Screening for Thyroid Disease

Recommendation Statement

U.S. Preventive Services Task Force

This statement summarizes the current U.S. Preventive Services Task Force (USPSTF) recommendations on screening for thyroid disease and the supporting scientific evidence, and updates the 1996 recommendations contained in the *Guide to Clinical Preventive Services, Second Edition: Periodic Updates*.[1] Explanations of the ratings and of the strength of overall evidence are given in Appendix A and Appendix B, respectively. The complete information on which this statement is based, including evidence tables and references, is available in the Systematic Evidence Review "Screening for Thyroid Disease,"[2] available through the USPSTF Web site (http://www.preventiveservices.ahrq.gov) and through the National Guideline Clearinghouse™ (http://www.guideline.gov). The recommendation statement and summary article are also available through the USPSTF Web site and in print through the AHRQ Publications Clearinghouse (call 1-800-358-9295 or e-mail ahrqpubs@ahrq.gov).

Recommendations made by the USPSTF are independent of the U.S. Government. They should not be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

This article first appeared in *Ann Intern Med.* 2004;140:125–127.

Summary of Recommendation

The U.S. Preventive Services Task Force (USPSTF) concludes the evidence is insufficient to recommend for or against routine screening for thyroid disease in adults. **I recommendation.**

The USPSTF found fair evidence that the thyroid stimulating hormone (TSH) test can detect subclinical thyroid disease in people without symptoms of thyroid dysfunction, but poor evidence that treatment improves clinically important outcomes in adults with screen-detected thyroid disease. Although the yield of screening is greater in certain high-risk groups (eg, postpartum women, people with Down syndrome, and the elderly), the USPSTF found poor evidence that screening these groups leads to clinically important benefits. There is the potential for harm caused by false positive screening tests; however, the magnitude of harm is not known. There is good evidence that overtreatment with levothyroxine occurs in a substantial proportion of patients, but the long-term harmful effects of overtreatment are not known. As a result, the USPSTF could not determine the balance of benefits and harms of screening asymptomatic adults for thyroid disease.

Clinical Considerations

- Subclinical thyroid dysfunction is defined as an abnormal biochemical measurement of thyroid hormones without any specific clinical signs or symptoms of thyroid disease and no history of thyroid dysfunction or therapy. This includes individuals who have mildly elevated TSH and normal thyroxine (T4) and triiodothyronine (T3) levels (subclinical hypothyroidism), or low TSH and normal T4 and T3 levels (subclinical hyperthyroidism). Individuals with symptoms of thyroid dysfunction, or those with a history of thyroid disease or treatment, are excluded from this definition and are not the subject of these recommendations.
- When used to confirm suspected thyroid disease in patients referred to a specialty endocrine clinic, TSH has a high sensitivity (98%) and specificity (92%). When used for screening primary care

Corresponding Author: Alfred O. Berg, MD, MPH, Chair, U.S. Preventive Services Task Force, c/o Project Director, USPSTF, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, e-mail: uspstf@ahrq.gov.

populations, the positive predictive value (PPV) of TSH in detecting thyroid disease is low; further, the interpretation of a positive test result is often complicated by an underlying illness or by frailty of the individual. In general, values for serum TSH below 0.1 mU/L are considered low and values above 6.5 mU/L are considered elevated.

- Clinicians should be aware of subtle signs of thyroid dysfunction, particularly among those at high risk. People at higher risk for thyroid dysfunction include the elderly, postpartum women, those with high levels of radiation exposure (> 20 mGy), and patients with Down syndrome. Evaluating for symptoms of hypothyroidism is difficult in patients with Down syndrome because some symptoms and signs (eg, slow speech, thick tongue, and slow mentation) are typical findings in both conditions.
- Subclinical hyperthyroidism has been associated with atrial fibrillation, dementia, and, less clearly, with osteoporosis. However, progression from subclinical to clinical disease in patients without a history of thyroid disease is not clearly established.
- Subclinical hypothyroidism is associated with poor obstetric outcomes and poor cognitive development in children. Evidence for dyslipidemia, atherosclerosis, and decreased quality of life in adults with subclinical hypothyroidism in the general population is inconsistent and less convincing.

Discussion

Subclinical thyroid disease is much more common than overt disease in primary care populations. Up to 5% of women and 3% of men have subclinical hypothyroidism[3]; prevalence increases with age and is greater in whites than in blacks.[3,4] Untreated hypothyroidism can lead to fatigue, weight gain, mental slowing, heart failure, and elevated lipid levels. Subclinical hyperthryroidism is less common, occurring in 1% of men over age 60 and 1.5% of women over age 60.[5] Untreated hyperthyroidism can lead to atrial fibrillation, congestive heart failure, osteoporosis, and neuropsychiatric problems.

The goal of screening for thyroid disease is to identify and treat patients at risk for the health consequences of thyroid dysfunction before they become clinically apparent. The USPSTF examined the evidence for screening people who have no known history of thyroid disease and no or few signs or symptoms. The USPSTF found no controlled studies that examined whether routine screening for thyroid disease in the primary care setting leads to improved symptoms or other health outcomes.

Screening for thyroid dysfunction can be performed using the medical history, physical examination, or any of several serum thyroid function tests. The TSH is usually recommended because it can detect abnormalities before other tests become abnormal. When used to confirm suspected disease in patients referred to an endocrine specialty clinic, the TSH test has a sensitivity above 98% and a specificity greater than 92% for the clinical and functional diagnosis.[6] The accuracy of TSH screening in primary care patients is difficult to evaluate, as TSH is often considered the "gold standard" for assessing thyroid function. A severe non-thyroid disease can lead to a false positive TSH test result; in a recent systematic review of TSH used to screen patients admitted to acute care and geriatric hospitals, the PPV of a low TSH was 0.24 for hyperthyroidism and 0.06 for hypothyroidism.[7] In screening programs, patients who have normal thyroxine levels but have mild elevations of TSH or low TSH levels often revert to normal over time.[8–10]

In a review of observational studies conducted among groups of people who were exposed to radiation from nuclear fallout, high doses of I^{131} ($\geq$ 20 mGy) were associated with hypothyroidism (Helfand, unpublished data, 2003). There is fair evidence to suggest that exposure to I^{131} increases the risk for developing thyroid antibodies, which may lead to autoimmune thyroid disease. The evidence for an increased risk for hypothyroidism from exposure to low doses of I^{131} (under 20 mGy) is inconclusive, making it difficult to ascertain whether there is a threshold dose of I^{131} exposure that confers no added risk for hypothyroidism. Thus, clinicians need to be extra vigilant in this high risk group for the possibility of thyroid dysfunction.

The USPSTF found no randomized trials of treatment for overt or subclinical hyperthyroidism. Limited evidence is available from observational studies regarding cardiac function parameters pre- and post-treatment, but no studies reported clinical outcomes.

The USPSTF found 14 randomized trials of treatment using levothyroxine (LT4) therapy,[2] but few were relevant to the issue of treating screen-detected subclinical hypothyroidism in primary care clinical settings, and most of the potentially relevant studies suffered from significant design flaws. Most studies were in groups of patients known to have thyroid disease (eg, patients with Hashimoto's thyroiditis or Graves' disease). Results from 3 small studies of women with screen-detected subclinical hypothyroidism showed no improved clinical outcomes (in 2 of the studies) and some improved clinical outcome in the third.[11–13] Data from observational studies, most of which were judged to be of poor quality, showed that treatment of subclinical hypothyroidism leads to modestly improved serum lipid levels.[14] No trials of treatment of subclinical hypothyroidism in pregnant women were identified.

A potential benefit of treating subclinical hypothyroidism is to prevent the spontaneous development of overt hypothyroidism, but this potential benefit has not been studied in clinical trials. If the potential benefit suggested by data from a longitudinal survey is real, the USPSTF estimates that in a reference population of 1,000 women screened, 3 cases of overt hypothyroidism would be prevented in 5 years, but 40 people would have taken medication for 5 years without a clear benefit.[2]

The potential harms of screening and treatment are principally the adverse effects of antithyroid drugs, radioiodine, thyroid surgery, and thyroid replacement therapy if detection and early treatment for subclinical disease are unnecessary. People with a false positive TSH test result (more common in those with a severe underlying illness or those who are frail or elderly) may be subjected to unnecessary treatment or may have adverse psychological consequences (eg, labeling). The USPSTF reviewed only the adverse effects of LT4 replacement therapy for mild thyroid failure and potential adverse effects of long-term treatment. These adverse effects were not carefully assessed in the randomized trials. Although some studies have suggested that women with a low TSH as a result of taking thyroid hormone replacement are at higher risk for developing osteoporosis,[15,16] a recent systematic review did not support this finding.[17] Overtreatment with LT4 is a potential risk: about 1 in 4 patients receiving LT4 are maintained unintentionally on doses sufficient to fully suppress TSH.[18,19] Data from the Framingham Study suggest that 1 excess case of atrial fibrillation might occur for every 114 patients treated with LT4 sufficient to suppress TSH.[19]

Recommendations of Other Groups

The American Thyroid Association recommends measuring thyroid function in all adults beginning at age 35 years and every 5 years thereafter, noting that more frequent screening may be appropriate in high-risk or symptomatic individuals.[20] The Canadian Task Force on the Periodic Health Examination recommends maintaining a high index of clinical suspicion for nonspecific symptoms consistent with hypothyroidism when examining perimenopausal and postmenopausal women.[21] The American College of Physicians recommends screening women older than age 50 with 1 or more general symptoms that could be caused by thyroid disease.[22] The American Association of Clinical Endocrinologists recommends TSH measurement in women of childbearing age before pregnancy or during the first trimester.[23] The American College of Obstetricians and Gynecologists recommends that physicians be aware of the symptoms and risk factors for postpartum thyroid dysfunction and evaluate patients when indicated.[24] The American Academy of Family Physicians recommends against routine thyroid screening in asymptomatic patients younger than age 60.[25]

References

1. U.S. Preventive Services Task Force. *Guide to Clinical Preventive Services*. 2nd ed. Alexandria, Virginia: International Medical Publishing, Inc.; 1996:209–218.

2. Helfand, M. *Screening for Thyroid Disease*. Systematic Evidence Review No. 23. (Prepared by the Oregon Health & Science University Evidence-based Practice Center under Contract No. 290-97-0018). Rockville, MD: Agency for Healthcare Research and Quality. (Available on the AHRQ Web site at: www.ahrq.gov/clinic/serfiles.htm).

3. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the

United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab*. 2002;87(2):489–499.

4. Tunbridge WM, Vanderpump MP. Population screening for autoimmune thyroid disease. *Endocrinol Metab Clin North Am*. 2000;29(2):239–253.
5. Helfand M, Redfern CC. Clinical guideline, part 2. Screening for thyroid disease: an update. American College of Physicians [published erratum appears in *Ann Intern Med.* 1999 Feb 2;130(3):246]. *Ann Intern Med*. 1998;129(2):144–158.
6. Helfand M, Crapo LM. Testing for suspected thyroid disease. In: Sox HC, ed. *Common Diagnostic Tests*. Philadelphia: American College of Physicians; 1990.
7. Attia J, Margetts P, Guyatt G. Diagnosis of thyroid disease in hospitalized patients. A systematic review. *Arch Intern Med*. 1999;159:658–665.
8. Jaeschke R, Guyatt G, Gerstein H, et al. Does treatment with L-thyroxine influence health status in middle-aged and older adults with subclinical hypothyroidism? *J Gen Intern Med*. 1996;11(12):744–749.
9. Parle J, Franklyn J, Cross K, Jones S, Sheppard M. Prevalence and follow-up of abnormal thyrotropin (TSH) concentrations in the elderly in the United Kingdom. *Clin Endocrinol (Oxf)*. 1991;34:77–83.
10. Sundbeck G, Jagenburg R, Johansson PM, Eden S, Lindstedt G. Clinical significance of a low serum thyrotropin concentration by chemiluminometric assay in 85-year-old women and men. *Arch Inten Med*. 1991;151:549–556.
11. Nystrom E, Caidahl K, Fager G, Wikkelso C, Lundberg PA, Lindstedt G. A double-blind cross-over 12-month study of L-thyroxine treatment of women with 'subclinical' hypothyroidism. *Clin Endocrinol*. 1988;29(1):63–75.
12. Ross DS. Bone density is not reduced during the short-term administration of levothyroxine to postmenopausal women with subclinical hypothyroidism: a randomized, prospective study. *Am J Med*. 1993;95(4):385–388.
13. Monzani F, Di Bello V, Caraccio N, et al. Effect of levothyroxine on cardiac function and structure in subclinical hypothyroidism: a double blind, placebo-controlled study. *J Clin Endocrinol Metab*. 2001;86(3):1110–1115.
14. Danese MD, Ladenson PW, Meinert CL, Powe NR. Clinical review 115: effect of thyroxine therapy on serum lipoproteins in patients with mild thyroid failure: a quantitative review of the literature. *J Clin Endocrinol Metab*. 2000;85(9):2993–3001.
15. Uzzan B, Campos J, Cucherat M, Nony P, Boissel JP, Perret GY. Effects on bone mass of long term treatment with thyroid hormones: a meta-analysis. *J Clin Endocrinol Metab*. 1996;81:4278–4289.
16. Faber J, Galloe AM. Changes in bone mass during prolonged subclinical hyperthyroidism due to L-thyroxine treatment: a meta-analysis. *Eur J Endocrinol*. 1994;130(4):350–356.
17. Greenspan SL, Greenspan FS. The effect of thyroid hormone on skeletal integrity. *Ann Intern Med*. 1999;130(9):750–758.
18. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med*. 2000;160(4):526–534.
19. Sawin C, Geller A, Wolf P. Low serum thyrotropin concentrations as a risk factor for atrial fibrillation in older persons. *N Engl J Med*. 1994;331:1249–1252.
20. Ladenson PW, Singer PA, Ain KB, et al. American Thyroid Association guidelines for detection of thyroid dysfunction. [erratum appears in *Arch Intern Med.* 2001 Jan 22;161(2):284]. *Arch Intern Med*. 2000;160(11):1573–1575.
21. Canadian Task Force on the Periodic Health Examination. Canadian Guide to Clinical Preventive Health Care. Ottawa: Canada Communication Group; 1994:611–618.
22. American College of Physicians. Clinical guideline, part 1. Screening for thyroid disease. *Ann Intern Med*. 1998;129(2):141–143.
23. AACE Thyroid Task Force. American Association of Clinical Endocrinologists medical guidelines for clinical practice for the evaluation and treatment of hyperthyroidism and hypothyroidism. *Endocrine Prac.* 2002;8:457–469.
24. American College of Obstetricians and Gynecologists. Thyroid disease in pregnancy. Technical Bulletin no. 37. Washington, DC: American College of Obstetricians and Gynecologists, 2002.
25. American Academy of Family Physicians. Summary of Policy Recommendations for Periodic Health Examinations. Leawood, KS: American Academy of Family Physicians; 2002.

Appendix A
U.S. Preventive Services Task Force—Recommendations and Ratings

The Task Force grades its recommendations according to one of 5 classifications (A, B, C, D, I) reflecting the strength of evidence and magnitude of net benefit (benefits minus harms):

A. The USPSTF strongly recommends that clinicians provide [the service] to eligible patients. *The USPSTF found good evidence that [the service] improves important health outcomes and concludes that benefits substantially outweigh harms.*

B. The USPSTF recommends that clinicians provide [the service] to eligible patients. *The USPSTF found at least fair evidence that [the service] improves important health outcomes and concludes that benefits outweigh harms.*

C. The USPSTF makes no recommendation for or against routine provision of [the service]. *The USPSTF found at least fair evidence that [the service] can improve health outcomes but concludes that the balance of benefits and harms is too close to justify a general recommendation.*

D. The USPSTF recommends against routinely providing [the service] to asymptomatic patients. *The USPSTF found at least fair evidence that [the service] is ineffective or that harms outweigh benefits.*

I. The USPSTF concludes that the evidence is insufficient to recommend for or against routinely providing [the service]. *Evidence that [the service] is effective is lacking, of poor quality, or conflicting and the balance of benefits and harms cannot be determined.*

Appendix B
U.S. Preventive Services Task Force—Strength of Overall Evidence

The USPSTF grades the quality of the overall evidence for a service on a 3-point scale (good, fair, poor):

Good: Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes.

Fair: Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes.

Poor: Evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes.

AHRQ Pub. No. 04-0530A
January 2004