CHAPTER 3

A randomized trial of PET scanning to improve diagnostic yield of direct laryngoscopy in patients with suspicion of recurrent laryngeal carcinoma after radiotherapy

R. de Bree, L. van der Putten, O.S. Hoekstra, D.J. Kuik, C.A. Uyl-de Groot, H. van Tinteren, C.R. Leemans & M. Boers

on behalf of the RELAPS Study Group

Contemporary Clinical Trials 2007;28:705-12
ABSTRACT
The RELAPS study (REcurrent LAryngeal carcinoma PET Study) was designed to determine whether FDG–PET is of value in the selection of patients for direct laryngoscopy under general anesthesia in patients with suspicion of recurrent laryngeal carcinoma after radiotherapy. In a randomized controlled clinical trial the current diagnostic practice, i.e. all patients undergo direct laryngoscopy, will be compared to a strategy in which FDG–PET selects the patients for laryngoscopy. All eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group NWHHT will participate in this multicenter trial. The study population consists of patients with clinical suspicion of recurrent T2–T4 laryngeal carcinoma after radiotherapy (without obvious signs of tumor) in whom a direct laryngoscopy under general anesthesia with taking of biopsies is indicated by the local physician. The primary efficacy endpoint is the difference in the number of futile indications for direct laryngoscopy between the conventional diagnostic arm and the FDG–PET based diagnostic arm. An indication for laryngoscopy is classified as futile if this laryngoscopy was negative and no recurrence was diagnosed within 6 months follow-up (gold standard). The FDG–PET based strategy may increase the risk of missing recurrent tumor compared to current practice. Safety endpoints include survival and morbidity due to laryngoscopy with taking of biopsies. Survival rates of both groups will have to be collected outside the time frame of the funded trial. Resectability of recurrent tumor and tumor negative surgical margins after total laryngectomy will be used as proxy endpoints. The trial will also compare quality of life and direct medical costs between both arms.
INTRODUCTION

Laryngeal carcinoma is the most common form of head and neck cancer. When treating patients for laryngeal cancer, the goal is not only to cure but also to preserve function. Early laryngeal cancer can usually be managed successfully with either radiotherapy or surgery. Carefully selected advanced lesions are initially treated by irradiation, with surgery reserved for salvage treatment. With the emphasis on preservation of organ and function, investigational treatment regimes using altered fractionation schedules of radiation and the combination of chemotherapy and radiation have recently emerged. There is a high local control rate of T1 laryngeal cancer treated with radiotherapy. For T2 to T4 laryngeal cancer treated with radiotherapy with curative intent, the rate of recurrence is considerable. In most cases where radical radiotherapy has failed, salvage surgery to treat recurrence remains a successful option.

In patients with symptoms or clinical abnormalities such as severe edema or necrosis, differentiation between recurrent carcinoma and sequelae of radiotherapy can be difficult. The need for biopsy can present a dilemma as the biopsy itself may exacerbate postradiotherapy changes and initiate superimposed infection, (peri)chondritis, failure to heal and further edema (1,2). On the other hand, symptomatic treatment with antibiotics and glucocorticoids will delay adequate treatment of a recurrence. Currently most physicians aggressively pursue potential recurrences, leading to a high rate of futile direct laryngoscopies and a waste of scarce health care resources.

Improvement of this situation requires increasing the a priori likelihood of a recurrence before laryngoscopy. In a retrospective cohort study, we defined the diagnostic accuracy and yield of signs and symptoms, the current noninvasive diagnostic techniques and direct laryngoscopy with biopsy in patients with suspicion of recurrent laryngeal carcinoma after radiotherapy. Overall, the current clinical diagnostic strategy led to the detection of a recurrence in only 45% of the direct laryngoscopies. We found that the accuracy of voice complaints, pain, dyspnoea, dysphagia, indirect laryngoscopy, videolaryngostroboscopy and CT or MRI was low (3,4).

The use of positron emission tomography (PET) with 18-fluorodeoxyglucose (FDG) seems to be promising. Because FDG–PET relies on the metabolic function of neoplastic cells, it can be an important tool in the detection of small, submucosal recurrences. Furthermore, it has been shown that FDG–PET is able to distinguish tumor recurrence from radiation sequelae in patients treated for laryngeal carcinoma (5).

In The Netherlands almost all laryngeal carcinomas are treated in the eight head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group NWHHT. There are no clear guidelines for the diagnostic policy when clinical suspicion of recurrent laryngeal carcinoma after radiotherapy exists. We surveyed the otolaryngologists and radiotherapists in these clinics to review current diagnostic practice. All respondents indicated that they perform direct laryngoscopy under general anesthesia when recurrence is suspected. However, 35% were dissatisfied with the current diagnostic path. Many indicated that they would like to have easier access to a PET-scan facility, since they expected that this technique might improve the diagnostic path.
The assessment and distribution of new technologies is a complex process. There may be either a low level of acceptance of new technologies in the medical community, or, more frequently, overuse. In the setting of recurrent laryngeal carcinoma a trial that randomly allocates patients to either conventional diagnostic work-up or to a work-up based on a new technique may be the best way to determine the value of FDG–PET. The major benefit of randomization lies in the creation of groups that are similar with respect to all known and unknown prognostic factors allowing an unbiased comparison of different strategies.

In our setting the achievable health gain comprises a reduction of the number of futile procedures. The RELAPS study (REcurrent LAryngeal carcinoma PET Study) will compare the current diagnostic practice, i.e. all patients undergo direct laryngoscopy, to a strategy in which FDG–PET selects the patients for laryngoscopy. The trial will also compare quality of life and direct medical costs. Randomization will be stratified by treating institute, T-stage (T2 vs T3–4) and smoking status.

**METHODOLOGY**

**Population**

All 8 head and neck cancer centers of the Dutch Head and Neck Oncology Cooperative Group will together recruit 150 patients for this trial. These centers treat more than 90% of all patients with laryngeal carcinoma in the Netherlands. Patients with clinical suspicion of recurrent laryngeal carcinoma after radiotherapy (without obvious signs of recurrent tumor), for whom a direct laryngoscopy under general anesthesia with taking of biopsies is indicated by the local physicians, are invited to participate. Patients have to be older than 18 years of age and give a written informed consent according to the local medical ethical committee regulations. Patients must have received and completed their radiotherapy for a histological proven T2–T4 laryngeal carcinoma for at least 4 months.

**Randomization**

In a prospective randomized controlled trial, two strategies are compared: 1) conventional strategy: direct laryngoscopy under general anesthesia with taking of biopsies; 2) FDG–PET based strategy: only direct laryngoscopy under general anesthesia with taking of biopsies if FDG–PET is positive or equivocal (Fig. 1). To randomize a patient, the otolaryngologists call the central data center. Eligible patients are randomized by a computer according to the method of minimization. Randomization will be stratified by treating institute, T-stage (T2 vs T3–4) and smoking status.
PET procedures

PET scans are performed in the local head and neck centers. For the purpose of this study, PET centers have to guarantee to perform the scan within 2 weeks. Prior to scanning, the patients need to fast for 6 h. The 20 min head and neck acquisition starts about 1 h after injection of 370 MBq FDG, with a scanned trajectory from skull base to clavicle. PET images are made with “state of the art techniques” according to local protocols. The data supplied by the physician contains only the stage, (sub)site and side of the laryngeal carcinoma of initial presentation. The scoring of laryngeal lesions is visually related to the neck background activity, using a three-point scale: negative (no abnormal activity more than background), equivocal and positive (enhanced tracer uptake in the larynx not compatible with physiological uptake). This interpretation is communicated to the referring physician by telephone and confirmed in writing.

Conventional strategy arm: all patients will undergo direct laryngoscopy. Per center it is optional to make a ‘blinded’ FDG–PET of each patient. Before the start of the study, each center has to decide to perform a FDG–PET in each patient or not. This FDG–PET will not be examined until the end of the study. If a biopsy shows recurrent tumor the patients will be scheduled for a total (or partial) laryngectomy. If the direct laryngoscopy is negative or equivocal a direct laryngoscopy will be repeated unless clinical symptoms or signs diminish over time. The reason for abandoning the second direct laryngoscopy has to be specified in the case report form.
**FDG–PET based strategy arm:** all patients undergo FDG–PET. When the result is positive or equivocal, a direct laryngoscopy under general anesthesia will be performed. If the biopsy shows recurrent tumor, the patient will be scheduled for a total (or partial) laryngectomy. In case of a negative or equivocal direct laryngoscopy, the direct laryngoscopy will be repeated without exception. If the FDG–PET is negative, expectative follow-up (no additional investigations) has to be maintained for at least 3 months. In case of a progression of clinical symptoms or signs within the first 3 months, a direct laryngoscopy is indicated.

**Follow-up**
Patients will visit the outpatient clinic on a regular basis (4–8 weeks). Follow-up of at least 6 months after inclusion is mandatory. Because repeated negative biopsies do not exclude the presence of a recurrent tumor, a combination of pathological examination and 6 months follow-up will be used as ‘gold standard’.

**Outcome parameters**
The primary efficacy endpoint is the difference in the number of futile indications for direct laryngoscopy under general anesthesia between the conventional diagnostic arm and the FDG–PET based diagnostic arm. An indication for laryngoscopy is classified as futile if this laryngoscopy was negative and no recurrence was diagnosed within 6 months follow-up. An indication is considered justified if a recurrence is diagnosed at laryngoscopy (procedure result true positive) but also if the recurrence is diagnosed within 6 months after a negative procedure (laryngoscopy result false negative).

To answer the question whether a FDG–PET based strategy can be more cost-effective and safe, secondary endpoints include costs, operability of a recurrence, surgical margins of the salvage laryngectomy and quality of life.

**Safety**
The FDG–PET-based strategy may increase the risk of missing recurrent tumor compared to current practice. The protocol calls for early referral for direct laryngoscopy in case of progression of clinical symptoms, which will minimize the delay of a diagnosis.

Safety endpoints include survival and morbidity due to laryngoscopy with taking of biopsies. Two, three and fiveyear survival rates of both groups will have to be collected outside the time frame of the trial. Resectability of recurrent tumor and tumor negative surgical margins after total laryngectomy will be used as proxy endpoints. We will consider the FDG–PET based strategy to be equivalent in safety when the outcome of these parameters are better, equal, or no more than 5% worse, compared to the conventional strategy.

**Quality of life**
Quality of life will be measured by questionnaires (EORTC QLQ-C30, QLQ-H&N35 and EQL-5D). The timing is: 1) before the direct laryngoscopy, 2) at the first visit in the outpatient clinic after direct laryngoscopy or FDG-PET, and 3) 6 months after randomization. The results will be compared between both arms (6–8).
Sample size calculation

Before the trial was designed, a pilot FDG–PET study was performed in which a percentage of futile indications for direct laryngoscopy of 38% was found [4, manuscript in preparation]. In the largest prospective accuracy study FDG–PET would have reduced the number of futile indications by two-thirds (66%). With this reduction (from 38% to 13%) as our aim, sample size calculation on Fisher’s Exact test with a two-sided significance level of 0.05 and a power of 85%, reveals a requirement of 59 patients per group. However, considering the nature of the experimental technique in terms of burden to patients or physicians a larger number of patients has to be accrued in the trial. Moreover, the data used for above mentioned power analysis are based on a limited number of reported patients. The trial is therefore open for an accrual period of 28 months, with a minimum of 150 patients.

Data-analysis

The number of futile indications for surgical procedures in both groups, will be simply expressed as a binomial value and can be tested using a Chi-square test (or non-parametric alternative). In secondary analyses, logistic regression will allow the inclusion of certain potentially confounding variables into the analyses, such as age of the patient, clinical stage at initial presentation, but also number of recurrent tumors and the percentage recurrences suitable for ‘salvage’ total laryngectomy at the end of the study. Test characteristics (sensitivity, specificity, positive and negative predictive values) of the two selection strategies will be compared. As laryngoscopy itself is not a perfect test, the presence or absence of recurrent disease will be determined in the whole of the 6-month follow-up. Thus, as indicated previously the selection strategy for laryngoscopy is considered true positive if a direct laryngoscopy is indicated and a recurrence is diagnosed within 6 months, even if the initial direct laryngoscopy was negative. In this case the indication was justified. The selection strategy is considered to be false positive if a direct laryngoscopy is indicated and no recurrence is diagnosed within 6 months. In this case the indication was futile. The selection strategy is considered to be true negative when a laryngoscopy is not indicated, not performed and no recurrence is diagnosed within 6 months. In this case the lack of indication is justified. Finally, the selection strategy is considered to be false negative when a laryngoscopy is not indicated (and not performed) but a recurrence is diagnosed within 6 months. In this case, the lack of indication is false (Table 1). Because of the expected low risk of PET imaging and the relative short accrual time, no interim statistical analysis is planned.

<table>
<thead>
<tr>
<th>Table 1. Comparison of strategy outcome with gold standard.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recurrence within 6 months</strong></td>
</tr>
<tr>
<td><strong>Yes</strong></td>
</tr>
<tr>
<td>True positive: laryngoscopy justified</td>
</tr>
<tr>
<td>False negative: lack of indication false</td>
</tr>
</tbody>
</table>

Strategy: indication for laryngoscopy

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>True positive: laryngoscopy justified</td>
<td>False positive: laryngoscopy futile</td>
</tr>
<tr>
<td>False negative: lack of indication false</td>
<td>True negative: lack of indication justified</td>
</tr>
</tbody>
</table>
Economic evaluation

The average total costs per patient will be determined for a period of 6 months or until local recurrence is diagnosed; from the moment of suspicion of recurrent laryngeal carcinoma after radiotherapy until the 6 months follow-up, local recurrence or death, whichever occurs first. Since salvage treatment is far more costly than the diagnostic procedures, the period after diagnosis of recurrence will not be included in this analysis. The study will focus on direct medical costs and record the following cost types: hospital days, daycare treatments, outpatient visits, medical procedures, diagnostic tests (including FDG–PET), surgical interventions (including laryngoscopy), radiotherapy, blood products and medications. The volumes of the mentioned cost items will be extracted from hospital databases and patient files. For the most relevant cost items new unit costs will be determined, reflecting real resource use, including a raise for overhead costs [9]. To determine these unit costs, the micro-costing method will be used. This method is based on a detailed inventory and measurement of all resources consumed (10). For items with low costs or a negligible influence (due to low average numbers), Dutch tariffs will be used. Costs of medication will be based on a Dutch pharmacotherapeutic reference (11). If the follow up period is less than 6 months, no discounting will be applied. Total costs, based on resource use multiplied by unit prices, will subsequently be calculated and compared for the diagnostic strategies. If FDG–PET proves to be clinically effective, sensitivity analysis based on accuracy measures provided by this study, together with information from the literature, will be performed.

Logistics

Prior to the study, a test-set of 30 FDG–PET scans will be sent to the participating nuclear medicine physicians in each center. In this way the inter-observer variation can be examined and if necessary the interpretation can be standardized. The presence of recurrent tumor will be scored (scale: negative, equivocal, positive). Medical-Ethics approval will be requested for each center. Before starting the trial a plenary meeting will be organized. During the trial a newsletter with accrual data will be sent every 3 months.

DISCUSSION

The introduction of new health technology requires justification of cost-effectiveness for specific clinical indications. A diagnostic imaging technique is considered ‘effective’ if it not only provides more accurate data than existing modalities, but also improves patient management, and ultimately it should contribute to have a favorable impact on health status at reasonable costs. Randomized clinical trials are extremely rare in the evolution of diagnostic tests. Acceptance of new diagnostic tests seems to be based on published accuracy studies for a certain indication. As a consequence, new and expensive but not optimally effective techniques often diffuse into clinical practice, potentially slowing down innovations in other areas (12). Accuracy studies provide information which is indicative but may be of little relevance to the actual effectiveness of a test (13).
We have documented an unsatisfactory diagnostic process in the follow-up after treatment of laryngeal carcinoma, and propose a trial that examines the potential for FDG–PET in this setting. In a pilot study we showed that PET might substantially reduce the number of futile indications for laryngoscopies. We felt a randomized trial was the only way to overcome the problem of variable and uncontrollable factors in the regular clinical diagnostic workup. In a randomized trial the experimental strategy is used in real clinical practice instead of a model, in which the application and decision making of an experimental strategy may be different.

In the proposed randomized trial eligibility criteria for patients have been kept to a minimum to guarantee fast accrual and generalizability. Patients without diagnostic dilemma are excluded: e.g. patients with clinically evident recurrent disease and patients with T1 tumors. In the latter, recurrence is highly unlikely due to effective radiotherapy. Immediately after radiotherapy FDG–PET may show increased uptake in the irradiated area so that it is less reliable at that stage. It has been shown that from 4 months after radiotherapy FDG–PET was a better predictor for the presence or absence of residual or recurrent carcinoma (14). Since relatively few patients (14–20% (3,5)) present with suspected recurrence within 4 months after completed radiotherapy, we decided to include only patients presenting with suspected recurrence beyond this time interval.

Patients are stratified for center, T-stage (T2 versus T3/4) and smoking, as these factors are likely to affect the rate of laryngoscopies (3).

PET scans are examined by the local specialists to improve applicability of the trial results. To examine the interobserver variation and standardize interpretation a test-set of 30 FDG–PET scans will be sent to the participating nuclear medicine physicians in all centers before the start of the study. A ‘blinded’ FDG–PET is optional in the conventional arm. These images will not be examined until the end of the study so that its result does not affect decision making in the control arm. At the same time, these data are suited to improve the estimation of test accuracy measures. Centers with ethical concerns in this respect will not use the option of the blinded PET scan in the control arm.

The follow-up period in this study is at least 6 months after FDG–PET; it is expected that recurrent tumor will become manifest in this period (3).

In the proposed study the results of the current ‘state of art’ techniques are analyzed. Therefore, participating centers are asked to scan with the best PET techniques available. Since in this trial a substantial number of patients will undergo PET/CT, the opportunity will be taken to compare PET alone, CT alone and integrated PET/CT in a homogeneous group of patients. Until now, there is only one report on PET/CT in a very heterogeneous group of patients with laryngeal cancer: different treatment modalities (radiotherapy with or without chemotherapy and surgery with or without radiotherapy) and different indications (locoregional recurrence, distant metastases, staging and response to treatment).

In the study design we expect that the delay due to a false negative FDG–PET result is minimal, because direct laryngoscopy is permitted if signs and symptoms progress. Therefore, we expect that this potential delay does not affect curability and prognosis. However, this has to be
evaluated. Endpoints to determine the safety of the FDG–PET based strategy, as compared to the conventional strategy, include morbidity due to laryngoscopy with taking of biopsies and survival. Compared to the conventional arm, the FDG–PET based selection strategy is considered to be safe when the overall survival of the FDG–PET arm is less than 5% lower than the conventional strategy. Since this follow-up period is too long to fit in the granted study time schedule, we will compare the 2, 3 and 5 year survival rates of both groups beyond the final report of the study.

Because of the limited follow-up time within the grant proxy prognostic indicators, e.g. operability of recurrent tumor and tumor negative surgical margins after total laryngectomy have to be used. It is debatable between which limits the alternative prognostic factors should be considered similar. In a retrospective study of direct laryngoscopies under general anesthesia for suspicion of recurrent laryngeal carcinoma after radiotherapy, 91% of the recurrences were eligible for salvage surgery (3). In the literature this figure ranges from 87–95% (15–18). In a retrospective analysis in our institute the results of 61 salvage total laryngectomies were analyzed. In 5% of the total laryngectomies, at least one of the surgical margins was positive (or close). If only T2–T4 recurrent laryngeal carcinomas were taken into account this figure was 6%. In the literature the incidence of positive margins in a salvage laryngectomy depends on the T stage, ranging from 4% to 20% (15–19). Because of the increasing application of larynx preservation treatments, nowadays more advanced tumor stages are treated with radiotherapy (with or without chemotherapy). In previous studies probably a relatively lower number of advanced stages have been included compared to what will happen in the present protocol. Therefore, the aforementioned figures may be higher in the present protocol. We will determine these endpoints (e.g. operability and surgical margins) as an alternative safety analysis at the end of the proposed study, to check if there is a clinically substantial difference. To make definitive conclusions on safety of the FDG–PET based strategy, we will await the long-term follow-up.

The decrease of quality of life and utility induced by the laryngoscopies and their associated morbidity will be assessed by the EORTC QLQ-C30, the QLQ-H&N35 and the EQL-5D questionnaires (6–8).

In accordance with the study assumption that diagnostic strategies should be equivalent, a cost minimization analysis will be performed to compare the FDG–PET based strategy to the conventional strategy. Savings are expected to come from requiring less laryngoscopies. If the clinical study does show a clinical difference between the strategies, this can be altered in a cost-effectiveness analysis. This will not influence the approach and the data collection of the cost analysis. The clinical study is aimed at investigating the impact of a diagnostic strategy within the hospital. Therefore, analysis will be focused on the hospital’s perspective, including only direct medical costs in the hospital.

The Dutch Head and Neck Oncology Cooperative Group has developed guidelines for laryngeal carcinoma in 1999 (20). Currently, there are no guidelines for the use of PET in head and neck cancer. Hopefully, this trial will be helpful in defining cost-effective use and the optimal application of this technique, when the guidelines for laryngeal carcinoma are updated by the Dutch Head and Neck Oncology Cooperative Group.
REFERENCES


