A Multiobserver Comparison of $^{99m}$TcO$_4$ and $^{123}$I Thyroid Imaging

Brahm Shapiro, Keith Britton, Alex Fountos, Maria Granowska, Laurie Hawkins, Penny Kiriaki, Saulo Londres, Cyril Nimmon, and Sophie Poston

Department of Nuclear Medicine and Radio-isotopes, St. Bartholomew's Hospital, London EC1, U.K.

Abstract. The thyroids of forty patients were imaged using 2 mCi (74 MBq) $^{99m}$Tc pertechnetate ($^{99m}$TcO$_4$) followed within one week by 2 mCi (74 MBq) $^{123}$I Iodide. The images obtained were evaluated by eight observers for 6 morphological criteria and assigned to 6 diagnostic categories with a confidence grading on a seven level scale (grade 1 being that for maximum confidence). Images were obtained with $^{123}$I for the same counts and the same times as those with $^{99m}$TcO$_4$ and compared using an index in which the number of diagnostic categories and the mean confidence grading within each category were taken into account. The mean index obtained for $^{99m}$TcO$_4$ images (9.2) was significantly greater ($P<0.05$, thus representing a lower observer confidence and less interobserver diagnostic agreement) than the mean indexes for $^{123}$I equal count images (5.6) and $^{123}$I equal time images (6.8). The diagnoses made on the basis of $^{123}$I images were more frequently concordant with the final diagnosis after six months follow up (75% for equal counts and 77% for equal time) than those with $^{99m}$TcO$_4$ (63%). The radioisotope of iodine most suited to modern nuclear medicine instrumentation and with the most favourable radiation dosimetry is $^{123}$I which is recommended for those areas of thyroid imaging where iodine is superior.

Introduction

Radioisotopes of iodine are theoretically preferable to pertechnetate for thyroid imaging in that they trace the entire metabolic pathway of thyroid hormonogenesis rather than the trapping mechanism alone. In spite of this, adequate thyroid images are obtainable using $^{99m}$TcO$_4$ which is very widely available at low cost (Dos Remedios, Weber and Jasko 1971). Of the radioisotopes of iodine available the most satisfactory for gamma camera thyroid imaging is $^{123}$I which is recommended for those areas of thyroid imaging where iodine is superior.

Materials and Methods

The $^{123}$I in the form of Na$^{123}$I was obtained from irradiation of $^{127}$I-diodomethane with 58 MeV protons from the Harwell variable energy cyclotron, the reaction being $^{127}$I (P, 5 n) $^{123}$Xe $\rightarrow ^{123}$I. This yielded a preparation free of $^{124}$I and having only 0.13% contamination with $^{124}$I (Cuninghame et al. 1978). Sodium pertechnetate was eluted from a standard $^{99}$Mo-$^{99m}$Tc generator (Union Carbide).

Forty consecutive patients who were referred for thyroid imaging and able to undertake two visits, were studied using on Ohio Nuclear series 100 gamma camera with a 3 mm pinhole collimator. A single anterior view with a skin to aperture distance of 7 cm was obtained. The image was recorded on Kodak NMB nuclear medicine film using the same intensity settings for each study. Each patient received 2 mCi (74 MBq) $^{99m}$TcO$_4$ by intravenous injection and imaging was commenced 20 min later. Pertechnetate images were obtained to a total of 200,000 counts or an imaging time of 1,200 s, which ever occurred first. Within one week the patients were again studied, this time using 2 mCi (74 MBq) $^{123}$I by intravenous injection and imaging commenced 20 min later. Images of $^{123}$I uptake were obtained with equal counts and over equal time as those obtained with $^{99m}$TcO$_4$. Seven cases were also imaged again in the same manner 2 h after the administration.
of $^{123}$I. No patients with retrosternal goitre appeared in this consecutive study.

All images were assigned a random number and presented without clinical or technical data to 8 trained observers who examined them for 6 morphological criteria, namely, the number of "cold" areas, number of "hot" areas, the presence of a pyramidal lobe, the presence of generalised patchy uptake, overall gland size (increased, normal or decreased) and thyroid tracer uptake in relation to the thyroid environment (increased, normal or decreased).

On the basis of these morphological criteria the observers chose one or more of 6 diagnostic categories as the most probable, namely, normal, single "cold" nodule and its site, single "hot" nodule and its site, multinodular goitre, appearances consistent with simple goitre or consistent with thyroiditis. The diagnosis was assigned a confidence grading on a seven level scale on which a grading of 1 represented maximum certainty and a grading of 7 maximal uncertainty to allow for the evident overlap between the diagnostic categories. All observers took part a training session in which the uniform descriptive and diagnostic nomenclature and confidence grading were agreed with reference to a series of typical thyroid scans on cases with known diagnosis and outcome. The test images were examined in batches to avoid observer fatigue.

An index was used to compare the findings of the $^{99m}$TcO$_4$ and $^{123}$I studies on the same patient. The index took into account: the number of diagnoses chosen by the panel of observers and the mean confidence levels assigned to each diagnostic category. (Index = sum of mean confidence levels of all the diagnostic categories chosen). Such an index rises rapidly with interobserver disagreement in diagnosis and with observer uncertainty of diagnosis.

Diagnoses made from the thyroid images examined in this way were compared with the final diagnosis made after 6 months follow up. This was based on the clinical assessment and natural history over 6 months, serum thyroxine, triiodothyronine, thyrotrophin and free thyroxine index and the presence of antibodies to thyroid components. In some cases histology and ultrasound were examined in batches to avoid observer fatigue. All images were assigned a random number and presented to 8 trained observers who examined them for 6 morphological criteria, namely, the number of diagnoses chosen by the panel of observers and the mean confidence levels assigned to each diagnostic category. (Index = sum of mean confidence levels of all the diagnostic categories chosen). Such an index rises rapidly with interobserver disagreement in diagnosis and with observer uncertainty of diagnosis.

Results

The mean index for $^{99m}$TcO$_4$ studies was 9.2 (range 1.1-23.5) that for $^{123}$I to equal counts was 5.6 (range 1.0-24.8) and that for $^{123}$I for equal time was 6.8

(range 1.25-24.8). The differences between the $^{99m}$TcO$_4$ and $^{123}$I studies were significant at the 5% level using Wilcoxon's signed rank test for paired samples. There was no significant difference between the mean indices for $^{123}$I images taken for equal counts or equal time as the $^{99m}$TcO$_4$ images, nor between $^{123}$I images obtained at 20 min and 2 h post injection.

The diagnoses made from $^{99m}$TcO$_4$ images was concordant with the final diagnosis in 63%. The $^{123}$I images containing equal counts were concordant with the final diagnoses in 75%, and 77% using $^{123}$I to the same exposure time as with $^{99m}$TcO$_4$. In the clinical setting the diagnoses based on $^{99m}$TcO$_4$ imaging was concordant with the final diagnosis in 69% and that based on $^{123}$I imaging (equal time and equal counts as the $^{99m}$TcO$_4$ images being examined together) concordant in 79%. The distribution of diagnoses and mean confidence levels are shown in Table 1.

The pyramidal lobe was visible more frequently on $^{123}$I images (19%) than on $^{99m}$TcO$_4$ images (11%) and in only one instance was a pyramidal lobe demonstrated by $^{99m}$TcO$_4$ which was also not demonstrated by $^{123}$I. Radiation dosimetry for $^{123}$I was based on the clinical setting the diagnoses based on $^{99m}$TcO$_4$ imaging was concordant with the final diagnosis in 69% and that based on $^{123}$I imaging (equal time and equal counts as the $^{99m}$TcO$_4$ images being examined together) concordant in 79%. The distribution of diagnoses and mean confidence levels are shown in Table 1.

The pyramidal lobe was visible more frequently on $^{123}$I images (19%) than on $^{99m}$TcO$_4$ images (11%) and in only one instance was a pyramidal lobe demonstrated by $^{99m}$TcO$_4$ which was also not demonstrated by $^{123}$I. Radiation dosimetry for $^{123}$I was based on the clinical setting the diagnoses based on $^{99m}$TcO$_4$ imaging was concordant with the final diagnosis in 69% and that based on $^{123}$I imaging (equal time and equal counts as the $^{99m}$TcO$_4$ images being examined together) concordant in 79%. The distribution of diagnoses and mean confidence levels are shown in Table 1.

Table 1. Distribution of diagnostic categories and mean confidence levels

<table>
<thead>
<tr>
<th>Diagnostic category</th>
<th>$^{99m}$TcO$_4$</th>
<th>$^{123}$I (equal counts)</th>
<th>$^{123}$I (equal time)</th>
<th>Final diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>Mean confidence</td>
<td>No.</td>
<td>Mean confidence</td>
</tr>
<tr>
<td>Normal</td>
<td>40</td>
<td>3.1</td>
<td>58</td>
<td>2.2</td>
</tr>
<tr>
<td>&quot;Cold&quot; nodule</td>
<td>68</td>
<td>2.7</td>
<td>91</td>
<td>1.7</td>
</tr>
<tr>
<td>&quot;Hot&quot; nodule</td>
<td>8</td>
<td>1.3</td>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td>Multinodular</td>
<td>71</td>
<td>3.5</td>
<td>99</td>
<td>2.4</td>
</tr>
<tr>
<td>Thyroiditis</td>
<td>72</td>
<td>4.1</td>
<td>32</td>
<td>3.2</td>
</tr>
<tr>
<td>Simple goitre</td>
<td>61</td>
<td>4.3</td>
<td>38</td>
<td>3.3</td>
</tr>
</tbody>
</table>

No = number of observations

The lower the value, the greater the confidence

Number of cases multiplied by number of observers