Background: The optimal approach for the diagnosis of hypothyroidism and hyperthyroidism in hospitalized patients is controversial.

Objectives: To estimate the prevalence of undiagnosed thyroid disease among inpatients, review the usefulness of clinical signs and symptoms, and elucidate the characteristics of the sensitive thyrotropin (thyroid-stimulating hormone) (sTSH) test in this population.

Methods: We undertook a systematic review of the literature by conducting a MEDLINE search covering January 1966 through December 1996. Searching was conducted in duplicate and independently. Specific inclusion and exclusion criteria were predetermined.

Results: Prevalence of thyroid disease among inpatients is approximately 1% to 2% and is similar to the outpatient population. Absence of clinical features of thyroid disease lowers the pretest likelihood and makes screening even less useful. Presence of clinical features, especially those specific for thyroid disease (eg, goiter), may increase the pretest likelihood and increase the yield of testing. Acute illness reduces the specificity of second-generation sTSH tests for thyroid disease. The positive likelihood ratio associated with an abnormal sTSH test result in ill inpatients is about 10 compared with about 100 in outpatients.

Conclusion: In unselected general medical, geriatric, or psychiatric inpatient populations, sTSH testing provides a low yield of true-positive and many false-positive results.

Arch Intern Med. 1999;159:658-665

HAVING A sensitivity and specificity of more than 99%,1 the sensitive thyrotropin (thyroid-stimulating hormone) (sTSH) assay currently is the best initial test for the diagnosis of thyroid disease in outpatients.2 Nonthyroidal illness (NTI) has a negative impact on both the sensitivity and specificity of sTSH. Recommendations on the use of the sTSH assay in patients with NTI range from limiting use of the test to patients with high clinical suspicion3-6 to generalized screening.7-9

Determining the optimal strategy for the use of any diagnostic test requires elucidating 3 variables: (1) the underlying risk of the target disease in the population of interest, (2) clinical variables that may raise or lower this risk, and (3) characteristics of the test in the population of interest. The first 2 variables lead to a pretest likelihood of disease and the third leads to a posttest likelihood of disease.

We applied this 3-fold framework to the diagnosis of thyroid disease in acutely ill, hospitalized patients with NTI. We conducted a systematic review that addressed the following 3 questions: (1) What is the prevalence of unrecognized thyroid disease? (2) What is the usefulness of clinical signs and symptoms in predicting thyroid disease? and (3) What is the sensitivity and specificity of the sTSH assay in the diagnosis of thyroid disease?

In acute illness, all thyroid indices have been shown to yield false-positive results.10 Because of this, we have decided that the gold standard for diagnosing thyroid disease in this population is follow-up after resolution rather than the actual thyroid index chosen in the particular study. We recognize that the new third-generation sTSH assay has greatly increased the sensitivity of diagnosis, mostly of hyperthyroidism, in an outpatient setting.

MATERIALS AND METHODS

Throughout, we sought studies reported in the English language that enrolled more than 50 patients. To address the prevalence of unrecognized thyroid disease in acutely ill, hospitalized patients with NTI. We conducted a systematic review that addressed the following 3 questions: (1) What is the prevalence of unrecognized thyroid disease? (2) What is the usefulness of clinical signs and symptoms in predicting thyroid disease? and (3) What is the sensitivity and specificity of the sTSH assay in the diagnosis of thyroid disease?

In acute illness, all thyroid indices have been shown to yield false-positive results.10 Because of this, we have decided that the gold standard for diagnosing thyroid disease in this population is follow-up after resolution rather than the actual thyroid index chosen in the particular study. We recognize that the new third-generation sTSH assay has greatly increased the sensitivity of diagnosis, mostly of hyperthyroidism, in an outpatient setting.

MATERIALS AND METHODS

Throughout, we sought studies reported in the English language that enrolled more than 50 patients. To address the prevalence of unrecognized thyroid disease in acutely ill, hospitalized patients with NTI. We conducted a systematic review that addressed the following 3 questions: (1) What is the prevalence of unrecognized thyroid disease? (2) What is the usefulness of clinical signs and symptoms in predicting thyroid disease? and (3) What is the sensitivity and specificity of the sTSH assay in the diagnosis of thyroid disease?

In acute illness, all thyroid indices have been shown to yield false-positive results.10 Because of this, we have decided that the gold standard for diagnosing thyroid disease in this population is follow-up after resolution rather than the actual thyroid index chosen in the particular study. We recognize that the new third-generation sTSH assay has greatly increased the sensitivity of diagnosis, mostly of hyperthyroidism, in an outpatient setting.
ill, hospitalized patients, we used the following MEDLINE search strategy: (explode) thyroid diseases and prevalence (text word) and 1 of the following: hospitalized or medical or inpatient (all text words) or atrial fibrillation or (explode) mental disorder or (explode) dementia (medical subject heading [MeSH] terms) or geriatrics (both MeSH and text word). We retrieved and included in the analysis all articles in which the study population was acutely ill or hospitalized with an NTI and those in which the reference standard included biochemical markers (with or without clinical features) and follow-up after resolution of the NTI. For this section, therefore, our main reference standard was long-term follow-up, and we did not limit ourselves to the sTSH assay as a diagnostic test.

To address the usefulness of clinical signs and symptoms in predicting thyroid disease in hospitalized patients, we used the following search strategy: (explode) thyroid diseases and (explode) cohort studies and either signs or symptoms (text words). We retrieved the articles in which the authors described clinical signs and symptoms in sufficient detail to calculate likelihood ratios (LRs). Because we found no articles addressing the usefulness of clinical signs and symptoms in the diagnosis of thyroid disease among inpatients, we included studies of outpatients only in this section.

To address the sensitivity and specificity of the sTSH assay in the diagnosis of thyroid disease in hospitalized patients, we combined thyrotropin (MeSH) and 1 of the following: hospitalized or medical or inpatient (text words) and either sensitivity or specificity (text words). We retrieved articles in which the study population was acutely ill or hospitalized with an NTI and the reference standard included second- or third-generation sTSH assays (ie, functional lower limit of 0.1 and 0.01, respectively) and follow-up after resolution of the NTI. We also searched the bibliographies of relevant articles and review articles for additional studies.

Two of us (J.A. and P.M.) carried out the searches and data extraction independently. We resolved discrepancies by consensus. In abstracting the data, we excluded postpartum patients and those with previously diagnosed thyroid disease, as well as subclinical hypothyroidism or hyperthyroidism. Subclinical thyroid disease was defined as abnormalities in 1 thyroid index (ie, sTSH), while other indices (ie, free thyroxine) were normal in a patient without overt symptoms of thyroid disease.

We graded the quality of the studies selected using the following 3 criteria: (1) follow-up of normal results included vs follow-up only of abnormal results; (2) greater than 60% of abnormal results followed up after resolution of NTI vs less than 60%; and (3) criteria for diagnosis of thyroid disease clearly and explicitly stated vs implicit or clinical diagnosis.

We pooled the results of individual studies using a weighted average (according to sample size) only if a formal test of heterogeneity using a $\chi^2$ test was not significant at $P<.05$.

## RESULTS

### Prevalence of Unrecognized Thyroid Disease

The MEDLINE search identified 110 titles. Thirty-one articles addressed the inpatient population, 16 of which had the required follow-up after resolution of the NTI, allowing us to calculate the prevalence of unrecognized thyroid disease. We grouped these 16 articles according to whether they studied general medical inpatients, geriatric inpatients, inpatients with atrial fibrillation, or psychiatric inpatients.

Two studies addressing the prevalence of thyroid disease in general medical inpatients met our inclusion criteria (Table 1). The first study is the most methodologically sound of all those included in this review. The authors screened 1580 general medical inpatients on admission with a sTSH level and found 17.2% to be outside the normal range. The investigators followed up 63% of the patients with abnormal results until resolution of the acute illness and found that, on retesting, 85% were euthyroid by criteria that combined thyroid indices and clinical examination. False-positive results were attributed to acute illness and drug interactions, in particular glucocorticoids. Investigators also followed up an arbitrary sample of 48% of the initially euthyroid group and found no new cases of thyroid disease.

<table>
<thead>
<tr>
<th>Population</th>
<th>Reference</th>
<th>Study Type</th>
<th>Diagnosis</th>
<th>Follow-up</th>
<th>Prevalence, %</th>
<th>Quality Criteria†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1580 Acutely ill, hospitalized patients</td>
<td>11</td>
<td>Prospective</td>
<td>Sensitive thyrotropin</td>
<td>329 Selected patients (both normal and abnormal TSH assay results)</td>
<td>0.55</td>
<td>Yes Yes Yes</td>
</tr>
<tr>
<td>363 Women aged &gt;50 y admitted to medical casualty ward</td>
<td>12</td>
<td>Prospective</td>
<td>Free thyroxine</td>
<td>37 of 53 patients with abnormal results were followed up after resolution of NTI</td>
<td>0.7</td>
<td>No Yes Yes</td>
</tr>
</tbody>
</table>

* Total number of patients is 1943. Total hyperthyroid prevalence is 0.58% and total hypothyroid prevalence is 0.90%. TSH indicates thyroid-stimulating hormone; NTI, nonthyroidal illness.

†A indicates patients with normal thyroid indices followed up; B, more than 60% of patients with abnormal thyroid indices followed up; and C, explicit diagnostic criteria used.
The second study tested 363 female medical inpatients on admission using free thyroxine and found 9.3% to be outside the euthyroid range. The investigators followed up 76% of these patients, 92% of whom were euthyroid or had previously diagnosed thyroid disease. Combining the 2 studies, we found a prevalence of thyroid disease of approximately 1.5%.

Six studies addressing the prevalence of unrecognized thyroid disease in the geriatric population met the inclusion criteria (Table 2). Three studies are prospective, and 3 are retrospective. Despite describing varying populations (eg, those drawn from a psychogeriatric unit and those referred for assessment of dementia), all studies indicated a high rate of false-positive initial thyroid indices and an overall estimate of true thyroid disease of approximately 2%.

Two studies describing the prevalence of thyroid disease in patients with atrial fibrillation met our criteria (Table 3). The first, a longitudinal cohort study, included both inpatients and outpatients. The diagnosis of thyroid disease was left to routine follow-up by the physician. The second study documented a much higher rate of hyperthyroidism and did not examine hypothyroidism. The authors used a shorter follow-up (6 weeks) and an older population and relied on TSH response to thyrotropin-releasing hormone. We believed that their population, methods, and results were too disparate to combine, and this was confirmed by a significant χ² test for heterogeneity.

Six studies addressing the prevalence of thyroid disease in admissions for acute psychiatric illness met our criteria (Table 4). In total, 1479 patients were studied, 181 (12.2%) of whom had initial abnormal thyroid indices. With a follow-up rate of about 90%, approximately 97% of these patients were subsequently found to be euthyroid.

In summary, the overall incidence of thyroid disease among inpatients was between 1.0% and 2.5%, with psychiatric inpatients at the lower end of this range and geriatric inpatients at the upper end. The highest incidence of hyperthyroidism was 0.6%, while hypothyroidism was up to 3 times more common, ie, approximately 1.8%.

### Usefulness of Clinical Signs and Symptoms of Thyroid Disease

No studies addressed the question of the usefulness of clinical signs and symptoms of thyroid disease in inpatients. We reviewed several studies that addressed this issue in outpatients, keeping the remaining inclusion criteria fixed. With the relaxed criteria, our MEDLINE search identified 147 titles, 8 of which met our inclusion criteria; 6 of these addressed individual signs and symptoms, and 2 addressed the total number of signs and symptoms in the diagnosis of thyroid disease.

Table 5 presents the 6 studies that provide information on...
individual signs and symptoms. The first study is the most comprehensive, listing 38 signs and symptoms of thyroid disease in 99 euthyroid, 131 hypothyroid, and 235 hyperthyroid patients from 6 specialty centers in the United States and Australia. We calculated LRs from these data and included clinical features with the highest predictive power. These LRs are similar to those calculated from results in 2 other studies, but they are probably inflated due to the fact that the clinicians were not blinded to the thyroid function test results before performing their physical examination.

Other studies found that clinical features were disappointingly poor in predicting thyroid disease. In a retrospective review of 982 charts, Schectman et al found a poor correlation between clinical features and thyroid disease. These investigators depended on the primary care physicians' records of the initial history and physical examination, and it was unclear whether specific signs and symptoms of thyroid disease were sought and not found or were not sought at all; the latter would underestimate test sensitivity. Two other studies also gave disappointing results, perhaps because they de-

### Table 3. Prevalence of Thyroid Disease in Patients With Atrial Fibrillation (N = 801)

<table>
<thead>
<tr>
<th>Population</th>
<th>Reference</th>
<th>Study Type</th>
<th>Diagnostic Criteria</th>
<th>Follow-up</th>
<th>Prevalence, %</th>
<th>Quality Criteria*</th>
</tr>
</thead>
<tbody>
<tr>
<td>726 Inpatients and outpatients</td>
<td>17</td>
<td>Cohort</td>
<td>Thyrotropin</td>
<td>Follow-up at 3 mo and yearly, clinical and biochemical diagnosis</td>
<td>0.96 1.5</td>
<td>No Yes No</td>
</tr>
<tr>
<td>75 Patients admitted with acute medical emergencies</td>
<td>18</td>
<td>Prospective</td>
<td>Free triiodothyronine, free thyroxine (thyrotropin-releasing hormone stimulation test evaluated)</td>
<td>42 Patients were followed up at 6 wk</td>
<td>5.3 . . . Yes Yes Yes</td>
<td></td>
</tr>
</tbody>
</table>

* A indicates patients with normal thyroid indices followed up; B, more than 60% of patients with abnormal indices followed up; C, explicit diagnostic criteria used; and ellipses, data not reported.

### Table 4. Prevalence of Thyroid Disease in Psychiatric Inpatients*

<table>
<thead>
<tr>
<th>Population</th>
<th>Reference</th>
<th>Study Type†</th>
<th>Diagnostic Criteria</th>
<th>Follow-up</th>
<th>Prevalence, %</th>
<th>Quality Criteria†</th>
</tr>
</thead>
<tbody>
<tr>
<td>98 Patients admitted to an acute psychiatric unit</td>
<td>19</td>
<td>Prospective</td>
<td>Total thyroxine</td>
<td>10 of 15 patients with abnormal results followed up for 14 d</td>
<td>0 . . . No Yes Yes</td>
<td></td>
</tr>
<tr>
<td>106 Patients admitted to an acute psychiatric unit</td>
<td>20</td>
<td>Retrospective</td>
<td>Total thyroxine, thyrotropin</td>
<td>1 Patient with elevated T4 level followed up after 4 d</td>
<td>0 0 No Yes Yes</td>
<td></td>
</tr>
<tr>
<td>386 Patients admitted to an acute psychiatric unit of a Veterans Affairs hospital</td>
<td>21</td>
<td>Retrospective?</td>
<td>FTI, FT3I thyrotropin</td>
<td>30 of 33 hyperthyroid patients normalized when followed up at 2 wk with repeated testing, no follow-up of 4 hypothyroid patients</td>
<td>0 1 No Yes No</td>
<td></td>
</tr>
<tr>
<td>480 Patients admitted to a psychiatric evaluation unit</td>
<td>22</td>
<td>Retrospective?</td>
<td>FTI, thyrotropin</td>
<td>31 of 43 hypothyroid patients and 35 of 42 hypothyroid patients followed up 2 wk into hospital stay</td>
<td>0.29 0.75 No Yes No</td>
<td></td>
</tr>
<tr>
<td>150 Patients admitted to a psychiatric inpatient unit</td>
<td>23</td>
<td>Prospective</td>
<td>Total thyroxine, total triiodothyronine, RT3U, FTI, FT3I</td>
<td>74 of 150 patients followed up at day 10, including all patients with abnormal results</td>
<td>0 26 No Yes Yes</td>
<td></td>
</tr>
<tr>
<td>259 Patients admitted to an adult psychiatric ward</td>
<td>24</td>
<td>Retrospective</td>
<td>Free thyroxine, free triiodothyronine, RT3U</td>
<td>9 Patients with abnormal results followed up</td>
<td>0 0.77 No Yes No</td>
<td></td>
</tr>
</tbody>
</table>

* Total number of patients is 1479. Total hyperthyroid prevalence is 0.09% and total hypothyroid prevalence is 0.97%.
† A indicates patients with normal thyroid indices followed up; B, more than 60% of patients with abnormal thyroid indices followed up; C, explicit diagnostic criteria used; ellipses, data not reported; ?, unclear; FTI, free thyroxine index; FT3I, free triiodothyronine index; and RT3U, triiodothyronine uptake.
Table 5. Signs and Symptoms of Thyroid Disease*

<table>
<thead>
<tr>
<th>Population</th>
<th>Reference</th>
<th>Study Type</th>
<th>Gold Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>99 Euthyroid, 131 hypothyroid, and 253 hyperthyroid outpatients</td>
<td>25</td>
<td>Retrospective?</td>
<td>Review of clinical, test, and biopsy data (unblinded)</td>
</tr>
<tr>
<td>314 79-year-old Swedish women (298 euthyroid and 16 hypothyroid) (community survey)</td>
<td>26</td>
<td>Prospective</td>
<td>Sensitive thyrotropin (unblinded)</td>
</tr>
<tr>
<td>982 Patients in primary care setting (46 hypothyroid and 936 euthyroid) (chart review)</td>
<td>28</td>
<td>Retrospective</td>
<td>Sensitive thyrotropin (unblinded)</td>
</tr>
<tr>
<td>1017 Thyrotoxic patients treated with radioactive iodine (postal questionnaire)</td>
<td>29</td>
<td>Retrospective</td>
<td>Hypothyroid if replacement therapy started due to clinical features and low protein-bound iodine</td>
</tr>
<tr>
<td>110 Outpatients with atrial fibrillation, (cardiology clinic)</td>
<td>27</td>
<td>Prospective</td>
<td>Diagnosis of hyperthyroid based on thyrotropin response to thyrotropin-releasing hormone and clinical criteria (unblinded)</td>
</tr>
<tr>
<td>1193 Elderly outpatients (questionnaire)</td>
<td>30</td>
<td>Prospective</td>
<td>Sensitive thyrotropin, total thyroxine</td>
</tr>
</tbody>
</table>

*LR indicates likelihood ratio; OR, odds ratio; A, patients with normal thyroid indices examined; B, more than 60% of patients with abnormal thyroid indices examined; C, explicit diagnostic criteria used; ellipses, data not reported; and ?, unclear.

Table 6. Number of Clinical Features of Thyroid Disease*

<table>
<thead>
<tr>
<th>Population</th>
<th>Reference</th>
<th>Study Type</th>
<th>Gold Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>135 Ambulatory patients in a family practice group; 10 had thyroid dysfunction</td>
<td>31</td>
<td>Retrospective chart review</td>
<td>Blinded review of thyroid indices by 2 clinicians</td>
</tr>
<tr>
<td>500 New patients presenting consecutively to a thyroid clinic; 21 had thyroid dysfunction</td>
<td>32</td>
<td>Retrospective chart review</td>
<td>Based on thyroid indices (unclear as to exact method, blinding, etc)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Signs and Symptoms (LR Calculated From Data)*</th>
<th>Quality Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1 (LR = 0.20)</td>
<td>Yes</td>
</tr>
<tr>
<td>2-4 (LR = 1.14)</td>
<td>Yes</td>
</tr>
<tr>
<td>0-1 (LR = 0.20)</td>
<td>Yes</td>
</tr>
<tr>
<td>3-4 (LR = 0.70)</td>
<td>Yes</td>
</tr>
<tr>
<td>1-2 (LR = 0.11)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*LR indicates likelihood ratio; A patients with normal thyroid indices followed up; B, more than 60% of patients with abnormal thyroid indices followed up; C, explicit diagnostic criteria used.

Another approach is to examine the total number of signs and symptoms of thyroid disease, rather than the individual signs and symptoms themselves. Two studies31,32 allowed us to calculate the relationship between the number of signs and symptoms and the probability of thyroid disease (Table 6). In a review31 of 135 family practice charts, there was a good correlation between the number of clinical features and the presence of thyroid disease. This was confirmed in a second study32 in a retrospective chart review of 500 patients seen in a thyroid clinic. In both studies, the absence of signs and symptoms significantly decreased the chance of thyroid disease.

In a study of the diagnostic properties of a clinical examination for thyroid disease, we would ideally want a prospective, blinded trial that included physician-generated information (as opposed to a self-administered questionnaire) that would best replicate the usual diagnostic process. Unfortunately, none of these studies combined all such elements.

Sensitivity and Specificity of sTSH Assay for Diagnosis of Thyroid Disease

Of 30 titles in the initial search, 7 studies and 1 abstract focused on inpatients. Of these, only 211,33 described the results in sufficient detail to allow us to calculate the LRs of sTSH in hospitalized patients (Table 7). The first study by Spencer et al11 used a second-generation TSH assay with a functional lower limit of 0.1. Table 7 shows the positive LRs associated with various ranges. In a follow-up study,33 some of the same serum samples plus additional inpatient serum samples were reassayed using a third-generation TSH test with a functional lower limit of 0.01. Of patients who truly had hyperthyroidism (as judged by follow-up) and an original TSH level of less than 0.1 µIU/mL, 22 of 23 had a TSH level of less than 0.005 µIU/mL using the third-generation assay. Of those who had NTI and an original TSH level of less than 0.1 µIU/mL, only 5 of 37 had a TSH level of less than 0.005 µIU/mL, indicating that the third-generation assay had additional power in detecting inpatients who were truly hyperthyroid. The LRs much greater than 1 are associated only with the extremes of the TSH range. Normal and mildly abnormal values are very useful in ruling out disease.
We applied the available evidence to the problem of diagnosing thyroid disease in hospitalized patients using a 3-part diagnostic strategy.

The Pretest Probability

We estimated the pretest probability of thyroid disease from the prevalence of unrecognized thyroid disease among inpatients, and found a prevalence of 1.0% to 2.5% (Tables 1-4). These data have a number of limitations, including varying reference standards in the diagnosis of thyroid disease, varying thyroid assays, and inconsistent follow-up that focused mainly on selected patients with abnormal initial biochemical test results. The degree of NTI of the patients varied between studies and was often not explicitly stated. The diagnosis of thyroid disease had to be inferred from the data presented.

Despite these caveats, estimates of thyroid disease were similar, and formal statistical testing did not show heterogeneity of estimates across studies. In addition, these estimates of prevalence among inpatients agree with those of recent reviews but are lower than those of population studies. This discrepancy is probably due to the more stringent criterion of thyroid disease in our search, ie, clinical and biochemical follow-up, as well as the inclusion of subclinical and biochemical thyroid disease in the population studies. We chose to include older studies, including those that relied on free thyroxine and early-generation TSH assays. As mentioned, we believe that follow up is the appropriate gold standard in diagnosis of thyroid disease in hospitalized patients after resolution of the acute illness.

Modifying the Pretest Probability

Clinicians modify pretest probability by noting the clinical signs and symptoms. Because there were no studies of the impact of clinical signs and symptoms on probabilities in inpatients, our options were to either neglect this area altogether or broaden our focus to include outpatients. Because we believe the biological relationship between underlying thyroid disease and signs and symptoms is similar in inpatients and outpatients, we chose the latter course. However, the specificity of symptoms and signs may decrease among inpatients, and the values of LRs we found should be treated as upper-limit values. Some findings, such as goiter, may retain their specificity in the hospitalized population.

Another reason the LRs in Tables 5 and 6 may be spuriously high is that examination of a large number of signs and symptoms in a relatively small population increases the likelihood of chance associations. Even more important, the lack of blinding to test results may cause a clinician to overinterpret physical signs that he or she expects to see. Studies that depend on

### Table 7. Likelihood Ratios Associated With sTSH Results*

<table>
<thead>
<tr>
<th>sTSH, µIU/mL</th>
<th>Hyperthyroid</th>
<th>Hypothyroid</th>
<th>Likelihood Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.1</td>
<td>11</td>
<td>34</td>
<td>7.7</td>
</tr>
<tr>
<td>0.1-0.6</td>
<td>0</td>
<td>124</td>
<td>0</td>
</tr>
<tr>
<td>&gt;0.6-6.7</td>
<td>0</td>
<td>103</td>
<td>0</td>
</tr>
</tbody>
</table>

* sTSH indicates sensitive thyrotropin (thyroid-stimulating hormone); plus sign, number of patients with disease; and minus sign, number of patients without disease.

COMMENT

We applied the available evidence to the problem of diagnosing thyroid disease in hospitalized patients using a 3-part diagnostic strategy.

The Pretest Probability

We estimated the pretest probability of thyroid disease from the prevalence of unrecognized thyroid disease among inpatients, and found a prevalence of 1.0% to 2.5% (Tables 1-4). These data have a number of limitations, including varying reference standards in the diagnosis of thyroid disease, varying thyroid assays, and inconsistent follow-up that focused mainly on selected patients with abnormal initial biochemical test results. The degree of NTI of the patients varied between studies and was often not explicitly stated. The diagnosis of thyroid disease had to be inferred from the data presented.

Despite these caveats, estimates of thyroid disease were similar, and formal statistical testing did not show heterogeneity of estimates across studies. In addition, these estimates of prevalence among inpatients agree with those of recent reviews but are lower than those of population studies. This discrepancy is probably due to the more stringent criterion of thyroid disease in our search, ie, clinical and biochemical follow-up, as well as the inclusion of subclinical and biochemical thyroid disease in the population studies. We chose to include older studies, including those that relied on free thyroxine and early-generation TSH assays. As mentioned, we believe that follow up is the appropriate gold standard in diagnosis of thyroid disease in hospitalized patients after resolution of the acute illness.

Modifying the Pretest Probability

Clinicians modify pretest probability by noting the clinical signs and symptoms. Because there were no studies of the impact of clinical signs and symptoms on probabilities in inpatients, our options were to either neglect this area altogether or broaden our focus to include outpatients. Because we believe the biological relationship between underlying thyroid disease and signs and symptoms is similar in inpatients and outpatients, we chose the latter course. However, the specificity of symptoms and signs may decrease among inpatients, and the values of LRs we found should be treated as upper-limit values. Some findings, such as goiter, may retain their specificity in the hospitalized population.

Another reason the LRs in Tables 5 and 6 may be spuriously high is that examination of a large number of signs and symptoms in a relatively small population increases the likelihood of chance associations. Even more important, the lack of blinding to test results may cause a clinician to overinterpret physical signs that he or she expects to see. Studies that depend on
physical signs and symptoms elicited before diagnostic test results found much lower LRs, in the range of 2 to 3. These studies do not, however, clarify whether clinical signs and symptoms were specifically sought.

The total number of signs and symptoms appears to be a more robust indicator of thyroid disease, with LRs between 7 and 19, although these estimates may also be inflated by the lack of blinding. Some descriptive studies suggest that elderly patients with known thyroid disease have fewer signs and symptoms than younger patients, and one might expect this effect in hospitalized patients as well.

Obtaining the Posttest Probability

Acute illness influences thyroid indices such that while the sensitivity of the test is maintained (a normal value carries an LR of 0), the specificity of the sTSH assay is reduced. The majority of abnormal sTSH test results obtained during acute illness return to normal once the illness has resolved. The LR for very abnormal results (≤0.1 or >20) is in the range of 7 to 11. In contrast, the LR of an abnormal TSH test result in outpatients is close to 100, emphasizing its superior test characteristics in that population.

Examples and Recommendations

Clinical examples help illustrate the application of this information. Suppose a general medical patient (prevalence of hypothyroidism, 0.6%) has no clinical features of thyroid disease (LR in the range of 0.2) and thus has a pretest likelihood of 0.12%. A second-generation sTSH assay yields a value of less than 0.1 (LR = 7.7). The resulting posttest likelihood of hyperthyroidism is now about 1%, sufficiently low that specific action is likely unwarranted.

Further consider a patient admitted with congestive heart failure who is noted to have a goiter. Using the prevalence of hypothyroidism calculated for general medical inpatients, ie, 0.6%, and modifying this by the LR of 5 associated with a goiter, we arrive at an estimated probability of hypothyroidism of 5%. If the TSH test result is greater than 20 (LR = 11.1), posttest probability is around 40%.

These examples stress the limitations of the sTSH test in the face of very low prior probability and suggest the following approach. First, in hospitalized patients in whom pretest probability is very low, thyroid testing should be deferred until patients have recovered from their acute illness. Second, among inpatients with numerous signs and symptoms of thyroid disease (data would support a value of at least 5, perhaps less in elderly patients) or inpatients with signs particular to thyroid disease, eg, goiter and eye signs, sTSH testing is appropriate. Third, clinicians should use the pretest-posttest framework to interpret sTSH test results.

We reviewed the prevalence of undiagnosed thyroid disease among inpatients and the LRs associated with various signs and symptoms of thyroid disease and clarified characteristics of the sTSH test in hospitalized patients. These data should help clinicians better interpret the physical examination and laboratory results that lead to the diagnosis of thyroid disease in patients whose clinical picture is confounded by acute illness.

Out thanks to John Booth, MD, FRCPC, for comments on an earlier version of the manuscript, and to Ms Deborah Maddock for assistance with manuscript preparation.

Accepted for publication July 14, 1998.

Reprints: Gordon Guyatt, MD, MSc, FRCPC, Clinical Epidemiology & Biostatistics, McMaster University Health Sciences Centre, 1200 Main St W, Room 2C12, Hamilton, Ontario, Canada L8N 3Z5 (e-mail: guyatt@fhs.csu.mcmaster.ca).

REFERENCES


22. Cohen KL, Swigart ME. Thyroid function screen-
ing in psychiatric patients. JAMA. 1979;242:254-257.