The role of TSH for $^{18}$F-FDG-PET in the diagnosis of recurrence and metastases of differentiated thyroid carcinoma with elevated thyroglobulin and negative scan: a meta-analysis

Chao Ma*, Jiawei Xie1*, Yanhui Lou1*, Yanyan Gao2, Shuyao Zuo and Xufu Wang

Department of Nuclear Medicine, Affiliated Hospital of Medical College Qingdao University, Qingdao 266003, Shandong Province, People’s Republic of China, 1Medical College Qingdao University, Qingdao 266021, People’s Republic of China and 2Department of Endocrinology, Affiliated Hospital of Medical College Qingdao University, Qingdao 266003, People’s Republic of China

(Correspondence should be addressed to C Ma; Email: mc_7419@hotmail.com)

*(C Ma, J Xie, Y Lou contributed equally to this work)

Abstract

Purpose: To establish the effects of TSH stimulation on the uptake of fluorine-18-labeled 2-fluoro-2-deoxy-D-glucose for differentiated thyroid carcinoma (DTC) with thyroglobulin-positive and scan negative metastases.

Materials and methods: We searched the MEDLINE, EMBASE and the Cochrane Library for prospective controlled trials using TSH stimulation as an intervention. The outcomes of positron emission tomography (PET)-positive lesions, tumor-to-background ratio, maximum standard uptake value of the detected lesions were extracted and synthesized, and patients with the altered clinical management were studied. A meta-analysis was carried out using the Review Manager software.

Results: Seven prospective controlled clinical trials with 168 patients were found. All studies had a low risk of bias. PET scans under TSH stimulation versus thyroid hormone suppression showed statistically significant differences in the number of patients with PET true-positive lesions (odds ratio (OR) 2.45, 95% confidence interval (CI) 1.23–4.90) and in the number of the PET-detected lesions (OR 4.92, 95% CI 2.70–8.95) and tumor-to-background ratios. PET scans taken under TSH stimulation altered clinical management in altogether 12/130 (9%) patients in five paired studies (OR 2.40, 95% CI 1.11–5.22).

Conclusion: The data indicate that TSH stimulation should be recommended for DTC patients undergoing PET scanning in these circumstances. However, further well-designed studies emphasizing on the clinical significance of altered management by PET under TSH stimulation are needed.

Introduction

Generally, thyroglobulin (Tg) and a radiiodine scan correlate well with each other in follow-up studies for differentiated thyroid carcinoma (DTC) after thyroid remnant ablation (1, 2). Undetectable Tg levels, with a negative radiiodine scan, suggest complete remission, whereas detectable or elevated Tg levels are often associated with the uptake of radioactive iodine in local or distant metastases. The scenario of thyroid cancer cells exhibiting no radiiodine uptake during radiiodine scintigraphy despite an elevated serum Tg level occurs in 10–15% of the patients in the follow-up of DTC patients. There are several possible explanations for this scenario (3–6). In current practice, radiiodine diagnostic scanning is less frequently used due to its inferior sensitivity, and has been supplanted by serum Tg and neck ultrasonography. Now, a new challenging scenario has emerged: that of the ultrasonography-negative, Tg-positive patients (6). Fluorine-18-labeled 2-fluoro-2-deoxy-D-glucose positron emission tomography ($^{18}$F-FDG-PET) is primarily used to detect DTC recurrence or metastases in the setting of raised Tg levels and scan-negative metastases, giving rise to a third new challenging scenario: the PET scan-negative, Tg-positive patients.

To maximize radiiodine uptake in malignant thyroid cells, it is necessary to elevate TSH levels by the withdrawal of thyroid hormone therapy (THW) or, more recently, to use recombinant human TSH (rhTSH). Do TSH stimulations done under THW or rhTSH before PET increase the detection of tumor sites? PET imaging done under TSH stimulation reveals more lesions in a greater number of patients with a significantly increased tumor-to-background ratio than paired scans done during TSH suppressive therapy (THS) (7–9). However, no significant scan outcomes were found in other studies (10, 11). Whether PET scanning
for the detection of Tg-positive and scan-negative metastases for DTC should be performed under TSH stimulation or suppression remains controversial. Therefore, there is a need to establish the effects of TSH stimulation on the uptake of FDG for DTC with Tg-positive and radioactive iodine-negative metastases.

Methods

Criteria for considering studies for this review

Patients of any age or sex who had elevated Tg levels and iodine-negative DTC after total or near-total thyroidectomy followed by radioactive iodine ablation and prospective controlled clinical trials comparing TSH stimulation and TSH suppression in any language were included. Patients with positive anti-Tg or uncontrolled diabetes mellitus were excluded.

We considered survival and the number of patients with PET true-positive lesions among patients with PET-detected recurrent and metastatic DTC and health-related quality of life as the primary outcomes; the number of the PET-detected lesions, tumor-to-background ratio, mean maximum standard uptake value (SUV_max) of the detected lesions, death from any cause, morbidity, adverse effects and costs were used as the secondary outcomes to evaluate the role of TSH in PET in the diagnosis of recurrence and metastases of DTC in patients with elevated Tg levels and negative scan.

Search methods for the identification of studies

We electronically searched The Cochrane Library (issue 2, 2009), MEDLINE (until August 2009) and EMBASE (until 2009) according to the Cochrane Search Strategy for the identification of trials using the words ‘TSH, FDG-PET, differentiated thyroid carcinoma, thyroglobulin, radiiodine, follow-up, rhTSH’. We also searched the databases of the ongoing trials: ‘Current Controlled Trials’ (www.controlled-trials.com – with links to other databases of ongoing trials). Additional key words of relevance could have been detected during any of the electronic or other searches. If this were the case, electronic search strategies would have been modified to incorporate these terms. We also tried to identify additional studies by searching the reference lists of included trials and (systematic) reviews, meta-analyses, and health technology assessment reports that were noticed.

Data collection and analysis

Data extraction and management To determine the studies to be assessed further, two authors independently scanned the abstract, title, or both sections of every record retrieved. All potentially relevant articles were investigated as full text. For studies that fulfilled the inclusion criteria, two authors independently abstracted relevant population and intervention characteristics using the standard data extraction templates with any disagreements to be resolved by discussion, or if required, by discussion with a third party. Possible disagreement was resolved by consensus, or through consultation with a third party. We explored the influence of individual bias criteria in a sensitivity analysis using the Cochrane Collaboration’s risk of bias tool (12). In cases of disagreement, the rest of the group was consulted, and a judgment was made based on the consensus.

Data synthesis and statistical analysis Data were summarized statistically if they were available, sufficiently similar and of sufficient quality. Statistical analysis was performed according to the statistical guidelines referenced in the newest version of the Cochrane Handbook for Systematic Reviews of Interventions (13). Odds ratios (ORs) were used for dichotomous data (the number of patients with FDG-PET-detected metastases and the number of the FDG-PET-detected lesions and other adverse effects). Mean differences were used for continuous data (mean SUV_max of the PET-detected lesions and quality of life).

In the event of substantial clinical, methodological, or statistical heterogeneity, study results were not combined by means of a meta-analysis. Heterogeneity was identified by visual inspection of the forest plots by using a standard χ² test and a significance level of α = 0.1 in view of the low power of such a test. Heterogeneity was specifically examined using I², where I² values of 50% and more indicate a substantial level of heterogeneity (14). When heterogeneity was found, we attempted to determine potential reasons for it by examining individual study and subgroup characteristics.

Figure 1 Flowchart of study selection.
Subgroup analysis and investigation of heterogeneity

Subgroup analysis was planned to be mainly performed if one of the primary outcome parameters demonstrated statistically significant differences between intervention and control groups. In any other case, subgroup analysis would have been clearly marked as a hypothesis-generating exercise. The following subgroup analysis was planned: PET scans taken under THW versus rhTSH.

Results

Results of the search

The electronic searches and hand searches identified 83 studies. Of these references, we excluded 70 citations based on their title, abstract, or both because they were not relevant to the question under study. After reading the full text of 13 potentially relevant publications, six were excluded because they were retrospective studies without controls and did not fulfill the inclusion criteria (7, 15–19). Seven prospective controlled clinical trials with 168 patients were included (8, 9, 11, 20–23), see Fig. 1.

Characteristics and quality of the included studies

Altogether 168 patients with elevated Tg levels and scan-negative DTC following thyroidectomy and thyroid remnant ablation participated in the seven trials. Summary data on age, gender, tumor pathology, and staging were reported for all the participants. No significant differences were found between comparison groups at baseline. The risk of bias of the seven included prospective controlled clinical trials was considered as low. The population characteristics and quality of the included trials are summarized in Tables 1 and 2 respectively.

Effects of TSH on the diagnostic performance of FDG-PET scanning in patients with elevated Tg levels and scan-negative DTC

The summary of the meta-analysis is given in Table 3.

Primary outcomes

PET scans under TSH stimulation (THW or rhTSH) vs THS showed statistically significant differences in the number of patients with PET true-positive lesions (seven studies (8, 9, 11, 20–23), OR 2.45, 95% confidence interval (CI) 1.23–4.90) as shown in Fig. 2. PET scans under TSH stimulation altered clinical management in altogether 12/130 (9%) patients in five paired studies (8, 9, 20, 21, 23) (OR, 2.40, 95% CI 1.11–5.22). No study analyzed survival and health-related quality of life.

Secondary outcomes

PET scans taken TSH stimulation (THW or rhTSH) versus THS showed statistically significant differences in the number of the PET-detected
lesions (four studies \(8, 9, 21, 23\)), OR 4.92, 95% CI 2.70–8.95) as shown in Fig. 3 and tumor-to-background ratio mean (OR 0.86, 95% CI 0.08–1.64). No significant difference was found in SUV\(_{\text{max}}\) of the detected lesions (four trials \(8, 21–23\), OR 0.02, 95% CI 0.45 to 0.41) during TSH stimulation and suppression as shown in Fig. 4. None of the trials investigated death from any cause, adverse effects, morbidity, or costs.

**Subgroup and sensitivity analyses** It was not possible to perform sensitivity analyses for age and gender due to the low number of studies.

**Discussion**

\(^{18}\text{F-FDG-PET}\) is primarily useful for the detection of DTC recurrence or metastases in the setting of raised Tg levels and radioactive iodine-negative metastases \(3, 24, 25\). In a recent meta-analysis, the pooled sensitivity and specificity for the DTC patients who presented with elevated serum Tg levels and negative iodine-131 scan were 0.885 (95% CI: 0.828–0.929) and 0.847 (95% CI: 0.715–0.934) respectively \(25\). In 2002, the Centers for Medicare and Medicaid Services (CMS) approved the use of \(^{18}\text{F-FDG-PET}\) in Tg-positive and radioiodine scan-negative patients with follicular cell-origin thyroid cancer \(17\). However, whether FDG-PET should be performed under TSH stimulation or not to detect Tg-positive and radioactive iodine-negative metastases of DTC remains controversial. We aimed to establish the effects of TSH stimulation on the diagnostic performance of PET in this clinical setting.

In our review, seven prospective controlled clinical trials \(8, 9, 11, 20–23\) involving 168 patients were included. Because of the low incidence, prolonged clinical course, and anticipated formidable sample sizes required for DTC, which has an overall excellent prognosis, it will be difficult to design and conduct a

### Table 2 Characteristics of the imaging factors and quality of the included studies. Each of the studies was carried out at a separate study center. They were free of selective reporting and no other bias was detected. Further allocation concealment for any of the studies was not mentioned. All these studies included diagnostic criteria of PET.

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Blinding</th>
<th>Verification of PET-detected lesions</th>
<th>PET examination time after rhTSH injection</th>
<th>(^{18}\text{F-FDG dose (MBq)} )</th>
<th>Time interval between two PET examinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8)</td>
<td>Paired</td>
<td>n</td>
<td>Biopsy</td>
<td>48 h</td>
<td>555</td>
<td>1 week</td>
</tr>
<tr>
<td>(9)</td>
<td>Paired</td>
<td>y</td>
<td>Iodine-131 post-therapy scan, CT, MRI, bone scan, and fine-needle biopsy</td>
<td>N/A</td>
<td>370</td>
<td>35 days</td>
</tr>
<tr>
<td>(11)</td>
<td>Prospective parallel</td>
<td>n</td>
<td>Multiple imaging, serological studies and biopsies</td>
<td>N/A</td>
<td>370</td>
<td>N/A</td>
</tr>
<tr>
<td>(20)</td>
<td>Paired</td>
<td>n</td>
<td>Iodine-131 post-therapy scan and CT</td>
<td>N/A</td>
<td>180 ± 10.5</td>
<td>Average 42 days (29–73)</td>
</tr>
<tr>
<td>(21)</td>
<td>Paired</td>
<td>n</td>
<td>Pathology</td>
<td>14/63, 24 h; 63, 48 h</td>
<td>222–888</td>
<td>5 days (3–13)</td>
</tr>
<tr>
<td>(22)</td>
<td>Prospective parallel</td>
<td>n</td>
<td>Multiple imaging including bone scan, CT, and ultrasonography</td>
<td>N/A</td>
<td>383.7 ± 47.4</td>
<td>N/A</td>
</tr>
<tr>
<td>(23)*</td>
<td>Paired</td>
<td>y</td>
<td>Histology, Iodine-131 post-therapy scan, CT, and MRI</td>
<td>48 h</td>
<td>4.5 MBq/kg body weight</td>
<td>&lt;1 month</td>
</tr>
</tbody>
</table>

\(^{18}\text{F-FDG-PET, fluorene-18-labeled 2-fluoro-2-deoxy-o-glucose positron emission tomography; y, yes; n, no; N/A, not applicable.}

*This was the only study reported in French. All other studies were reported in English.

### Table 3 Summary of data analyses.

<table>
<thead>
<tr>
<th>Outcome or subgroup</th>
<th>Studies</th>
<th>No. of patients with detected tumors</th>
<th>Statistical method</th>
<th>Effect estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>PET true-positive patients – TSH stimulation vs suppression</td>
<td>(8, 9, 11, 20–23)</td>
<td>127</td>
<td>OR (M-H, fixed, 95% CI)</td>
<td>2.45 (1.23, 4.90)</td>
</tr>
<tr>
<td>Detected lesions – TSH stimulation vs suppression</td>
<td>(8, 9, 21, 23)</td>
<td>223</td>
<td>OR (M-H, fixed, 95% CI)</td>
<td>4.92 (2.70, 8.95)</td>
</tr>
<tr>
<td>Tumor-to-background ratio – TSH stimulation vs suppression</td>
<td>(8, 20)</td>
<td>26</td>
<td>Mean difference (IV, fixed, 95% CI)</td>
<td>0.86 (0.08, 1.64)</td>
</tr>
<tr>
<td>Mean SUV(_{\text{max}}) – TSH stimulation vs suppression</td>
<td>(8, 21–23)</td>
<td>251</td>
<td>Mean difference (IV, fixed, 95% CI)</td>
<td>−0.02 (−0.45, 0.41)</td>
</tr>
</tbody>
</table>

THS, thyroid hormone suppression therapy; THW, thyroid hormone withdrawal; rhTSH, recombinant human TSH; PET, positron emission tomography; OR, odds ratio.
randomized trial on this issue with a large patient sample (6). The seven included prospective controlled clinical trials, in which age, gender, tumor pathology, and tumor staging were similar at baseline and no significant heterogeneity between trials was found, had a similar overall low risk of bias. The results showed the statistically significant patient-based and lesion-based differences of PET under TSH stimulation with enhanced tumor-to-background ratios compared with PET under TSH suppression. The clinical management was altered in altogether 12/130 (9%) patients who were imaged by PET scans taken under TSH stimulation versus TSH suppression in five paired studies (8, 9, 20, 21, 23). The limited data that is available suggest that TSH stimulation significantly improves the sensitivity of FDG-PET in the detection of DTC recurrence and metastases in patients with elevated Tg levels and iodine-negative disease. The data indicate that TSH stimulation should be recommended for DTC patients undergoing PET scanning in these circumstances.

The accuracy of OR estimate is in part determined by the sample size, and in general, the larger the sample size, the better. Five included studies had relatively small patient sample (8, 9, 11, 20, 22), which was compatible with the low incidence of the disease (26). The clinical significance of PET under TSH stimulation also depends on the knowledge of the range of possible treatments, costs, and side effects. No end points such as survival and mortality were reported in the included trials. The duration of the follow-up was 18 months (21), and this was not mentioned in three trials (11, 22, 23). It should be noted that one of the studies had a follow-up of only 10 days based on the collected data on iodine-131 post-therapy scan taken 10 days after treatment for the verification of the PET-detected lesions. The authors did not mention the examination time of computed tomography (CT), magnetic resonance imaging (MRI), bone scan, and fine-needle biopsy for the verification of the PET-detected lesions (9).

As we know, the FDG-PET scan is expensive and is not covered by most public health care systems in China. The cost-effectiveness and health-related quality of life should also be considered and analyzed in future studies. In addition, PET scans were performed under variable conditions. The dose of the administered FDG, time interval between injection and scanning, and time interval between the two PET scans varied between the studies. However, variation in the administered dose may affect counting statistics, but not the mean values. None of the studies described the reproducibility of PET. The calculation of the sample size was reported in only one trial (21). Under this circumstance, PET examinations done under TSH stimulation for DTC patients with elevated Tg levels and iodine-negative scan are not well established. Further well-designed studies emphasizing the clinical significance of altered management by PET under TSH stimulation are needed.

Increased uptake of FDG by DTC metastases can be expected under TSH stimulation, owing to the higher metabolic demand of the stimulated thyroid tissue. This was confirmed by the improved diagnostic performance of FDG-PET and the tumor-to-background ratios observed under TSH stimulation in our analysis. However, we could not explain why there was no significant difference in the mean SUV max of the

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>TSH stimulation Events</th>
<th>TSH suppression Events</th>
<th>Odds ratio</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8)</td>
<td>99</td>
<td>99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>(23)</td>
<td>97</td>
<td>95</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>(21)</td>
<td>102</td>
<td>108</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>(9)</td>
<td>11</td>
<td>11</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>223</td>
<td>223</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Total events</td>
<td>209</td>
<td>165</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Heterogeneity: $\chi^2 = 2.72, df = 3 (P = 0.44); I^2 = 0$

Test for overall effect: $Z = 5.21 (P < 0.00001)$
PET-detected lesions; this may have been due to the experimental error.

rhTSH contributes substantially to the diagnostic approach to thyroid cancer, offering an alternative to THW by avoiding the morbidity of hypothyroidism. Clinical studies have shown that administration of rhTSH promotes radiiodine uptake and Tg production by thyroid cells with an efficacy that is comparable to that of hypothyroidism for diagnosing residual or recurrent cancer (27–29). In North America and Europe, rhTSH is approved for use before Tg testing or diagnostic radioactive iodine scintigraphy in patients on thyroid hormone suppressive therapy (30). Currently, we cannot draw a conclusion about the effects of rhTSH-versus THW-stimulated PET scans on the detection of DTC recurrence and metastases in patients with elevated Tg levels and negative scan due to the unavailability of well-designed trials, except for only one retrospective study. In this retrospective study (19), the diagnostic performance of rhTSH stimulation versus THW in 15 patients showed no significant difference in tumor-positive cases displayed by PET, which indicates that rhTSH is as effective as THW in detecting DTC recurrence and metastases in patients with elevated Tg levels and negative scan.

We initially planned to carry out an analysis of sensitivity and specificity of pooled data. However, most studies did not supply enough valid data to enable the calculation of sensitivity and specificity. The accuracy of PET in the diagnosis of DTC recurrence and metastases in patients with elevated Tg levels and negative scan was confirmed partially by biopsy of accessible PET-detected lesions, surgical pathology, clinical follow-up, and response to treatment, or by comparison with conventional diagnostic methods. Not all the PET-detected lesions could be systematically verified by histology in all the included trials. Should negative PET examinations, which were also not confirmed in a valid way, in patients with elevated Tg levels and iodine-negative scan be considered false negative? This was neither feasible nor ethical because of the high number of lesions detected in some patients, and it is a well-known limitation precluding true sensitivity and specificity analyses (7, 21, 26).

With respect to the correlation of Tg levels and PET scan results, studies showed a higher specificity of PET in patients with higher Tg levels than in patients with lower Tg levels (10, 11), because serum marker levels and tumor load are positively related (26). However, PET/CT scanning performed under TSH stimulation is an effective method for the detection of DTC recurrence and direct surgical interventions, even in those with persistently elevated but relatively low Tg levels (13 and 14 µg/l) (19). A recent study suggested that the sensitivity of rhTSH-stimulated PET/CT is low in the diagnosis of recurrent disease in DTC patients with a low Tg level (10 µg/l). No correlation was observed between PET/CT findings and Tg levels (31). Therefore, correlation of Tg levels and PET scan results obtained during THW versus rhTSH also needs to be studied further.

Conclusions

The results obtained from the seven prospective controlled clinical trials suggest that TSH stimulation done under thyroid hormone withdrawal or rhTSH slightly and significantly improves the diagnostic performance of PET for the detection of Tg-positive and radiiodine-negative metastases of DTC. Therefore, TSH stimulation may be recommended for DTC patients undergoing PET scanning in these circumstances. However, the clinical significance remains uncertain due to the small sample sizes, short durations of follow-up, and the absence of systematic histological verification of all the PET-detected lesions in the included trials.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by the National Nature Science Foundation of China (No. 30700185).

Acknowledgements

We would like to acknowledge Dr S Leboulleux for the paper information offered. We are especially grateful to Dr Albert Driedger from the Department of Nuclear Medicine/Oncology, University of Western Ontario, London, Canada, and Dacia Richmond from Cross Cancer Institute, Edmonton, Canada, for their comments and advice on the review.