Post-operative neck ultrasound and risk stratification in differentiated thyroid cancer patients with initial lymph node involvement

Charlotte Lepoutre-Lussey¹,*, Dina Maddah¹,*, Jean-Louis Golmard², Gilles Russ¹, Frédérique Tissier³, Christophe Trésallet⁴, Fabrice Menegaux⁴, André Aurengo⁴ and Laurence Leenhardt¹

Departments of ¹Nuclear Medicine, ²Biostatistics, ³Pathology and ⁴Endocrine Surgery, Institute of Cancer, Pitie-Salpetriere Hospital, Pierre et Marie Curie University, 47-83 Bd de l’hôpital, 75013 Paris, France
*(C Lepoutre-Lussey and D Maddah contributed equally to this work)

Correspondence should be addressed to L Leenhardt
Email laurence.leenhardt@psl.aphp.fr

Abstract

Objective: Cervical ultrasound (US) scan is a key tool for detecting metastatic lymph nodes (N1) in patients with papillary thyroid cancer (PTC). N1-PTC patients are stratified as intermediate-risk and high-risk (HR) patients, according to the American Thyroid Association (ATA) and European Thyroid Association (ETA) respectively. The aim of this study was to assess the value of post-operative cervical US (POCUS) in local persistent disease (PD) diagnosis and in the reassessment of risk stratification in N1-PTC patients.

Design: Retrospective cohort study.

Methods: Between 1997 and 2010, 638 N1-PTC consecutive patients underwent a systematic POCUS. Sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) of POCUS for the detection of PD were evaluated and a risk reassessment using cumulative incidence functions was carried out.

Results: After a median follow-up of 41.6 months, local recurrence occurred in 138 patients (21.6%), of which 121 were considered to have PD. Sensitivity, specificity, NPV, and PPV of POCUS for the detection of the 121 PD were 82.6, 87.4, 95.6, and 60.6% respectively. Cumulative incidence of recurrence at 5 years was estimated at 26% in ETA HR patients, 17% in ATA intermediate-risk patients, and 35% in ATA HR patients respectively. This risk fell to 9, 8, and 11% in the above three groups when the POCUS result was normal and to <6% when it was combined with thyroglobulin results at ablation.

Conclusion: POCUS is useful for detecting PD in N1-PTC patients and for stratifying individual recurrence risk. Its high NPV could allow clinicians to tailor follow-up recommendations to individual needs.

Introduction

In patients with papillary thyroid cancer (PTC), initial lymph node (LN) involvement (N1) is frequent (50–70% of cases) (1). Although the disease-specific survival rate is good, exceeding 90% at 10 years (1, 2), the outcome for N1-PTC patients is variable, with a risk of locoregional recurrence and persistent disease (PD) varying from 2 to 40% (3, 4, 5). The risk of recurrence of N1-PTC is classified to be intermediate in the American Thyroid Association (ATA) guidelines (6) and to be high by the European Thyroid Association (ETA) (7) and French (SFE) societies (8). This risk stratification is based on the pathological tumor node metastasis (pTNM) classification. Recent reports have proposed individual risk stratification based on data obtained 8–12 months after RAI (5, 9), while others outline the utility of a post-operative staging using stimulated thyroglobulin (s-Tg) tests to rule out PD after...
surgery (10, 11, 12). Neck ultrasonography (US) is the most sensitive method for detecting cervical LN metastasis (1, 13) and its usefulness just before surgery and 6–12 months after radioiodine ablation (RAI) is well established (6). This paper presents further evidence regarding the interest of post-operative cervical US (POCUS) in the management of these patients as recently highlighted (14). The objective of this retrospective study was to assess the value of POCUS in the detection of PD in N1-PTC patients, and to compare its value with postoperative s-Tg (RAI-Tg) and post-therapeutic whole-body scans (tWBS). Our hypothesis is that POCUS could be useful in redefining the initial recurrence risk based on the initial pTNM staging and thus enable clinicians to tailor RAI and follow-up recommendations to individual needs.

**Subjects and methods**

**Patients**

Between 1997 and 2010, 731 N1-PTC patients were consecutively referred to our center for RAI after thyroidectomy, and systematically underwent a POCUS examination. Their medical records were retrospectively reviewed.

Ninety-three patients were excluded: 75 without follow-up in our institution, eight without examination during the first 6 months after radioiodine therapy, four with no primary tumor discovered (Tx), and six with unavailable data. Adequate clinico-pathological and follow-up records of the remaining 638 patients were analyzed. The clinical characteristics of the patients are given in Table 1.

**Initial surgery, pTNM staging, and risk stratification**

Among our patients, 324 (50.6%) had undergone initial thyroidectomy at our institution (C Trésallet and F Menegaux) and 314 at other centers, and were then referred to our department. All tumors were classified using the 2010 pTNM scoring system (Table 1). The surgical procedure was guided by clinical and imaging findings, including at least a systematic preoperative cervical ultrasound. All patients had total thyroidectomy and 97% of them had LN dissection (LND). The extent of the LND took into account the clinical and preoperative ultrasound (US) data and remained at the discretion of the attending surgeon.

All patients had LN involvement, discovered either at systematic LND or fortuitously at histological examination in 97 and 3% of cases respectively (Table 1), and were thus considered to be at a high risk (HR) of recurrence as stated by the ETA consensus (7). According to the ATA guidelines, patients were considered to be at a HR of recurrence (macroscopic tumor invasion and/or incomplete tumor resection and/or distant metastases) in 49.4% of the cases (315/638 patients) or an intermediate risk (N1 without the above-mentioned criteria) in 50.6% of the cases (323/638 patients). Because of the retrospective nature of the study and the many changes in pTNM classification, we could not isolate the subgroup of microscopic invasion of tumor into the perithyroidal soft tissues at initial surgery. Consequently, we decided by default to classify all pT3 patients as ATA HR.

The location of the metastatic LN and the presence of an extra-capsular LN involvement (N1-ECI) were recorded when available. We studied the relevance of dividing the patients into two subgroups of LN involvement (LN score): a low-risk (LR) subgroup (LR-N1) including N1a patients without N1-ECI vs a HR subgroup (HR-N1) represented by

---

**Table 1  Characteristics of patients (n = 638).**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (median (range))</td>
<td>42.2 (11.3–82.2)</td>
</tr>
<tr>
<td>Sex (females/males (%))</td>
<td>482/156 (75.5/24.4)</td>
</tr>
<tr>
<td>Thyroidectomy</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>633 (99.2)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>5 (0.8)</td>
</tr>
<tr>
<td>LN excision (%)</td>
<td></td>
</tr>
<tr>
<td>Central (C)</td>
<td>106 (16.6)</td>
</tr>
<tr>
<td>Lateral (L)</td>
<td>75 (11.8)</td>
</tr>
<tr>
<td>C + L</td>
<td>438 (68.6)</td>
</tr>
<tr>
<td>No LN dissection</td>
<td>19 (3)</td>
</tr>
<tr>
<td>pTNM stage (%)</td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>247 (38.71)</td>
</tr>
<tr>
<td>pT2</td>
<td>81 (12.70)</td>
</tr>
<tr>
<td>pT3</td>
<td>292 (45.77)</td>
</tr>
<tr>
<td>pT4</td>
<td>18 (2.82)</td>
</tr>
<tr>
<td>Multifocality (%)</td>
<td>320 (50.16)</td>
</tr>
<tr>
<td>Vascular invasion (%)</td>
<td>70 (10.97)</td>
</tr>
<tr>
<td>Primary tumor extrathyroidal extension (ETE) (%)</td>
<td>275 (43.17)</td>
</tr>
<tr>
<td>LN involvement (%)</td>
<td></td>
</tr>
<tr>
<td>N1a</td>
<td>289 (45.3)</td>
</tr>
<tr>
<td>N1b</td>
<td>345 (54.1)</td>
</tr>
<tr>
<td>Unspecified</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>Extra-capsular LN involvement (LN-ECI)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>208 (32.6)</td>
</tr>
<tr>
<td>No</td>
<td>361 (56.6)</td>
</tr>
<tr>
<td>Unspecified</td>
<td>69 (10.8)</td>
</tr>
<tr>
<td>LN score</td>
<td></td>
</tr>
<tr>
<td>High-risk N1 (HR-N1)</td>
<td>408 (64)</td>
</tr>
<tr>
<td>Low-risk N1 (LR-N1)</td>
<td>230 (36)</td>
</tr>
<tr>
<td>Radioiodine ablation (RAI) (%)</td>
<td>638 (100)</td>
</tr>
<tr>
<td>Median follow-up (months) (range)</td>
<td>41.6 (6–171)</td>
</tr>
</tbody>
</table>

LR-N1, N1a patients without LN-ECI; HR-N1, N1b patients and N1a patients with LN-ECI.
all others (i.e. N1b patients and N1a patients with N1-ECI). Patients with unspecified location and/or unspecified N1-ECI were classified by default in the LR-N1 group.

**Post-operative cervical ultrasonography**

POCUS was systematically and prospectively performed for all patients after surgical treatment and just before RAI. It was performed with a real-time US scanner (Toshiba (Tokyo, Japan) or Siemens (Munich, Germany)) with a 7.5–13 MHz bandwidth linear transducer and by experienced sonographers trained in neck US. The examination included both the central and lateral neck compartments, studied with gray-scale and color Doppler. Abnormal LNs were systematically drawn onto a diagram of the neck compartments as recommended by the European guidelines (14). All US scan images were reviewed by a single experienced sonographer who was unaware of the final outcome (L. Leenhardt).

According to the recent US European guidelines (14), the US result was considered suspicious when at least one of the following features was detected:

i) a vascular hypoechogenic mass in the thyroid bed, with or without microcalcifications or cystic areas,

ii) a LN with at least one of the following signs: microcalcifications, cystic areas, peripheral or anarchic vascularization, echogenicity similar to thyroid tissue (hyperechoic area).

The US examination was considered normal in the absence of these signs (14). The location (central or lateral neck) and the size (small (S) and long axis (L)) of the LN were recorded. Criteria of intermediate significance such as L:S ratio < 2 and absence of hilum were not considered suspicious in the absence of the four major echo-Doppler signs aforementioned. In cases of multiple suspicious LN, the largest was analyzed.

**RAI treatment and follow-up**

All patients received an ablative $^{131}$I dose (RAI) of 3.7 GBq within 12 months of surgery, on average, 90 days after surgery (± 48 days).

Tg levels were measured using the immuno-radiometric technique, BRAHMS, coupled with the systematic detection of anti-Tg antibodies. s-Tg levels were assessed at the time of radioiodine ablation (RAI-Tg) after 4 weeks of thyroid hormone withdrawal or after stimulation by rTSH (Thyrogen) during the follow-up. An RAI-Tg value > 10 µg/l was considered suspicious for persistent or recurrent disease (12). Patients with interfering anti-Tg antibodies were not excluded, as our aim was to assess the POCUS value in all N1-PTC patients.

A WBS was performed 5 days after administration of the ablative $^{131}$I dose of 3.7 GBq (tWBS). Images were obtained using an APEX HELIX SPX camera (Elscint, Nicosia, Cyprus) until 2007 and a TDM INFINIA camera (GE Medical System) afterwards. Radioiodine uptake outside the thyroid bed was considered suspicious for metastatic LN persistence or early recurrence.

Median follow-up was 41.6 months (range 6–171 months). Follow-up protocol consisted of a systematic first check-up 6 months after radioiodine ablation, including clinical examination, neck ultrasonography, and s-Tg (rTSH Tg or Tg level after levothyroxine ($\mathrm{-T_4}$) withdrawal). A WBS after a diagnostic $^{131}$I dose (dWBS) under thyroid-stimulating hormone (TSH) stimulation was also performed for 378 patients at 6 months (58%). For 59 patients (9%), it was the only imaging examination performed at 6 months (without cervical ultrasonography). Subsequent check-ups were systematically performed under TSH stimulation 12 months later and every 3 years for 7 years and subsequently, under suppressive treatment, every 3 years (on site or via periodic correspondence with patients or their referring physicians). Other imaging examinations (cervical computed tomography, $^{18}$FDG positron emission tomography) were prescribed at the discretion of the attending physician. All patients were receiving TSH suppressive therapy. Follow-up information and date of death were obtained either from the medical records or from town council registers.

**Clinical endpoints**

The diagnosis of recurrence was assessed by histological analysis after surgery, by fine-needle aspiration biopsy (FNAB; malignant cytological result and/or Tg washout measurement > 10 µg/l (15)), or positive $^{131}$I cervical uptake after a $^{131}$I diagnostic (11.1 MBq) or therapeutic (3.7 GBq) dose. WBS was considered abnormal, suspicious for recurrence, when uptake was seen outside the thyroid bed at ablation or if cervical uptake was persistent after the first radioiodine treatment (dWBS and/or tWBS at second treatment).

Recurrence was divided into two groups depending on the date of occurrence (Fig. 1):

i) Persistence (PD) if disease was detected at ablation and/or at 6 months of follow-up without documented disease-free period.

ii) Late recurrence (l-RD) if no PD was detected and when recurrence was detected after a normal check-up at 6 months (negative cervical US, s-Tg ≤ 1 µg/l (and/or...
in case of detectable anti-Tg antibody, if the d-WBS was normal).

The POCUS was performed, on average, 90 days (± 48 days) after surgery. Its diagnostic value was calculated taking only PD into account and regarding the entire follow-up of the patient. In this analysis, POCUS results were considered:

i) True positive (TP) if POCUS and other cervical US performed during the follow-up were suspect for the same structure until the PD or l-RD evidence.

ii) True negative if POCUS and all other cervical US performed during the follow-up were normal without recurrence and/or when the l-RD occurred tardily after more than 1-year follow-up with both POCUS and 6-month cervical US (US-6) negative.

iii) False positive (FP) if POCUS was suspect without recurrence during the entire follow-up of the patient.

iv) False negative if POCUS was normal with PD or l-RD occurrence detected as a result of other cervical US and/or dWBS and/or tWBS after a second RAI treatment during the first year of follow-up.

Statistical analysis

Descriptive statistics used numbers and percentages for qualitative variables; mean and s.d.s or medians and interquartile intervals for quantitative ones. The diagnostic value of POCUS, RAI Tg, and tWBS was evaluated retrospectively according to PD occurrence.

The interest of POCUS for PD was first evaluated using the usual parameters: sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV). Then, a complete analysis of the diagnostic factors of recurrence was carried out, with χ²-tests or Fisher’s exact tests for qualitative variables and Student’s t-tests for quantitative ones, followed by a stepwise logistic regression. Variables were entered into the stepwise regression when their P value was < 0.10 in the univariate analysis, and they were retained in the final model when the P value of the Wald test was < 0.05. The relationships between sensitivity or specificity and the timing of POCUS and location of abnormal findings were also assessed using univariate and bivariate logistic regressions. Cumulative incidence of recurrence functions was estimated by one-minus the Kaplan–Meier survival function estimate. For late recurrence and overall survival, two survival analyses were carried out with the same method: first, a univariate analysis of each potential prognostic factor, involving log-rank tests for qualitative variables and univariate Cox models for quantitative ones. In a second step, variables with a P value lower than 0.10 in the univariate step were entered in a stepwise Cox model regression. Variables significant with a P value lower than 0.05 with the Wald tests were retained in the final models. All the tests were two sided. Computations were carried out using the SAS V9 statistical package (SAS, Cary, NC, USA).

Results

Recurrence occurred in 138 patients (21.6%), who were divided into two groups: 121 PD (88%) and 17 l-RD (12%). Recurrence was diagnosed by histology for 116 patients (86%), positive ¹³¹I cervical uptake for 13 patients (9.6%), and FNAB for nine patients (5%). The average time period between the primary tumor diagnosis and the recurrence onset was 18 months (range 3–144 months; s.d. 17.7). Among the 121 PD, 26 were located in the central neck compartment only, 63 in the lateral neck, and 32 concerned both the central and lateral compartments.

Diagnostic value of POCUS and comparison with tWBS and RAI-Tg

For the detection of PD, sensitivity, specificity, NPV, and PPV of POCUS were 82.6, 87.4, 95.6, and 60.6% respectively. Data are detailed in Table 2.
The location of abnormal findings (central or lateral neck) did not significantly influence POCUS sensitivity and specificity even when correlated with the timing of POCUS (P=0.74).

The timing of POCUS alone did not significantly influence its sensitivity (P=0.91); however, the specificity increased with the time lapse between POCUS and surgery and exceeded 90% when performed 90 days after surgery (Table 3).

POCUS revealed LN metastasis (i.e. a TP POCUS result) in 100/638 (15.7%) patients, among which 58 were located in the central neck (exclusively or with associated lateral involvement). The mean size of suspect LN was 16.3 (G7.1) mm for the largest (L) diameter and 8.8 (G3.8) mm

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Risk factor for locoregional persistent disease (PD).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
<td>PD Yes (n=121)</td>
</tr>
<tr>
<td>Age (n (%))</td>
<td></td>
</tr>
<tr>
<td>≥ 45 years</td>
<td>54 (44.6)</td>
</tr>
<tr>
<td>&lt; 45 years</td>
<td>67 (55.4)</td>
</tr>
<tr>
<td>Sex (n (%))</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>41 (33.9)</td>
</tr>
<tr>
<td>F</td>
<td>80 (66.1)</td>
</tr>
<tr>
<td>LN dissection (n (%))</td>
<td></td>
</tr>
<tr>
<td>Central (C)</td>
<td>15 (12.4)</td>
</tr>
<tr>
<td>Lateral (L)</td>
<td>14 (11.6)</td>
</tr>
<tr>
<td>C + L</td>
<td>92 (76)</td>
</tr>
<tr>
<td>Multifocal involvement (n (%))</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>76 (62.8)</td>
</tr>
<tr>
<td>No</td>
<td>45 (37.2)</td>
</tr>
<tr>
<td>Vascular invasion (n (%))</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>19 (15.7)</td>
</tr>
<tr>
<td>No</td>
<td>102 (84.3)</td>
</tr>
<tr>
<td>ETE primary tumor (n (%))</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>75 (62)</td>
</tr>
<tr>
<td>No</td>
<td>46 (38)</td>
</tr>
<tr>
<td>Tumor size (mm), mean ± s.d.</td>
<td>23.6 ± 17.8</td>
</tr>
<tr>
<td>pTx classification (n (%))</td>
<td></td>
</tr>
<tr>
<td>pT1</td>
<td>28 (23.1)</td>
</tr>
<tr>
<td>pT2</td>
<td>13 (10.7)</td>
</tr>
<tr>
<td>pT3</td>
<td>70 (57.9)</td>
</tr>
<tr>
<td>pT4</td>
<td>10 (8.3)</td>
</tr>
<tr>
<td>LN involvement (n (%))</td>
<td></td>
</tr>
<tr>
<td>N1a</td>
<td>34 (28.1)</td>
</tr>
<tr>
<td>N1b</td>
<td>87 (71.9)</td>
</tr>
<tr>
<td>LN-ECI (n (%))</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>65 (53.7)</td>
</tr>
<tr>
<td>No</td>
<td>47 (38.8)</td>
</tr>
<tr>
<td>Unspecified</td>
<td>9 (7.5)</td>
</tr>
<tr>
<td>LN score</td>
<td></td>
</tr>
<tr>
<td>High-risk N1 (HR-N1)</td>
<td>103 (85.1)</td>
</tr>
<tr>
<td>Low-risk N1 (LR-N1)</td>
<td>18 (14.9)</td>
</tr>
<tr>
<td>Initial distant metastasis (n (%))</td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>114 (94.2)</td>
</tr>
<tr>
<td>M1</td>
<td>7 (5.8)</td>
</tr>
<tr>
<td>POCUS (n (%))</td>
<td></td>
</tr>
<tr>
<td>Abnormal result</td>
<td>100 (82.6)</td>
</tr>
<tr>
<td>Normal result</td>
<td>21 (17.4)</td>
</tr>
<tr>
<td>tWBS (n (%))</td>
<td></td>
</tr>
<tr>
<td>Abnormal result</td>
<td>48 (39.7)</td>
</tr>
<tr>
<td>Normal result</td>
<td>73 (60.3)</td>
</tr>
<tr>
<td>RAI-Tg (n (%))</td>
<td></td>
</tr>
<tr>
<td>&gt; 10 μg/l</td>
<td>83 (68.6)</td>
</tr>
<tr>
<td>&lt; 10 μg/l</td>
<td>38 (31.4)</td>
</tr>
</tbody>
</table>

C, central; L, lateral; ETE, extrathyroidal extension; LN-ECI, extra-capsular LN involvement; RAI-Tg, thyroglobulin at radioiodine ablation.
for the smallest (S) diameter. These patients with TP POCUS were treated for their PD, on average, 15 months after the first surgery and 13 months after the POCUS.

Sixty-five patients (10%) had a suspect POCUS without recurrence and were thus considered to have FP results. Their characteristics and outcomes are summarized in Table 4. Among these 65 FP POCUS patients, 26 had an abnormal tWS and the median RAI-Tg was 52 μg/l (Table 4).

Finally, 473 POCUS patients (71%) were considered normal and 452 of them (95.6%) never developed a recurrence. Among the total series, 21 POCUS results were considered false negative (FN, 3.3%). The POCUS was normal for all 21 patients, but for 16 of them the subsequent US control performed during the first year of follow-up revealed a suspicious LN (cytologically or histologically proven) and for five of them a positive 131I cervical uptake after a second radioiodine treatment was observed. Five patients among the 21 FN POCUS had an abnormal cervical tWBS uptake at ablation. For two of these cases, PD was diagnosed after a positive 131I cervical uptake and for three of them PD was proven histologically.

An uptake outside the thyroid bed was detected at tWBS in 149/638 patients. Among these, 48 patients had PD proven by histology, cytology, or persistence on tWBS after a second RAI treatment. The remaining 101 patients with abnormal tWBS were considered FP of tWBS in the absence of PD or proven recurrence during subsequent follow-up. Seventy-three patients with normal cervical tWBS uptake had at least one LN metastasis detected, thanks to POCUS. Sensitivity, specificity, NPV, and PPV of tWBS for the diagnosis of PD were 39.7, 80.5, 85, and 32% respectively.

At ablation, RAI-Tg was undetectable in 79 patients (12.4%), under 10 μg/l in 206 (32.3%) and 463 (72.6%) patients respectively. Median RAI-Tg was 52 μg/l (range 0.15–300) and 13 μg/l (range 0.15–300) in FP and FN POCUS patients respectively. One hundred and seventy-one patients (27%) had both a RAI-Tg under 1 μg/l and a negative POCUS. Among them, tWBS was abnormal in 28 cases and a locoregional recurrence was detected in six cases: two patients with PD and four with l-RD. A RAI-Tg higher than 10 μg/l was chosen as a cut-off value with a good prediction of PD (P<0.0001). Sensitivity, specificity, NPV, and PPV of RAI-Tg ≥10 μg/l for prediction of PD were 68.6, 82.2, 91.8, and 47.4% respectively.

Sensitivity and specificity of POCUS were significantly higher than those of tWBS (P<0.0001 and P=0.0008 respectively) and of RAI-Tg (P=0.001 and P=0.01 respectively).

### Predictive factors for PD, late recurrence, and patients’ outcome

In N1 PTC patients, only a suspect POCUS result and a RAI-Tg >10 μg/l were significantly predictive of PD (P<0.0001) at multivariate analysis (Table 2). None of the other factors listed on Table 2 were significant.

### Table 3  Sensitivity and specificity of POCUS according to the timing for POCUS after surgery.

<table>
<thead>
<tr>
<th>Time period between surgery and POCUS (days)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 60</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td>60–90</td>
<td>92</td>
<td>86</td>
</tr>
<tr>
<td>90–120</td>
<td>79</td>
<td>91</td>
</tr>
<tr>
<td>&gt; 120</td>
<td>80</td>
<td>92</td>
</tr>
<tr>
<td>P value</td>
<td>0.91</td>
<td>0.0039</td>
</tr>
</tbody>
</table>

### Table 4  Characteristics of patients with false-positive (FP) POCUS.

| Total numbers | 65/638 (10%) | 38/65 (58%) |
| Number of patients with repeated cervical US (%) | 28/65 (43%) | 28/638 (4.3%) |
| Number of FNAB (%) | 11/65 (17) | 11/638 (1.7) |
| Number of negative reoperation (%) | 2/65 (3) | 2/638 (0.3%) |
| Number of additional RAI (%) | 8/65 (12) | 8/638 (1.2) |
| Number of patients with abnormal tWBS (%) | 26/65 (40) |
| For all | 7/65 (11) |
| For patients with normal US-6 | 19/65 (29) |
| Median RAI-Tg (μg/l) (range) | 52 (0.15–300) |
| For all | 19 (0.15–84.1) |
| For patients with normal US-6 | 77 (0.2–300) |
| Suspect LN characteristic | 18 (28) |
| LN with microcalcification (%) | 46 (72) |
| LN with peripheral or anarchic vascularization (%) | 13 (20) |
| LN with cystic change (%) | 18 (28) |
| Hyperchogenic LN (%) | 46/65 (70) |
| Endpoint | 2/65 (3) |
| Cured (%) | 16/65 (26) |
Only three factors were predictive of late recurrence (l-RD) in the univariate analysis: RAI-Tg ≥10 μg/l (P=0.0091), initial N1b status (P=0.0191), and initial M1 status (P=0.0131). Only RAI-Tg ≥10 μg/l remained predictive in the multivariate analysis (HR 3.353, 95% CI (1.278–8.793); P=0.0139).

At the end of the study, 626 (98.1%) patients were still alive and 492 patients (77.1%) considered cured: normal neck US and a s-Tg level ≤1 μg/l (and/or a normal d-WBS in case of detectable anti-Tg antibody). Among the total series, 12 (1.9%) patients died; seven (1.1%) of these were from their thyroid cancer and two from another cause. The cause of death was unknown for three patients. All those patients who died from their thyroid cancer had developed a recurrence. Overall survival of the entire cohort was superior by 99% at 5 years. Factors significantly influencing survival in univariate analysis were age at diagnosis (P<0.0001), the occurrence of recurrence (HR 5.085, 95% CI (1.25–20.67), P=0.0139) and pT4 initial stage (P<0.0001), the occurrence of recurrence (HR 5.085, 95% CI (1.25–20.67), P=0.018) even in the case of persistence (PD) (97.4 vs 100%, P=0.005). In the multivariate analysis, only age at diagnosis ≥45 years (HR 8.841, 95% CI (1.05–74.28); P=0.048) and pT4 initial stage still influence the survival (HR 8.062, 95% CI (2.03–32.02); P=0.003).

**Risk stratification**

Among the recurrence occurred in 138 patients, cumulative incidence estimated by the Kaplan–Meier method was 11.4% (69 patients) during the first year of follow-up, 22% (119 patients) within 3 years, and 26.2% (131 patients) within 5 years.

Taking into account the initial pTNM risk stratification, 26% of ETA HR patients, 17% of ATA intermediate risk patients, and 35% of ATA HR patients developed a recurrence at 5 years (Table 5).

When the recurrence risk was reassessed with POCUS and its result was normal, only 9% of ETA HR patients, 8% of ATA intermediate, and 11% of ATA HR patients had locoregional recurrence at 5 years.

Patients with a RAI-Tg value under 10 μg/l also had a lower risk of recurrence estimated at 5 years, which represented 13% for ETA HR, and 9 and 18% for ATA intermediate and HR patients respectively.

Finally, a combination of normal POCUS and RAI-Tg value under 10 μg/l conferred a reassessed risk of recurrence lower than 6% for the three groups (Table 5).

**Table 5 Five-year cumulative incidence (%) of recurrence according to ATA and ETA initial stratification depending on POCUS and RAI-Tg results.**

<table>
<thead>
<tr>
<th></th>
<th>ATA</th>
<th>ETA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate N (n = 323)</td>
<td>High N (n = 315)</td>
<td>High N (n = 638)</td>
</tr>
<tr>
<td>5-year cumulative incidence of recurrence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk reassessment of recurrence depending on POCUS result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal POCUS (n = 473)</td>
<td>17%</td>
<td>35%</td>
</tr>
<tr>
<td>Abnormal POCUS (n = 165)</td>
<td>8%</td>
<td>11%</td>
</tr>
<tr>
<td>Risk reassessment of recurrence depending on Tg IRA result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tg IRA &lt;10 μg/l (n = 463)</td>
<td>56%</td>
<td>77%</td>
</tr>
<tr>
<td>Tg IRA ≥10 μg/l (n = 175)</td>
<td>9%</td>
<td>18%</td>
</tr>
<tr>
<td>Risk reassessment of recurrence depending on POCUS and RAI-Tg result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal POCUS and RAI-Tg &lt;10 μg/l (n = 387)</td>
<td>4%</td>
<td>6%</td>
</tr>
<tr>
<td>Abnormal POCUS or RAI-Tg ≥10 μg/l (n = 251)</td>
<td>42%</td>
<td>63%</td>
</tr>
</tbody>
</table>

Recurrence: locoregional recurrence including persistent disease (PD) and late recurrence (l-RD).

**Discussion**

Over the past few years, new concepts have emerged concerning the usefulness of neck ultrasonography in the management of PTC patients. They focus on three main points:

i) Neck US as the best diagnostic tool for detecting recurrence when combined with Tg measurement. In particular, it gives a more accurate diagnosis than a diagnostic WBS (16, 17).

ii) Neck US as a two-stage dynamic risk assessment for recurrence: first at the post-operative period to tailor radioiodine ablation indications to individual needs, secondly during the follow-up (especially the first 2 years) to categorize response to therapy.

iii) The concept that N1 patients have a heterogeneous risk of recurrence and that specific characteristics of LN metastases could be used to stratify the risk of recurrence in PTC.

**Comparison of POCUS with RAI-Tg and tWBS**

Our results showed that POCUS accuracy is better than that achieved by tWBS. This could partly be due to the fact...
that before 2007, most patients were explored by planar whole-body scintigraphy. This could lead to an under-estimation of tWBS diagnostic value in this study (18, 19, 20). An abnormal POCUS is highly predictive of PD with high sensitivity, especially when combined with RAI-Tg levels. Rosario et al.’s results (12, 21) showed that a RAI-Tg threshold value > 10 µg/l is predictive of PD, but this report only included LR patients without initial LN metastasis. This finding should be taken with caution for patients with interfering anti-Tg antibodies. In our study, although some patients presented such antibodies, they were not excluded and RAI-Tg levels remained closely associated with the risk of recurrence.

Finally, in multivariate analysis, only the POCUS and the RAI-Tg results remain predictive of PD (Table 2). Neither the tWBS, nor any other classical risk factors such as initial pTNM, male sex, and age are significantly associated with PD. Likewise, initial metastatic status (19 patients) does not influence PD risk in multivariate analysis. We chose not to exclude these patients, considering that POCUS is a part of their management and its diagnostic value for PD and recurrence should be evaluated.

**POCUS contribution to patient management**

A normal POCUS enabled us to reassure 461 patients (70.7%). Moreover, more than 95% of patients with a normal POCUS never developed a recurrence. This high NPV of POCUS when combined with an undetectable RAI-Tg was also reported by Nascimento et al. (22) (whose series included 44.2% of N1 patients). The author used this high NPV in the RAI decision-making and follow-up.

Regarding TP POCUS results, patients were reoperated, on average, 13 months after POCUS. This time lapse can be explained by the necessity for a systematic US and FNAB confirmation 3–6 months after RAI treatment and for some patients because of the small size of suspect LN detected. Some series outline that LN reoperation achieves biochemical complete remission in only 27–51% of cases and sometimes requires several operations (23, 24) and other series support the fact that detecting LN after thyroidectomy for PTC is of no clinical importance (25). Although the long-term effectiveness of LN reoperation has not been proved in our study, it could alter the survival rate in patients with recurrence, especially for radioiodine refractory recurrence, located in the thyroid bed or in the neck soft tissues (26, 27). Whilst reoperation could be criticized, POCUS is necessary to identify patients with HR of recurrence and recent European US guidelines should help identify those requiring FNAB and surgery (14).

A false-positive POCUS result (64/638 (10%) in our series) risks creating anxiety for both patients and clinicians and can lead to unnecessary examinations. Several factors could account for this rate: the(se) small LN metastases could have been destroyed by the RAI treatment or could remain dormant (25). Another explanation could be the timing of POCUS, which was carried out, on average, 90 days after surgery (+ 48 days). Too early, a POCUS could lead to the misinterpretation of post-surgery tissue swelling. In our study, we showed a better specificity of POCUS when the examination was done 3 months after surgery without modification of the sensitivity. However, among the 65 patients who had a suspect POCUS considered to be FP (64/638; 10%), only 11 (1.7% of the entire cohort) underwent a FNAB and two (0.3%) had negative neck surgery.

Finally, based on the latest ATA guidelines (6) that recommend a preoperative neck US to assess the central and lateral neck compartments for all patients undergoing thyroidectomy for malignant cytology, one could argue the relevance of doing a POCUS if this preoperative US is done proficiently. In our series, a preoperative US was systematically done for all patients before surgery. However, because of the retrospective design of the study,
conducted from 1997 to 2010, a systematic second neck US following FNAB malignant results was not always performed as suggested in the recent recommendation and not always in a reference center. Furthermore, it has been reported that preoperative US identifies only half the number of LNs found at surgery, due to the presence of the overlying thyroid gland (28). Moreover, this POCUS can reveal LN metastases that were not detectable at surgery but whose growth had developed since due to the TSH stimulation required for radioiodine treatment.

POCUS contribution to risk reassessment

N1 patients have recently been shown as a heterogeneous group regarding recurrence risk (29), since the location of LN involvement (central N1a vs lateral N1b), the existence of LN-ECI, the size and number of initial LN metastasis (29, 30) are all factors that determine this risk. Recently, some authors have introduced a new concept of delayed risk stratification (DRS) based on data obtained 8–12 months after RAI, the time of the first diagnostic control (5, 9). This DRS better defined the patient’s risk of recurrence and interestingly (DRS) led to a significant change in risk estimates in N1-PTC patients (5).

In our series, N1a patients without N1-ECI showed a lower risk of PD than that of other LN (Table 2). Therefore, early risk stratification could be created by re-stratifying N1-PTC patients with normal POCUS from intermediate to LR of PD according to the ATA classification. Indeed, only 3.2% (21/482) of patients with a negative POCUS developed a recurrence during the first year of follow-up. Furthermore, a combination of normal POCUS and RAI-Tg value under 10 μg/l conferred a 5-year cumulative incidence risk of recurrence of <6% whatever the initial pTNM group. Thus, the recurrence risk is comparable with the ATA LR group PTC (5, 6).

Consequently, early risk reassessment with POCUS could help to individually tailor radioiodine treatment to individual needs. As suggested in Fig. 2, patients with a normal POCUS and RAI-Tg <10 μg/l present less risk of recurrence (5%) and could be treated with a low dose (1.1 GBq) of radioiodine as has been suggested in three recent studies (31, 32, 33). However, other patients could benefit from a more traditional radioiodine treatment in order to destroy any neoplastic foci.

In conclusion, despite some limitations inherent to its retrospective design (about 10% patients lost during follow-up, impossibility to isolate the subgroup of ATA intermediate pT3 subgroup with microscopic perithyroidal invasion), the results of this large cohort study highlight the crucial diagnostic role of POCUS in detection of PD. A normal POCUS result combined with a low RAI-Tg level allows the clinician to reassure the patient by re-stratifying his case into the LR N1-PTC category. Thus, our results support the introduction of a systematic POCUS in individualized risk assessment to improve therapeutic interventions and follow-up management.

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 research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

References

2 Mazzaferri EL & Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. American Journal of Medicine 1994 97 418–428. (doi:10.1016/0002-9343(94)90321-2)
4 Mazzaferri EL. An overview of the management of papillary and follicular thyroid carcinoma. Thyroid 1999 9 421–427. (doi:10.1089/thy.1999.9.421)
9 Castagna MG, Maino F, Cipri C, Belardini V, Theodoreopoulou A, Cevenini G & Pacini F. Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients.
with undetectable postoperative stimulated thyroglobulin level. *Endocrine-Related Cancer* 2011 **18** R29–R40. (doi:10.1677/ERC-10-0292)


27 Mazzaferri EL & Kloos RT. Clinical review 128: Current approaches to primary therapy for papillary and follicular thyroid cancer. *Journal of Clinical Endocrinology and Metabolism* 2001 **86** 1447–1463. (doi:10.1210/jcem.86.4.7407)


29 Randolph G, Duhl QY, Heller KS, Livoi-VI, Mandal SJ, Steward D, Tufano JP & Tuttle RM. The prognostic significance of nodal metastases from papillary thyroid carcinoma can be stratified based on the size and number of metastatic lymph nodes, as well as the presence of extranodal extension ATA surgical committee’s taskforce on thyroid cancer nodal surgery. *Thyroid* 2012 **22** 1144–1152. (doi:10.1089/thy.2012.0043)


33 Rosario PW & Calasori MR. Thyroid ablation with 1.1 GBq (30 mCi) iodine-131 in patients with papillary thyroid carcinoma at intermediate risk for recurrence. *Thyroid* 2014. In Press.