The Proportion of Malignancy in Incidental Thyroid Lesions on 18-FDG PET Study: A Systematic Review and Meta-analysis

Smriti Nayan, MD1, Jayant Ramakrishna, MD1, and Michael K. Gupta, MD, FRCSC1

Abstract

Objective. To evaluate through a systematic review and meta-analysis the malignancy rates of thyroid incidentalomas identified in adults by 18-fluorodeoxyglucose positron emission tomography, computed tomography (18-FDG PET-CT) imaging studies.

Data Sources. The literature search was conducted using OVID Medline, EMBASE, the Cochrane Library, Google Scholar, Pubmed, and reference list review (inception to April 2013) by 2 independent review authors.

Review Methods. Studies with adults undergoing 18-FDG PET scan identifying a thyroid incidentaloma with definitive histological or cytological results reported were included.

Results. Thirty-one studies with a total of 197,296 PET studies and 3659 focal thyroid incidentalomas were identified with 1341 having definitive cytopathology or histopathology. The pooled proportion of malignancy was calculated as 19.8% (95% confidence interval [CI], 15.3%-24.7%) with 15.4% (95% CI, 11.4%-20.0%) of the total cases being papillary thyroid cancer. Distant metastases represented 1.1% (95% CI, 0.6%-1.8%) of the total cases.

Conclusions. Our systematic review and meta-analysis suggests that the incidence of malignancy is high in thyroid incidentalomas identified through 18-FDG PET imaging studies. Thyroid incidentalomas identified through 18-FDG PET require thorough investigation.

Keywords

thyroid, incidental findings, positron emission tomography, papillary thyroid cancer, systematic review

Received December 15, 2013; revised February 17, 2014; accepted March 18, 2014.

Introduction

Whole body 18F-fluorodeoxyglucose (FDG) positron emission tomography (PET) has become an important diagnostic imaging study when assessing patients with malignancies.1,2 The standard imaging protocol in most centers for oncological 18-FDG PET-CT is a whole body scan from the midbrain to the midthighs.3 PET studies monitor the metabolism of FDG in the body. Malignant cells, being more metabolically active, have higher rates of glycolysis and glucose transport. Consequently, cancer tissue has an increased uptake of FDG when compared to normal tissues.2,4 These properties of 18-FDG allow PET-CT studies to aid in tumor staging, detecting metastases or recurrences, and monitoring treatment response.1

The increased use of whole body PET studies, especially whole body PET-CT studies, has resulted in an increase in the discovery of incidental lesions. These lesions are often termed incidentalomas.1 Reports in the literature indicate that on 18-FDG PET-CT, the overall prevalence of incidental, clinically significant lesions of malignant or premalignant pathology, ranges between 1.2% and 2.0%.2 The most common sites of incidental findings are the gastrointestinal tract, thyroid, and lung.2

Usually, the thyroid gland shows low-grade 18-FDG uptake.5 Thyroid incidentalomas on 18-FDG PET scans are of greater concern when characterized by focal uptake of FDG within the thyroid as compared to diffuse uptake, which is more indicative of benign pathology such as autoimmune thyroiditis.2 Published reports have estimated the prevalence of thyroid incidentalomas on 18-FDG PET between 1.2% and 4.3%.2,4,6 However, there is a lack of consensus on the proportion of malignancy in thyroid incidentalomas, with a wide range of reported rates.2,7,8

Physicians and surgeons specializing in the treatment of thyroid nodules must understand the clinical significance of thyroid incidentalomas identified on 18-FDG PET-CT.
scans. Defining the proportion of malignancy in thyroid incidentalomas is important when counseling patients, determining further investigations, and developing management strategies. The objective of this systematic review and meta-analysis is to examine the literature to determine the proportion of malignancy in thyroid incidentalomas identified through 18-FDG PET-CT studies in adult patients.

### Methods

This systematic review and meta-analysis was performed in accordance with a predetermined protocol consisting of eligibility criteria, a search strategy, and statistical analysis. The review conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines 2009 and to the guidelines for systematic review of observational studies from the Meta-Analyses of Observational Studies in Epidemiology (MOOSE) group.

Our primary aim was to perform a pooled analysis of the proportion of malignancy in thyroid incidentalomas.

### Literature Search Strategy

The literature search was conducted using OVID Medline (1950 through April 2013), EMBASE (1980 to April 2013), Pubmed (inception to April 2013), Google Scholar (inception to April 2013), and the Cochrane Library (Cochrane Database of Systematic Reviews, 2013 Issue 4) by 2 independent reviewers. Similar strategies were used when searching all the databases. Relevant articles and abstracts were selected and reference lists reviewed from these sources. Recent review articles or meta-analyses were searched for additional relevant publications.

The literature search of the electronic databases combined the following specific keywords in different combinations: thyroid mass, benign, malignant, fine needle aspiration, biopsy [FNAB], surgery, PET scan, 18-FDG, positron emission tomography, incidental, incidentaloma, fluorodeoxyglucose. The following study designs and publication types were included: retrospective studies, prospective cohort studies, randomized controlled trials, systematic reviews, and meta-analyses. To ensure all relevant published articles were captured, the literature search was not limited by study design or publication date.

### Study Selection Criteria

The three authors reviewed the studies using the search strategy to identify articles for inclusion. Disagreements on article inclusion were resolved through discussion and consensus. Reasons for exclusion were noted for each article omitted from this study. Articles were included in the systematic review if they were fully published reports or abstracts of studies evaluating adult patients (>18 years of age) undergoing PET scan with a focal thyroid incidentaloma identified at the time of the scan with no prior history of thyroid disease. Included trials reported on the diagnostic investigation tests undertaken as well as the final histopathological or cytopathological result of the thyroid incidentaloma under investigation.

Articles were excluded if they were non-human trials, were published in a language other than English, were case reports (n < 10), did not use PET imaging studies to initially identify the incidentaloma, did not characterize the thyroid mass through diagnostic investigation, did not provide a definitive pathology result, or included patients less than 18 years of age.

### Data Extraction

Data were extracted from fully published reports by the 3 review authors using prescribed tables. Data collected included total number of focal thyroid incidentalomas, number of malignant lesions (both by cytology and histopathology), number of benign lesions (by cytology and histopathology), number of indeterminate lesions. Further data were collected on the breakdown of malignant incidentalomas. Only data on focal thyroid incidentalomas were analyzed, as the risk of malignancy is higher with focal uptake in the thyroid as compared to diffuse uptake in the thyroid.

Trial setting, mean or median standard uptake value (SUV) values, patient characteristics, and demographics were all collected for secondary analysis. Any disagreement was resolved through discussion and consensus.

Cytopathology or histopathology results were included for focal thyroid incidentalomas based on definitive results on fine needle aspiration biopsy or surgery.

To be counted as either benign or malignant in the analysis, cases had to have either a definitive cytopathology result on fine needle aspiration biopsy or a histopathology result from surgery. Definitive cytopathology was considered to be all lesions except for the following 5 reported results: indeterminate, atypia of undetermined significance, nondiagnostic, follicular neoplasms, or suspicious for follicular neoplasm. If definitive histopathology was not available for the latter 5 cytopathology results, then these lesions were not counted as either benign or malignant.

### Quality Assessment

The assessment of the quality of the individual studies is important for any systematic review. Two independent review authors performed quality assessment based on the Methodological Index for Non-Randomized Studies (MINORS) Scale for Non-Randomized Studies and the Oxford Level of Evidence. MINORS is a validated instrument designed to assess the methodological quality of nonrandomized. The scale was developed by surgeons in light of the large number of observational studies in surgery. The scale has high external validity and is able to identify high-quality studies, making it comparable to the CONSORT standards for randomized trials. Any disagreement between the reviewers with regards to the quality assessment of the studies was resolved through consensus.

### Statistical Analysis

Using the extracted data, 2 types of proportions were calculated. One type of proportion used the number of all focal thyroid incidentalomas as the denominator. A second type...
used all focal incidentalomas with definitive cytopathology or histopathology as the denominator. The decision was made to report both proportions as it was felt that neither on its own was free of bias, as typically a large number of focal thyroid incidentalomas do not undergo complete investigation to render a final diagnosis. Completely excluding this large group of incidentalomas from the analysis adds significant bias to the analysis. It was felt that the true proportion likely is found between the 2 reported proportions.

The random effects model of pooled proportions was calculated using the DerSimonian-Laird method. All statistical calculations were done using the StatsDirect software (Cheshire, UK). A random effects model was used as the heterogeneity of the data was large. Confidence intervals of the pooled proportions were calculated using the Greenland-Robin variance formula.

Subgroup analyses were performed where possible to determine possible sources of heterogeneity in the data. A subgroup analysis was performed on papers based on methodological quality. The papers were divided into groups based on their MINORS score and pooled proportions across groups were compared. A subgroup analysis on sample size and proportion of incidentalomas with definitive diagnosis was also performed.

For the subgroup analyses, the proportions calculated used the denominator of all incidentalomas.

Results

Literature Search

Two thousand one hundred and twenty-five studies were identified initially through the search strategy. These studies were than narrowed down to 63 studies. Thirty-one studies were included for final analysis (Figure 1), and 32 studies were excluded as they did not meet the inclusion criteria. All the studies identified were retrospective, noncomparative cohort studies.

Trial Characteristics

All published studies involved a single center and were conducted between 1996 and 2013. A total of 197,296 PET studies were described. Three thousand six hundred and fifty-nine focal thyroid incidentalomas were identified, and 1340 (36.6%) of these had definitive cytopathology or histopathology reported (Supplemental Table S1 at www.otojournal.org; and Table 1). A total of 2319 (63.4%) of the incidentalomas did not have final cytopathology or histopathology reported. Various reasons were stated in the individual studies, including patients having more serious diseases that required treatment or patient refusal to pursue further investigation and thus was excluded from analysis.

Figure 1

The 18-FDG PET-CT studies were conducted as part of oncological workups as screening studies in healthy subjects, or were a mixed population of healthy patients and oncology patients.

There were more benign than malignant lesions, 861 and 479, respectively. The most common type of malignant lesion was papillary thyroid cancer with a total of 391 identified through fine needle aspiration biopsy. The second most common cause of malignancy was metastatic disease from a primary other than thyroid (n = 28) and follicular thyroid carcinomas (n = 28). The least common malignancy was anaplastic thyroid cancer with only 3 reported cases. There were 13 non-thyroid malignancies such as lymphoma, 9 Hurthle cell carcinomas, and 8 medullary thyroid carcinomas. The results are summarized in Table 1.

Quality Assessment

The studies were rated using the MINORS scale, and the scores ranged between 10 and 16 (Supplemental Table S2 at www.otojournal.org). The mean rating for all of the studies on the MINORS scale was 14. All the studies were rated as 2 on the Oxford Level of Evidence as they were all retrospective noncomparative cohort studies, which had the primary aim of evaluating a diagnostic tool.

Proportion of Malignancy of the Thyroid Incidentalomas

Proportion of all incidentalomas. As mentioned in the methods, 2 different proportions were calculated. The first used all incidentalomas as the denominator. A random effects model was used in all calculations due to high degrees of heterogeneity. The calculated pooled proportion of malignancy in the thyroid incidentalomas was 20% (95% CI, 15%-25%, I² = 89%) (Figure 2). The pooled proportion of papillary thyroid cancer was 15% (95% CI, 11.4%-20.0%, I² = 88.9%) (Figure 3), and the pooled proportion of metastases to the thyroid from another primary was calculated to be 1% (95% CI, 1%-2%, I² = 43.3%) (Figure 4).

Proportion of diagnosed incidentalomas. The second proportion calculated used only incidentalomas with a definitive diagnosis as the denominator. Again a random effects model was used. The pooled proportion of malignant incidentalomas was calculated to be 37% (95% CI, 33%-41%, I² = 54.0%) (Figure 5). The pooled proportion of papillary thyroid cancer was calculated to be 29% (95% CI, 25%-34%, I² = 59.1%) (Figure 6). The pooled proportion of metastatic incidentalomas was 2% (95% CI, 1%-4%, I² = 33.6%) (Figure 7).

Subgroup analysis. A subgroup analysis demonstrated that the proportion of malignancy in studies with only oncology patients was not significantly different than the proportion of malignancy in studies with only healthy patients undergoing PET imaging as screening. The pooled proportions of malignancy in the healthy group of patients was 30% (95% CI, 4%-68%) (Figure 8) and in the oncology group of patients, 14% (95% CI, 9% to 18%, I² = 86.9%) (Figure 9).

A subgroup analysis based on MINORS scores was also performed. All the papers scored either a 10, 12, 13, 14, 15, or 16 on the MINORS assessment. There was only a single paper in the group scoring 10 and the group scoring 12 so the results of these groups could not be pooled. There was
no significant difference in pooled proportion of malignancy across the groups that could be analyzed as confidence intervals overlapped between all groups (MINORS score 13: 19%, 95% CI, 11%-27%; MINORS score 14: 17%, 95% CI, 9%-27%; MINORS score 15: 25%, 95% CI, 14%-39%; MINORS score 16: 13%, 95% CI, 5%-24%).

We also performed a subgroup analysis based on sample size. We compared the pooled proportion of malignancy between those studies including greater than 100 thyroid incidentalomas and those papers with less than 100 thyroid incidentalomas. We found a significant difference between the 2 groups, with the confidence intervals not crossing. The pooled proportion of malignancy in the group with greater than 100 incidentalomas was 11% (95% CI, 7%-15%, I² = 97%). The pooled proportion of malignancy in the group with less than 100 incidentalomas was 49% (95% CI, 40%-58%, I² = 79.8%).

We performed a final subgroup analysis based on the proportion of all incidentalomas in each study with a definitive diagnosis. The pooled proportion of malignancy was compared between studies with greater than 50% of incidentalomas with definitive diagnosis and that of studies with less than 50% of incidentalomas with a definitive diagnosis. Again there was a significant difference, with the confidence

---

**Table 1. Summary Table.**

<table>
<thead>
<tr>
<th>Number of Incidentalomas (%) of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of PET scans</td>
</tr>
<tr>
<td>Number of focal thyroid incidentalomas</td>
</tr>
<tr>
<td>FNA or surgical histopathology available</td>
</tr>
<tr>
<td>Number of focal thyroid incidentalomas with no FNA or surgical histopathology</td>
</tr>
<tr>
<td>Benign</td>
</tr>
<tr>
<td>Malignant</td>
</tr>
<tr>
<td>Indeterminate</td>
</tr>
<tr>
<td>Papillary thyroid cancer</td>
</tr>
<tr>
<td>Follicular thyroid cancer</td>
</tr>
<tr>
<td>Hurthle cell carcinoma</td>
</tr>
<tr>
<td>Medullary thyroid cancer</td>
</tr>
<tr>
<td>Anaplastic thyroid cancer</td>
</tr>
<tr>
<td>Mets</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Abbreviations: PET, positron emission tomography; FNA, fine needle aspiration.
intervals not crossing. Studies with greater than 50% of incidentalomas with a definitive diagnosis had a pooled proportion of malignancy calculated at 55% (95% CI, 49%-62%, $I^2 = 63.4\%$). The studies with less than 50% of incidentalomas with a definitive diagnosis had a pooled proportion of malignancy of 11% (95% CI, 7%-14%, $I^2 = 84.8\%$).

**Discussion**

The increased use of 18-FDG PET-CT studies has, not surprisingly, caused a rise in the number of thyroid incidentalomas identified. Determining the proportion of malignant lesions in thyroid incidentalomas is an important clinical question, which lacks consensus in the literature.

Our review found the proportion of malignant focal incidentalomas of the thyroid to be 20%. The majority of malignant lesions were papillary thyroid cancer. The proportion of thyroid incidentalomas being papillary thyroid carcinoma was 15%. Interestingly, subgroup analysis found that if a definitive diagnosis is vigorously pursued this proportion is higher. In studies that had a greater than 50% of included lesions with a definitive diagnosis the pooled proportion of malignancy was calculated to be 55% and was significantly higher than those studies with less than 50% of lesions with a definitive diagnosis. Consequently, these incidentalomas warrant further clinical investigations to definitively rule out the possibility of malignancy.

The chance of a focal hypermetabolic lesion in the thyroid representing distant spread from the primary was small at 1.1%. As a consequence, the investigation of focal lesions in the thyroid should not delay or alter treatment of primary malignancy elsewhere in the body. The proportion of malignancy in incidentally discovered thyroid lesions on 18F-FDG PET-CT in healthy patients without a known malignancy was higher in these patients compared to those who had a known malignancy, 30% versus 14%, respectively. However, this difference did not reach statistical significance as the confidence intervals overlap (Figures 8 and 9). Interestingly, 2 studies in our review used 18-FDG PET-CT as a screening study in healthy patients, which is not routine practice in North America.

In all cases, the benign lesions identified in these studies were diagnosed on the basis of FNAB. Fine needle aspiration biopsy is not the gold standard of diagnosis. However, FNABs have high sensitivity and specificity as reported by many studies. The systematic review reflects current

Figure 2. The pooled proportions of malignancy in thyroid incidentalomas.

Pooled proportion = 20% (95% CI, 15%-25%), $I^2 = 88.7\%$.

Figure 3. The pooled proportions that are papillary thyroid carcinoma.

Pooled proportion = 15% (95% CI, 11%-20%), $I^2 = 88.9\%$. 
clinical practice in most centers in that FNABs described as benign did not undergo further surgery to obtain a histopathological diagnosis.

18-FDG PET studies assess lesions by evaluating the maximum standard uptake value (SUVmax). Studies in the past have evaluated SUVmax values to help differentiate if a lesion is benign or malignant. Generally, malignant lesions have a higher SUVmax. Yet, Hurthle cell adenomas and follicular adenomas also have high SUVmax. Thus, controversy remains surrounding the relevance of SUVmax values as there can be considerate overlap between benign and malignant lesions. For this reason, we did not evaluate SUVmax values in our systematic review to help differentiate between benign and malignant lesions, but rather focused on and recommend a definitive pathology result to aid with diagnosis.

The primary limitation of this systematic review was that the majority of focal thyroid incidentalomas did not have a definitive cytopathological or histopathological diagnosis, limiting the identification of the true proportion of malignancy. For this reason, we did not evaluate SUVmax values in our systematic review to help differentiate between benign and malignant lesions, but rather focused on and recommend a definitive pathology result to aid with diagnosis.

The primary limitation of this systematic review was that the majority of focal thyroid incidentalomas did not have a definitive cytopathological or histopathological diagnosis, limiting the identification of the true proportion of malignancy. Frequently, patients with an indeterminate, atypia of undetermined significance, follicular neoplasm or suspicious for follicular neoplasm, and non-diagnostic cytopathology results were not investigated further to help identify the definitive pathology.

These cases were treated by calculating 2 proportions. One proportion included these incompletely diagnosed incidentalomas in the denominator; another excluded them entirely. We recognized that both proportions were subject to bias and possibly confounding. Certainly a strong argument can be made to exclude these cases entirely. However, we felt the proportion that included these undiagnosed cases was pragmatic and represented what was seen in day-to-day clinical practice and was interesting information to present.

A further limitation of this systematic review is the heterogeneity found during systematic analysis. Subgroup analysis found that a significant source of this heterogeneity was the workup of incidentalomas. Studies that vigorously investigated incidentalomas to render a final diagnosis had a significantly different proportion of malignancy than those that did not. We also found that papers with sample size greater than 100 were found to have a lower pooled proportion of malignancy than those studies with less than 100
incidentalomas. It is possible the difference in pooled proportion was due to differences in the threshold to diagnose an incidentaloma; however, there was no way to analyze that with the existing data. Methodological quality was not found to cause significant differences on subgroup testing.

The global ideal score on the MINORS scale averaged 14 for all the studies together, indicating that the methodological quality of the studies was not strong. All the studies were rated as 2 on the Oxford Level of Evidence Scale.

Lastly, the proportion of malignancy of thyroid nodules in this study was, at 19.8%, higher than the standard proportion of around 5% for all thyroid nodules. However, this difference should be interpreted with caution. This review does not evaluate the ability of 18-FDG PET-CT to discern the malignancy of thyroid nodules. As some authors have mentioned, using these studies to extrapolate 18-FDG PET-CT to determine malignancy of thyroid nodules introduces significant verification bias. This systematic review is not afflicted by verification bias as the review is not concerned with the ability of 18-FDG PET to discern malignancy in thyroid nodules; rather the study aims to provide information to clinicians that can be useful in managing patients who have already had 18-FDG PET-CT studies showing focal thyroid incidentalomas. The ability of 18-FDG PET-CT to diagnose malignancy in known thyroid nodules is a

---

**Figure 6.** The pooled proportion of papillary thyroid cancer of all definitively diagnosed incidentalomas.

Pooled proportion = 29% (95% CI, 25%-34%), I²=59.1%.

**Figure 7.** The pooled proportion of distant metastases of all definitively diagnosed incidentalomas.

Pooled proportion = 2% (95% CI, 1%-4%) I²=33.6%.

**Figure 8.** The pooled proportions of malignancy in healthy subjects.

Pooled proportion = 30% (95% CI, 4%-68%).
different question than that addressed in this review and requires careful study design to avoid verification bias.

**Conclusion**

Our systematic review suggests that the proportion of malignancy is high in focal thyroid incidentalomas as identified through 18-FDG PET imaging studies. Subgroup analysis suggests that thorough investigation by clinicians to render a final diagnosis in these lesions is very important.

**Author Contributions**

Smriti Nayan, substantial contributions to conception and design, acquisition of data, analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; final approval of the version to be published; Jayant Ramakrishna, acquisition of data, revising it critically for important intellectual content, final approval of the version to be published; Michael K. Gupta, substantial contributions to conception and design, acquisition of data, analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; final approval of the version to be published.

**Disclosures**

**Competing interests:** None.

**Sponsorships:** None.

**Funding source:** None.

**Supplemental Material**

Additional supporting information may be found at www.otojournal.org/supplemental.

**References**


80. Kim SK, Hwang TS, Yoo YB, et al. Surgical results of thyroid nodules according to a management guideline based on the

