 95% CI, 0.40 to 1.37) and those with p16-negative tumors (hazard ratio, 0.98; 95% CI, 0.58 to 1.66) (Fig. 3).

**Patterns of Relapse**

The 2-year rate of locoregional control was 91.9% (95% CI, 88.5 to 95.3) in the surveillance group and 91.4% (95% CI, 87.8 to 95.0%) in the planned-surgery group. In the latter group, the 2-year rate of locoregional control was 90.4% (95% CI, 86.0 to 94.7) among patients who underwent neck dissection after chemoradiotherapy and 94.8% (95% CI, 89.0 to 100) among patients who underwent neck dissection before chemoradiotherapy.

Documented recurrence in the nodes only (without concurrent disease in the primary site) occurred in 1 patient in the planned-surgery group and in 3 patients in the surveillance group. Distant metastases were identified in 23 patients in the planned-surgery group and in 21 patients in the surveillance group.

**Figure 2. Kaplan–Meier Estimates of Overall Survival, According to Trial Group.**

In Panel B, the numbers of patients shown are the numbers of patients in those groups at randomization.
### Figure 3. Overall Survival, According to Subgroup.

The size of the squares corresponds to the number of patients with an event. The diamond incorporates the point estimate and the 95% confidence interval of the overall effect. TPF denotes docetaxel plus cisplatin and fluorouracil.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>PET-CT Surveillance</th>
<th>Planned Surgery</th>
<th>Hazard Ratio for Death (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timing of neck dissection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before chemoradiotherapy</td>
<td>13/76 (17.1)</td>
<td>18/77 (23.4)</td>
<td>0.67 (0.33–1.35)</td>
<td>0.29</td>
</tr>
<tr>
<td>After chemoradiotherapy</td>
<td>47/206 (22.8)</td>
<td>44/205 (21.5)</td>
<td>1.03 (0.68–1.56)</td>
<td></td>
</tr>
<tr>
<td><strong>Test of interaction between 2 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tumor stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 or T2</td>
<td>28/162 (17.3)</td>
<td>18/160 (11.3)</td>
<td>1.50 (0.84–2.68)</td>
<td>0.04</td>
</tr>
<tr>
<td>T3 or T4</td>
<td>32/116 (27.6)</td>
<td>42/116 (36.2)</td>
<td>0.70 (0.44–1.11)</td>
<td></td>
</tr>
<tr>
<td>Occult disease</td>
<td>0/4</td>
<td>2/6 (33.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test of heterogeneity among 3 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nodal stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2a or N2b</td>
<td>36/221 (16.3)</td>
<td>42/222 (18.9)</td>
<td>0.83 (0.53–1.29)</td>
<td></td>
</tr>
<tr>
<td>N2c or N3</td>
<td>24/61 (39.3)</td>
<td>20/60 (33.3)</td>
<td>1.11 (0.61–2.01)</td>
<td>0.45</td>
</tr>
<tr>
<td><strong>Test of interaction between 2 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cancer site</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral cavity</td>
<td>1/4 (25.0)</td>
<td>1/7 (14.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oropharynx</td>
<td>46/240 (19.2)</td>
<td>42/236 (17.8)</td>
<td>1.05 (0.69–1.59)</td>
<td></td>
</tr>
<tr>
<td>Larynx</td>
<td>7/18 (38.9)</td>
<td>8/19 (42.1)</td>
<td>0.76 (0.27–2.16)</td>
<td></td>
</tr>
<tr>
<td>Hypopharynx</td>
<td>6/15 (40.0)</td>
<td>8/14 (57.1)</td>
<td>0.45 (0.15–1.32)</td>
<td></td>
</tr>
<tr>
<td>Occult disease</td>
<td>0/5</td>
<td>3/6 (50.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test of heterogeneity among 5 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.26</td>
</tr>
<tr>
<td><strong>Chemotherapy schedule</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concomitant platin</td>
<td>31/163 (19.0)</td>
<td>35/169 (20.7)</td>
<td>0.88 (0.54–1.43)</td>
<td></td>
</tr>
<tr>
<td>Concomitant cetuximab</td>
<td>5/14 (35.7)</td>
<td>5/14 (35.7)</td>
<td>1.03 (0.30–3.60)</td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant and concomitant platin</td>
<td>1/10 (10.0)</td>
<td>4/9 (44.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant TPF with concomitant platin</td>
<td>22/89 (24.7)</td>
<td>18/85 (21.2)</td>
<td>1.10 (0.59–2.04)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1/5 (20.0)</td>
<td>0/5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test of heterogeneity among 5 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>54/223 (24.2)</td>
<td>50/237 (21.1)</td>
<td>1.15 (0.78–1.69)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>6/59 (10.2)</td>
<td>12/45 (26.7)</td>
<td>0.29 (0.11–0.77)</td>
<td></td>
</tr>
<tr>
<td><strong>Test of interaction between 2 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 yr</td>
<td>6/37 (16.2)</td>
<td>9/43 (20.9)</td>
<td>0.77 (0.27–2.15)</td>
<td></td>
</tr>
<tr>
<td>50–59 yr</td>
<td>28/135 (20.7)</td>
<td>23/113 (20.4)</td>
<td>0.96 (0.55–1.68)</td>
<td></td>
</tr>
<tr>
<td>60–69 yr</td>
<td>22/90 (24.4)</td>
<td>24/107 (22.4)</td>
<td>1.09 (0.61–1.95)</td>
<td></td>
</tr>
<tr>
<td>≥70 yr</td>
<td>4/20 (20.0)</td>
<td>6/19 (31.6)</td>
<td>0.58 (0.16–2.14)</td>
<td></td>
</tr>
<tr>
<td><strong>Test of heterogeneity among 4 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td><strong>p16 status</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.97</td>
</tr>
<tr>
<td>Positive</td>
<td>17/164 (10.4)</td>
<td>23/171 (13.5)</td>
<td>0.74 (0.40–1.37)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>33/62 (53.2)</td>
<td>25/49 (51.0)</td>
<td>0.98 (0.58–1.66)</td>
<td></td>
</tr>
<tr>
<td>Status not known</td>
<td>10/56 (17.9)</td>
<td>14/62 (22.6)</td>
<td>0.76 (0.34–0.70)</td>
<td></td>
</tr>
<tr>
<td><strong>Test of heterogeneity among 3 groups</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>All patients</td>
<td>60/282 (21.3)</td>
<td>62/282 (22.0)</td>
<td>0.92 (0.65–1.32)</td>
<td></td>
</tr>
</tbody>
</table>
Surgical Complications, Serious Adverse Events, and Global Quality of Life

A total of 22 surgical complications after neck dissection were noted in the surveillance group, as compared with 83 in the planned-surgery group. The rate of complications among patients who underwent neck dissection in the surveillance group was 42% (95% CI, 24 to 59). The rates of complications and severe complications in the planned-surgery group were 38% (95% CI, 30 to 46) and 26% (95% CI, 0.15 to 0.41), respectively, with similar rates among the patients who underwent planned surgery after chemoradiotherapy (39% [95% CI, 29 to 49] and 26% [95% CI, 0.13 to 0.28], respectively) and among patients who underwent planned surgery before chemoradiotherapy (35% [95% CI, 21 to 49] and 23% [95% CI, 12 to 40]). A smaller percentage of modified radical neck dissections were performed in the groups of patients who underwent planned surgery after chemoradiotherapy than in the group of patients who underwent planned surgery before chemoradiotherapy (Table S5 in the Supplementary Appendix). A total of 282 serious adverse events occurred: 169 in the planned-surgery group and 113 in the surveillance group (Table S8 in the Supplementary Appendix).

There was a small difference in global health status scores on the EORTC QLQ-C30 questionnaire in favor of the surveillance group at 6 months after randomization (mean change in the score at 6 months in the surveillance group relative to the planned-surgery group, 4.94; P=0.03). This difference narrowed at 12 months (mean change, 3.03; P=0.09) and disappeared by 24 months (mean change, −0.81; P=0.85) (Fig. S1 in the Supplementary Appendix).

Cost-Effectiveness

Over the 2-year minimum follow-up period, PET-CT–guided surveillance was more cost-effective than planned neck dissection. The per-person cost saving was £1,492 (approximately $2,190 in U.S. dollars), with an additional 0.08 QALYs per person.

Discussion

There has been lack of clarity about treatments after chemoradiotherapy and wide variation in the clinical treatment of patients with squamous-cell carcinoma of the head and neck who have advanced nodal disease and who have received chemoradiotherapy for primary treatment. Our trial showed that PET-CT–guided surveillance was noninferior to planned neck dissection and was equally effective in both HPV-positive and HPV-negative patient groups. Patients in the surveillance group were not disadvantaged by undergoing delayed neck dissection; the global quality-of-life scores and rates of surgical complications were similar in this group and in the group of patients who underwent earlier planned neck dissection. PET-CT surveillance resulted in far fewer operations; approximately 80% of patients were spared neck dissection. Surveillance was also more cost-effective than planned neck dissection over the trial period.

There was a high concordance between local radiologic and central laboratory assessments in this multicenter, nationwide trial. This high concordance confirms the feasibility of PET-CT–guided surveillance in routine clinical practice.

Furthermore, our trial may actually underestimate the benefit of PET-CT–guided surveillance in patients with advanced head and neck cancer. Our protocol recommended neck dissection in patients with equivocal responses (PET-CT–negative residual masses or mild FDG uptake in normal-sized nodes) because of the high failure rate (37%) reported among patients with equivocal responses on CT after chemoradiotherapy. However, a recent study suggests that nodal disease may take longer to involute in patients with HPV-positive disease. It is therefore conceivable that patients in our trial who had HPV-positive tumors and equivocal PET-CT findings (especially with enlarged nodes) at the 3-month assessment might have achieved a cure without neck dissection if they had undergone PET-CT at a later time. Other researchers have reported that nodes with no FDG uptake have very high rates of regional control (93%), especially in HPV-positive disease.

We recommend that patients with an equivocal FDG uptake should continue to undergo neck dissection, especially if they have HPV-negative disease. However, patients with HPV-positive cancers who have enlarged nodes but no FDG uptake after chemoradiotherapy may be considered for close follow-up with serial CT or PET-CT; this strategy may spare even more patients from undergoing a neck dissection.

Our results are consistent with those associ-
ated with other management approaches that are based on nodal response and that use CT. These approaches have shown high rates of complete response to chemoradiotherapy, with relatively low proportions of patients undergoing neck dissection and good rates of long-term control of locoregional disease. However, unlike our trial, these studies were often limited because they were retrospective studies from single, high-throughput institutions. In addition, they performed earlier (at 8 weeks) assessment of the nodal response (which is known to be less accurate than later assessment), and they used histopathological findings of neck dissection specimens after chemoradiotherapy as evidence of persistent disease (this method is known to overestimate persistence).11,24

Studies that have compared PET-CT with CT show higher efficacy of PET-CT in patients with advanced head and neck cancer. In one study, PET-CT was superior to CT in high-risk patients who had a clinically significant history of smoking or alcohol use or HPV-negative, nonoropharyngeal primary tumors. In most countries, these groups constitute the majority of both HPV-positive and HPV-negative patients.25,26 In another study, PET-CT was significantly more accurate than CT in identifying a complete nodal response, especially in patients with HPV-positive disease (93% vs. 50%).21 A prospective single-institution study also showed that PET-CT surveillance was more cost-effective than CT surveillance, regardless of HPV status.24 In the same study, CT followed by PET-CT in patients who did not have a response was shown to be only marginally more cost-effective than PET-CT alone, and with small changes in baseline assumptions and costs, CT followed by PET-CT ceased to be cost-effective as compared with PET-CT alone.24 The relative efficacy and cost-effectiveness of CT, as compared with those of PET-CT–based approaches, require further evaluation.

At the time of the inception of this trial, it was not possible to calibrate standard uptake values among various scanning systems. Interpretation of the results presented here is therefore limited by the fact that we could not undertake assessments of standard uptake values in determining the patients’ response to therapy. Since calibration among systems is now feasible, we are undertaking a retrospective evaluation of standard uptake values in the scans used in this study to assess whether this improves the accuracy of the response assessments.

When extrapolating the data presented here to routine clinical practice, clinicians should note that few patients in our trial had low-prevalence, N3 (stage IVb) disease. Although 5 of the 9 patients with stage N3 disease in the PET-CT surveillance group had complete responses, extrapolation of a PET-CT–guided surveillance policy to this higher-risk group of patients cannot currently be justified because of the small number of such patients in the trial.


Supported by academic grants from the National Institute for Health Research Health Technology Assessment Programme (06/302/129) and Cancer Research UK (C19677/A9674, for tissue-sample collection).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank all the patients and their clinicians for their participation in the trial; the trial steering committee: Dr. Jonathan Ledermann, Mr. Malcolm Babb, Dr. Michael Langman (deceased), Dr. Patrick Bradley, and Mrs. Ethel Culling (independent members) and Drs. Roger A’Hearn, Vincent Gregoire, and Alfonso Balm (independent data and safety monitoring committee); Ms. Jo Grumett and Mr. Dharmesh Patel (Warwick Clinical Trials Unit) for trial coordination; Mrs. June Jones and Ms. Lynda Wagg (the Institute of Head and Neck Studies and Education, University of Birmingham) for collection of health economic data; Mr. Bal Sanghera and Ms. Wendy Sookham (Mount Vernon Hospital) for coordinating the central PET-CT laboratory; the University of Warwick and University Hospitals Coventry and Warwickshire NHS Trust for acting as legal cosponsors of the study; and Dr. John Chester (Cardiff University) and Dr. Kevin Harrington (Royal Marsden Hospital, London) for reviewing an earlier version of the manuscript and Ms. Gemma Jones for transcribing an earlier version of the manuscript.

REFERENCES


