Screening and Treatment of Subclinical Hypothyroidism or Hyperthyroidism
Comparative Effectiveness Review
Number 24

Screening and Treatment of Subclinical Hypothyroidism or Hyperthyroidism

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. HHSA 290-2007-10057-I

Prepared by:
Oregon Evidence-based Practice Center
Portland, OR

Investigators:
Bruin Rugge, M.D.
Howard Balshem, M.S.
Raj Sehgal, M.D.
Rose Relevo, M.L.S.
Paul Gorman, M.D.
Mark Helfand, M.D.

AHRQ Publication No. 11(12)-EHC033-EF
October 2011
This report is based on research conducted by the Oregon Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA 290-2007-10057-I). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted for which further reproduction is prohibited without the specific permission of copyright holders.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the Children’s Health Insurance Program (CHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews (CERs) of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strengths and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see http://effectivehealthcare.ahrq.gov/reference/purpose.cfm.

AHRQ expects that CERs will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family’s health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Kim Marie Wittenberg, M.A.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Acknowledgments

We would like to thank Kim Marie Wittenberg, M.A., Christine Chang, M.D., Mary Barton, M.D., and David Hickam, M.D., for their helpful comments on early versions of this report, and Sasha Walia for administrative support of the project.
Effectiveness of Screening and Treatment of Subclinical Hypothyroidism or Hyperthyroidism

Structured Abstract

Objectives. This report focused on four questions:
1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?
2. What are the harms of screening for subclinical thyroid dysfunction?
3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?
4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism?

Data Sources. A research librarian searched MEDLINE, the Cochrane Register Database of Systematic Reviews, and the Database of Abstracts of Reviews of Effects from inception to May 2010 for systematic reviews. Additionally, MEDLINE, AGELINE (AARP.org), EMBASE, the Cochrane Central Register of Controlled Trials, CINAHL, and World Health Organization’s Global Health Library from 2002 to May 2010 were searched for new studies. Finally, additional materials were sought by searching for regulatory information, clinical trial registries, conference proceedings, and other sources of gray literature.

Review Methods. Studies were selected based on predetermined eligibility criteria, with two investigators reviewing abstracts and full articles. English and non-English studies were eligible for inclusion. Randomized controlled trials (RCTs) and observational studies were used to determine the benefits of screening, and RCTs, controlled trials, cohort studies, case-controlled studies, and observational studies were reviewed to determine screening harms and treatment benefits and harms. Two investigators assessed study data and quality.

Main Results. Three systematic reviews met our inclusion criteria; none found screening or treating subclinical thyroid dysfunction to be beneficial. Since 2002, no studies have directly assessed the benefits or harms of screening. Six good- to fair-quality studies found that treating subclinical hypothyroidism did not improve quality of life, blood pressure, or body mass index (BMI). The findings regarding lipids were inconsistent, with two studies showing a small benefit to treatment and two studies finding no benefit to treatment. Two small poor-quality studies that evaluated the benefits of treating subclinical hyperthyroidism met our inclusion criteria. One study found a small change in the mean daytime systolic blood pressure, while the other found a small increase in BMI with treatment. Treatment harms were neither systematically evaluated nor well described.

Conclusions. Currently there are no studies that evaluate the benefits and harms of screening for subclinical thyroid dysfunction in the primary care setting. Studies of treatment tend to be small and of short duration, and they have failed to demonstrate improvement in quality of life, blood pressure, or weight. The data concerning lipids is inconsistent, but at best, treatment might cause a modest (about 5-percent) improvement in lipid measurements. The lack of any formal data on the harms of treatment makes it difficult to balance the benefits against the harms of treatment. Further research is needed to determine if screening and/or treating subclinical thyroid dysfunction is beneficial or harmful.
# Contents

Executive Summary .................................................................................................................. ES-1  
Introduction .............................................................................................................................. 1  
  Background ............................................................................................................................ 1  
  Description of Condition ....................................................................................................... 1  
Prevalence and Course of Mild Thyroid Dysfunction ............................................................. 4  
Strategies for Detecting and Managing Subclinical Thyroid Dysfunction ......................... 5  
  Screening Strategies ............................................................................................................ 5  
  Management Strategies ...................................................................................................... 5  
Guidelines on Early Detection and Treatment ....................................................................... 6  
Rationale for Screening and Treatment ................................................................................... 8  
Scope of Review .................................................................................................................... 11  
Key Questions and Analytic Framework ............................................................................... 11  
Methods ................................................................................................................................ 14  
  Topic Development ............................................................................................................ 14  
  Search Strategy .................................................................................................................. 14  
  Selection of Systematic Reviews ....................................................................................... 14  
Data Abstraction and Quality Rating of Systematic Reviews ............................................... 15  
  Individual Study Selection .................................................................................................. 16  
Data Abstraction and Quality Rating of Individual Studies ................................................. 17  
Data Synthesis ..................................................................................................................... 17  
External Review .................................................................................................................... 17  
Results .................................................................................................................................. 18  
  Systematic Reviews ............................................................................................................ 18  
  Individual Studies ............................................................................................................... 20  
Key Question 1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality? ...................................................................................................................... 24  
Key Question 2. What are the harms of screening for subclinical thyroid dysfunction? .......... 24  
Key Question 3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes? .................................... 26  
  Subclinical Hypothyroidism ................................................................................................ 26  
  Subclinical Hyperthyroidism ............................................................................................... 37  
Key Question 4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism? ......................................................................................... 43  
Discussion .............................................................................................................................. 45  
  Benefits of Screening/Treatment ....................................................................................... 45  
  Harms of Screening/Treatment (Key Questions 2 and 4) ..................................................... 51  
Strength of Evidence ............................................................................................................. 51  
  Risk of Bias ........................................................................................................................ 51  
  Consistency ......................................................................................................................... 51  
  Directness .......................................................................................................................... 52  
  Applicability ...................................................................................................................... 52  
  Precision ............................................................................................................................. 52  
Limitations .............................................................................................................................. 52  
Summary of Review Findings ............................................................................................... 53  
Emerging Issues/Next Steps ................................................................................................. 54
Future Research ........................................................................................................................ 54
References ........................................................................................................................................ 56
Abbreviations .................................................................................................................................. 61

Tables
Table A. Summary of Evidence for Subclinical Hypothyroidism .............................................. ES-7
Table B. Summary of Evidence for Subclinical Hyperthyroidism ............................................... ES-9
Table 1. Symptoms and Signs of Overt Thyroid Dysfunction .................................................... 2
Table 2. Classification of Thyroid Dysfunction: Biochemical Definition .................................. 2
Table 3. Evidence for the Association of Subclinical Hypothyroidism (TSH 4.5-10 mIU/L) and Adverse Health Outcomes and Quality of Evidence for Risks and Benefits of Treatment: Findings of the American Association of Clinical Endocrinologists, the American Thyroid Association, and the Endocrine Society Panel ........................................................... 7
Table 4. Included Systematic Reviews ....................................................................................... 15
Table 5. Criteria for Inclusion/Exclusion of Studies in the Review ............................................ 16
Table 6. Systematic Review Study Concordance ....................................................................... 20
Table 7. Screening Evidence Table ........................................................................................... 25
Table 8. Treatment Evidence Table (Subclinical Hypothyroidism) .......................................... 27
Table 9. Quality of Life (Subclinical Hypothyroidism) ............................................................... 29
Table 10. Weight (kg)/Body Mass Index (BMI) (kg/m²) (Subclinical Hypothyroidism) ............ 30
Table 11. Blood Pressure (mmHg) (Subclinical Hypothyroidism) ............................................ 31
Table 12. Total Cholesterol (mg/dL) (Subclinical Hypothyroidism) ........................................... 32
Table 13. Low-Density Lipoprotein (LDL) Cholesterol (mg/dL) (Subclinical Hypothyroidism) .......................................................................................................................... 33
Table 14. High-Density Lipoprotein (HDL) Cholesterol (mg/dL) (Subclinical Hypothyroidism) .......................................................................................................................... 34
Table 15. Triglycerides (mg/dL) (Subclinical Hypothyroidism) .................................................. 35
Table 16. Treatment Evidence Table (Subclinical Hyperthyroidism) ........................................ 38
Table 17. Body Mass Index (BMI) (kg/m²) (Subclinical Hyperthyroidism) ................................. 39
Table 18. Blood Pressure (mmHg) (Subclinical Hyperthyroidism) ............................................ 40
Table 19. Bone Mineral Density (g/cm²) (Subclinical Hyperthyroidism) ................................... 41
Table 20. Lipids (mg/dL) (Subclinical Hyperthyroidism) ............................................................ 42
Table 21. Summary of Evidence for Subclinical Hypothyroidism ............................................. 46
Table 22. Summary of Evidence for Subclinical Hyperthyroidism .......................................... 48

Figures
Figure A. Analytic Framework ........................................................................................................ ES-2
Figure 1. Analytic Framework ......................................................................................................... 13
Figure 2. Systematic Review Literature Flow Diagram ................................................................. 19
Figure 3. Individual Study Literature Flow Diagram ........................................................................ 22
Figure 4. Foreign Language Literature Flow Diagram ...................................................................... 23

Appendixes
Appendix A. Lipid Conversion Factors
Appendix B. Search Strategies
Appendix C. Criteria for Assessing Internal Validity of Individual Studies
Appendix D. Excluded Studies
Appendix E. Screening and Case-Finding Recommendations of Other Groups
Appendix F. Results of Search for Systematic Reviews
Executive Summary

Background

Mildly elevated or decreased serum thyroid stimulating hormone (TSH, also called thyrotropin) levels are the most common abnormalities related to thyroid function. Subclinical thyroid dysfunction, defined as an abnormal TSH with normal levels of serum thyroid hormones (T3 and T4), affects 5 percent of women and 3 percent of men. Subclinical hypothyroidism is defined as a high TSH and normal T3/T4, and subclinical hyperthyroidism as having a low or undetectable TSH and normal T3/T4. Subclinical thyroid dysfunction has been shown to be a risk factor for the later development of overt thyroid disease. In addition, a high TSH level may be a risk factor for coronary events, elevated cholesterol levels, and increased rates of congestive heart failure, while a low TSH level is a risk factor for atrial fibrillation and osteoporosis. Therefore, it has been proposed that screening for and treating subclinical thyroid dysfunction might lead to a decrease in the morbidity associated with overt thyroid disease, heart disease, and possibly osteoporosis. To date, evidence-based reviews have recommended against routine screening and treatment of subclinical thyroid dysfunction, primarily based on the lack of evidence that treating subclinical thyroid dysfunction improves patient-centered outcomes. However, some experts, while acknowledging that evidence to support treatment is lacking, suggest that screening could decrease morbidity and mortality, and perceive the potential for harm as both minor and preventable. They argue it would be best to screen for and treat subclinical thyroid dysfunction until there are sufficient data to address this question definitively.

Topic Development

This topic was nominated by the public as part of the Effective Health Care (EHC) Program of the Agency for Healthcare Research and Quality (AHRQ). The U.S. Preventive Services Task Force (USPSTF) also expressed interest in updating its 2004 recommendations. After receiving an initial set of key questions and an analytic framework for the process of screening and treating subclinical hypothyroidism and subclinical hyperthyroidism from AHRQ, we revised the key questions and analytic framework with the input of external technical experts, members of the USPSTF, and additional input from AHRQ (Figure A). After this input, AHRQ personnel approved the scope and key questions for the report.

A 2004 review of screening for thyroid disease for the USPSTF established that subclinical thyroid dysfunction is quite prevalent, may be responsible for morbidity, and can be detected with a serum TSH assay, a readily available, reliable, and acceptable test. However, in 2004, it remained unclear whether treating subclinical thyroid dysfunction would reduce morbidity. Consequently, this current review focuses on whether new evidence demonstrates that treatment improves clinically important outcomes in adults with screen-detected thyroid disease.
Figure A. Analytic framework

Key Questions
1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?
2. What are the harms of screening?
3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?
4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism?

Note: A. fibrillation = atrial fibrillation; CAD = coronary artery disease; CHF = congestive heart failure; DEXA = dual-energy x-ray absorptiometry; NNS = number needed to screen; NNT = number needed to treat; S = subclinical; TSH = thyroid stimulating hormone.
Key Questions

We reviewed published studies to answer the following key questions:

**Key Question 1.** Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?

**Key Question 2.** What are the harms of screening? Specifically, how frequently and how severely do patients screened for subclinical thyroid dysfunction experience adverse psychological impacts or other harms of workup from screening?

**Key Question 3.** Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes? We were primarily interested in the comparative effectiveness of a strategy of routine treatment versus active surveillance to prevent the possible complications of untreated subclinical thyroid dysfunction.

**Key Question 4.** What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism? Specifically, what are the consequences of overtreatment, including effects on bone mineral density and incidence of atrial fibrillation, and how frequently do they occur?

**Methods**

**Data Sources**

A research librarian searched MEDLINE, the Cochrane Register of Systematic Reviews, and the Database of Abstracts of Reviews of Effects to identify systematic reviews of screening and treatment of subclinical hypothyroidism or subclinical hyperthyroidism, with no limits on dates. A number of systematic reviews have addressed this topic. Three reviews—Helfand (2004), the USPSTF review; Surks et al. (2004); and Villar et al. (2007), a Cochrane review—adequately reflected the state of the evidence through 2006. Three reviews had largely concordant findings.

We also searched MEDLINE, AGELINE (AARP.org), EMBASE (embase.com) and the Cochrane Central Register of Controlled Trials to identify studies regarding screening and treatment of subclinical hypothyroidism or subclinical hyperthyroidism published from 2002 to May 2010. Additional materials were sought by searching for regulatory information, clinical trial registries, conference proceedings, and other sources of gray literature. Additional studies were identified from citations in relevant articles, discussions with experts, and requests to pharmaceutical companies.

We also performed a supplementary search of the foreign language literature. For this search we included CINAHL and the World Health Organization Global Health Library.

**Study Selection**

We defined the target population as community-living nonpregnant adults without a history of thyroid disease or symptoms of overt hypothyroidism or hyperthyroidism who are representative of adults who might be seen in primary care settings. Intermediate outcomes of interest for subclinical hypothyroidism were lipid levels and blood pressure; intermediate outcomes of interest for subclinical hyperthyroidism were bone mineral density and blood
pressure. Final outcomes of interest for subclinical hypothyroidism were weight change; measures of well-being, including but not limited to cognition and memory; cardiovascular morbidity; and progression to overt disease. Final outcomes of interest for subclinical hyperthyroidism were weight change; measures of well-being, including but not limited to cognition and memory; cardiovascular morbidity; progression to overt disease; fractures; and atrial fibrillation.

The quality of each systematic review was assessed using the Assessment of Multiple Systematic Reviews (AMSTAR) checklist. The quality of each study was assessed using criteria established by the USPSTF, and poor quality studies were excluded from review except for subclinical hyperthyroidism, for which no good or fair quality studies were found. Information regarding the population, setting, treatments, and outcomes was all abstracted.

Data Synthesis

We assessed overall strength for each body of evidence addressing a particular outcome of each key question using the guidance from the Strength of Evidence chapter of the AHRQ Effective Health Care Program Methods Guide for Effectiveness and Comparative Effectiveness Reviews. To assign an overall strength of evidence (high, moderate, low, or insufficient), we considered the number, quality, and size of studies; consistency of results between studies; and directness of evidence.

Findings

Findings are summarized in Tables A and B.

Key Question 1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?

We identified no randomized controlled trials (RCTs) or observational studies comparing the outcomes of screening versus not screening for subclinical hypothyroidism or subclinical hyperthyroidism in the general population.

Key Question 2. What are the harms of screening? Specifically, how frequently and how severely do patients screened for subclinical thyroid dysfunction experience adverse psychological impacts or other harms of workup from screening?

Information about the harms of screening remains sparse. We identified no RCTs or controlled observational studies that evaluated harms associated with screening for subclinical thyroid dysfunction. Two natural history studies suggest that a significant number of individuals with subclinical thyroid dysfunction will have normal thyroid function if followed up at about 3 years.

Key Question 3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?

Taken together, the 3 reviews listed earlier evaluated 14 controlled trials of treatment for subclinical hypothyroidism. The most recent of the three studies, a good-quality Cochrane review, found that, while in the short term lipid profiles and left ventricular function may improve after treatment with levothyroxine, treatment did not improve health-related quality of
life or symptoms, and the trials were not suitable to assess survival or cardiovascular mortality and morbidity. The other two previous reviews had similar results.

Six trials of treatment for subclinical hypothyroidism were published from 2002 to 2010; none of these were included in the 2004 USPSTF review,¹ and four were not included in the 2007 Cochrane review.² None of the trials evaluated long-term cardiovascular outcomes, and none had more than 1 year of followup, the minimum time needed to compare immediate treatment versus a strategy of active surveillance with annual testing. The largest trial included 120 patients. The trials used different TSH cutoffs to diagnose subclinical thyroid dysfunction and different dosages of medication, and they assessed effects over different time periods. None of the included studies were conducted on U.S. populations; most patients included in the studies were recruited from specialty clinics rather than from primary care settings.

Four trials evaluated the effect of treatment on lipids; they had inconsistent findings. In two of the four studies, there were modest reductions in total cholesterol and low density lipoprotein (LDL), but in two others, there was no improvement for any lipoproteins. Two other studies measured well-being; two looked at blood pressure; and four examined changes in weight or body mass index (BMI). Studies evaluating these measures consistently found no evidence of benefit.

Two controlled trials assessed the efficacy of treatment of subclinical hyperthyroidism; both of these were of poor quality. Both assessed changes in blood pressure. One also evaluated change in BMI. The other evaluated patient-reported fatigue, nervousness, sweating, change in appetite, and tremors; lipids; and bone mineral density. Evidence of efficacy was inconsistent.

**Key Question 4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism? Specifically, what are the consequences of overtreatment, including effects on bone mineral density and incidence of atrial fibrillation, and how frequently do they occur?**

None of the studies of treatment for either subclinical hypothyroidism or subclinical hyperthyroidism systematically evaluated harms. An assessment of harms was likely not a part of any of the studies’ protocol, nor does it appear that study participants were provided with a list of potential harms and asked to identify those that they experienced. Some patients initially diagnosed with subclinical hypothyroidism may revert to normal levels of TSH without treatment, suggesting unnecessary treatment as a possible harm from therapy with levothyroxine.
Table A. Summary of evidence for subclinical hypothyroidism

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study type: Number of studies</th>
<th>Risk of bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect (Strength of evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?</td>
<td>No studies directly compared screening with an alternative strategy for detecting thyroid dysfunction</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>KQ 2. What are the harms of screening for subclinical thyroid dysfunction?</td>
<td>2 studies had indirect evidence about overdiagnosis</td>
<td>NA</td>
<td>No major inconsistency</td>
<td>Indirect; low to moderate applicability to primary care</td>
<td>NA</td>
<td>No RCTs were identified; two natural history studies demonstrating indirect evidence of potential harm were included</td>
<td>Insufficient</td>
</tr>
<tr>
<td>KQ 3. Does treatment of patients with subclinical hypothyroidism detected by screening affect outcomes?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>Cardiovascular events, coronary artery disease, and heart failure</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>Overall quality of life</td>
<td>RCT: 2 169</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; moderate applicability to primary care setting</td>
<td>Imprecise: small studies of 100 and 69 subjects</td>
<td>No effect</td>
<td>No effect (Low)</td>
</tr>
<tr>
<td>Key Question</td>
<td>Study type: Number of studies</td>
<td>Risk of bias</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Comments</td>
<td>Magnitude of effect (Strength of evidence)</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Changes in mood/cognition</td>
<td>RCT: 2 169</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; moderate applicability to primary care setting</td>
<td>Imprecise: small studies of 100 and 69 subjects</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Weight/BMI changes</td>
<td>RCT: 4 305</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect</td>
<td>Imprecise; the largest study had 100 subjects; the smallest had 23</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Blood pressure changes</td>
<td>RCT: 2 195</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; low applicability to asymptomatic patients in primary care settings</td>
<td>Imprecise</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Changes in lipid levels</td>
<td>RCT: 4 379</td>
<td>Medium</td>
<td>Inconsistent</td>
<td>Indirect; low applicability to asymptomatic patients in U.S. primary care settings</td>
<td>Imprecise</td>
<td>Small effect for LDL and total cholesterol (Low)</td>
<td></td>
</tr>
<tr>
<td>KQ4. What are the harms of treatment of subclinical hypothyroidism?</td>
<td>None of the trials in this update addressed the harms of levothyroxine treatment</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Insufficient</td>
<td></td>
</tr>
</tbody>
</table>

Note: BMI=body mass index; LDL=low-density lipoprotein; NA=not applicable; RCT=randomized controlled trial.
Table B. Summary of evidence for subclinical hyperthyroidism

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study type: Number of studies</th>
<th>Number of subjects</th>
<th>Risk of bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect (Strength of evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>KQ 2. What are the harms of screening for subclinical thyroid dysfunction?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>KQ 3. Does treatment of patients with subclinical hyperthyroidism detected by screening affect outcomes?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular events, including angina, atrial fibrillation, and other clinically significant arrhythmias</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Fractures</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Overall quality of life</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Changes in mood/cognition</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Key Question</td>
<td>Study type: Number of studies Number of subjects</td>
<td>Risk of bias</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Comments</td>
<td>Magnitude of effect (Strength of evidence)</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Weight/BMI changes</td>
<td>Controlled trial: 1 14</td>
<td>High</td>
<td>NA</td>
<td>Direct</td>
<td>Imprecise</td>
<td></td>
<td>About 1% greater decrease in BMI in treated compared with placebo group; absolute change in BMI in treated group of 0.5 kg/m² (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Blood pressure changes</td>
<td>RCT: 1 20</td>
<td>High</td>
<td>Inconsistent</td>
<td>Indirect; treated subjects in 1 study included 2 patients with Graves disease and 8 with autonomous nodules</td>
<td>Imprecise</td>
<td></td>
<td>2.58 mmHG reduction in daytime systolic blood pressure from 1 study; no change in 2nd study (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Changes in bone density (as measured by DEXA scan)</td>
<td>RCT: 1 20</td>
<td>High</td>
<td>NA</td>
<td>Indirect; treated subjects included 2 patients with Graves disease and 8 with autonomous nodules</td>
<td>Imprecise</td>
<td></td>
<td>No effect (Insufficient)</td>
<td></td>
</tr>
</tbody>
</table>
Table B. Summary of evidence for subclinical hyperthyroidism (continued)

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study type: Number of studies</th>
<th>Risk of bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect (Strength of evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changes in lipid levels</td>
<td>RCT: 1 20</td>
<td>High</td>
<td>NA</td>
<td>Indirect; treated subjects included 2 patients with Graves disease and 8 with autonomous nodules</td>
<td>Imprecise</td>
<td></td>
<td>No effect (Insufficient)</td>
</tr>
</tbody>
</table>

KQ4. What are the harms of treatment of subclinical hyperthyroidism?

<table>
<thead>
<tr>
<th>Study type: Number of studies</th>
<th>Risk of bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect (Strength of evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>Insufficient</td>
</tr>
</tbody>
</table>

Note: DEXA=dual-energy x-ray absorptiometry; NA=not applicable; RCT=randomized controlled trial.
Discussion

The findings of this review are consistent with those of a previous Cochrane review (2007),\(^3\) and an older systematic review conducted for the USPSTF and the Institute of Medicine (2004)\(^1\) and the Surks review (2004).\(^2\) We found that the benefits and harms of screening for subclinical thyroid dysfunction remain inadequately studied. We found no RCTs assessing the benefits or harms of screening for thyroid dysfunction in the general population. The evidence was insufficient to assess or quantify the effect of screening on cardiovascular events and cardiac risk factors.

With regard to treatment for subclinical hypothyroidism, no RCTs have directly compared well-defined strategies for routine or selective treatment with a strategy of watchful waiting (active surveillance). No trials have tested the theory that early treatment of subclinical hypothyroidism can prevent coronary events or other heart disease. The applicability of these trials to decision making in primary care settings in the United States was poor.

We considered the evidence insufficient to estimate an effect size or to draw conclusions with regard to benefits of treatment for lipids. For all other outcomes, we assessed the reviewed studies as indicating that there is no benefit of treatment over watchful waiting for either subclinical hypothyroidism or subclinical hyperthyroidism, and we rated the quality of that evidence as low for subclinical hypothyroidism and insufficient for subclinical hyperthyroidism. Evidence was insufficient to assess the long-term harms of treatment for either subclinical hypothyroidism or subclinical hyperthyroidism.

Currently, it is unclear whether screening and early treatment for thyroid disease is better than not screening or watchful waiting when a TSH is mildly abnormal. For patients who have been screened and have a mildly elevated TSH level, the balance of benefits and harms of treatment vs. active surveillance is unclear. Recent studies indicate that, in elderly individuals, a mildly high TSH may be a predictor of longevity and possibly of better functional status. A well-designed RCT or cohort study comparing the outcomes of well-defined alternative strategies—routine or selective treatment vs. active surveillance—is needed to determine which strategy has better outcomes.

References


Surks MI, Ortiz E, Daniels GH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA Jan 14;291(2):228-238.

Introduction

Background

A mildly elevated thyroid stimulating hormone (TSH – also called thyrotropin) concentration is the most common thyroid function test abnormality encountered in everyday practice. Most patients who have a mildly elevated TSH have a normal free thyroxine (T4) level. The treatment of such patients is controversial, particularly when they have few or no symptoms and no other clinical evidence of thyroid disease. Less frequently, clinicians encounter patients who have a low or undetectable serum TSH and normal triiodothyronine (T3) and free T4 levels. The management of these patients is also unclear.

The main purpose of this review is to compare the effectiveness of different strategies for managing individuals who have mildly elevated serum TSH concentrations and those who have mildly diminished TSH concentrations with normal free T3 and/or free T4. We also address the evidence for whether the primary care clinician should screen for thyroid function in patients who have no specific indication for thyroid testing and who come to the clinician for other reasons. For the purposes of this review, we considered overt thyroid disease to be a well-defined clinical entity that has clear signs and symptoms, and thus, outside the scope of our review.

This topic was also the focus of a 2004 U.S. Preventive Services Task Force review.

Description of Condition

Disorders of the thyroid gland are among the most common endocrine conditions that U.S. clinicians evaluate and treat. Hyperthyroidism or hypothyroidism affects about five percent of adults in the United States. The thyroid gland is involved in metabolic homeostasis in adults. It accomplishes this through secretion of two hormones, thyroxine (T4) and triiodothyronine (T3), and is regulated by thyroid stimulating hormone (TSH), which is secreted by the anterior pituitary. Hypothyroidism is the under-secretion of thyroid hormones, while hyperthyroidism is the over-secretion of these hormones.

Symptoms of overt hypothyroidism are subtle and nonspecific and may include fatigue, feeling cold, weight gain, hair loss, poor concentration, dry skin, and constipation (Table 1). If overt hypothyroidism is allowed to progress due to lack of treatment or under-treatment, then myxedema coma, a life-threatening condition, can occur. Myxedema coma is generally seen in the elderly and may be precipitated by factors that impair respiration; it is marked by hypothermia, hypoventilation, decreased level of consciousness, and sometimes seizures and death.

Symptoms of overt hyperthyroidism may include palpitations, heat intolerance, and sweating, weight loss, hyperactivity, and fatigue. Thyroid storm is a life-threatening condition that results from an acute illness superimposed on undiagnosed or under-treated hyperthyroidism. It is accompanied by fever, delirium, seizures, and coma.
Table 1. Symptoms and signs of overt thyroid dysfunction

<table>
<thead>
<tr>
<th></th>
<th>Hypothyroidism</th>
<th>Hyperthyroidism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symptoms</strong></td>
<td>Coarse, dry skin and hair</td>
<td>Nervousness and irritability</td>
</tr>
<tr>
<td></td>
<td>Cold intolerance</td>
<td>Heat intolerance</td>
</tr>
<tr>
<td></td>
<td>Constipation</td>
<td>Increased frequency of stools</td>
</tr>
<tr>
<td></td>
<td>Deafness</td>
<td>Muscle weakness</td>
</tr>
<tr>
<td></td>
<td>Diminished sweating</td>
<td>Increased sweating</td>
</tr>
<tr>
<td></td>
<td>Physical tiredness</td>
<td>Fatigue</td>
</tr>
<tr>
<td></td>
<td>Hoarseness</td>
<td>Blurred or double vision</td>
</tr>
<tr>
<td></td>
<td>Paresthesias</td>
<td>Erratic behavior</td>
</tr>
<tr>
<td></td>
<td>Periorbital puffiness</td>
<td>Restlessness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heart palpitations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Restless sleep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Decrease in menstrual cycle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased appetite</td>
</tr>
<tr>
<td><strong>Signs</strong></td>
<td>Slow cerebration</td>
<td>Distracted attention span</td>
</tr>
<tr>
<td></td>
<td>Slow movement</td>
<td>Tremors</td>
</tr>
<tr>
<td></td>
<td>Slowing of ankle jerk</td>
<td>Tachycardia</td>
</tr>
<tr>
<td></td>
<td>Weight gain</td>
<td>Weight Loss</td>
</tr>
<tr>
<td></td>
<td>Goiter</td>
<td>Goiter</td>
</tr>
</tbody>
</table>

Subclinical thyroid dysfunction includes subclinical hypo- and hyperthyroidism. Since the development of sensitive TSH assays, these conditions have been defined as follows (Table 2):  
- High or low serum thyroid stimulating hormone (TSH) levels  
- Normal free T4 and T3 levels  
- The absence of signs and symptoms of overt thyroid dysfunction.³

For the purposes of this report, the term ‘subclinical thyroid dysfunction’ is used to define the state of having an abnormal TSH in the context of normal free T4 and T3 levels. It includes those with ‘sub-clinical thyroid disease,’ i.e. those who have a high risk of disease progression or other adverse consequences, but also those whose prognosis is not well understood.

Patients with subclinical hypothyroidism are further categorized into those with mildly elevated TSH (4.5–10 mIU/L), and those with markedly increased serum TSH levels (>10 mIU/L).⁴ Similarly, TSH levels below the lower reference limit can be classified as “low but detectable” (serum TSH 0.1-0.4mIU/L) and “clearly low serum TSH” (less than 0.1mIU/L).

Table 2. Classification of thyroid dysfunction: Biochemical definition

<table>
<thead>
<tr>
<th>Condition</th>
<th>TSH</th>
<th>Thyroid hormones</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overt hyperthyroidism</td>
<td>&lt; 0.1 mIU/L or undetectable</td>
<td>Elevated T4 or T3</td>
<td></td>
</tr>
<tr>
<td>Overt hypothyroidism</td>
<td>&gt; 4.5 mIU/L</td>
<td>Low T4</td>
<td></td>
</tr>
<tr>
<td>Subclinical hyperthyroidism</td>
<td>TSH &lt; 0.1 mIU/L</td>
<td>Normal T4 and T3</td>
<td>Clearly low serum TSH</td>
</tr>
<tr>
<td></td>
<td>0.1 mIU/L &lt;= TSH &lt; 0.4 mIU/L</td>
<td>Normal T4 and T3</td>
<td>Low but detectable</td>
</tr>
<tr>
<td>Subclinical hypothyroidism</td>
<td>4.5 mIU/L &lt; TSH &lt; 10 mIU/L</td>
<td>Normal T4</td>
<td>Mildly elevated TSH</td>
</tr>
<tr>
<td></td>
<td>TSH &gt;= 10 mIU/L</td>
<td>Normal T4</td>
<td>Markedly elevated TSH</td>
</tr>
</tbody>
</table>

Abbreviation: TSH=thyroid stimulating hormone.
Although widely used in the literature, these thresholds are arbitrary. For example, longitudinal studies have demonstrated that the risk of progression to overt hypothyroidism increases as the initial serum TSH level increases, but do not support the common notion of a threshold at either 4.5 mIU/L or at 10 mIU/L. Nevertheless, because many guidelines and studies use these thresholds, we use them in this report.

The biochemical definition has several limitations. First, the term “subclinical” usually implies that symptoms and signs are absent, whereas, in actual practice, the more common situation is that patients have nonspecific symptoms such as cold intolerance or feeling tired. As Cooper has noted, “Although subclinical hypothyroidism is the term most frequently used..., it is not necessarily apt, since on close questioning many patients disclose mild nonspecific symptoms. Mild hypothyroidism may be a more appropriate term for this very common syndrome, which is defined by an isolated elevated serum TSH level in the setting of normal serum thyroid hormone levels, in the presence or absence of symptoms.” Importantly, the presence of such symptoms does not mean they are related to the finding of an abnormal TSH test. In epidemiologic studies, individuals who had one or two nonspecific symptoms, such as cold intolerance or feeling “a little tired,” were no more likely to have subclinical thyroid dysfunction, or be at increased health risk, than were asymptomatic individuals in the general population.

Second, the biochemical, TSH-based definition of subclinical thyroid dysfunction does not take into account clinical factors that affect the natural history. Subclinical hypothyroidism can be caused by recent hospitalization for severe illness; previously treated Graves disease or nodular thyroid disease; thyroiditis and recovery from thyroiditis; or untreated adrenal insufficiency. The lack of clinical context has caused confusion about the applicability of evidence from patients with overt thyroid disease to asymptomatic, otherwise healthy individuals who have an abnormal TSH level. For example, a frequently cited randomized trial of treatment for subclinical hypothyroidism recruited patients who had undergone treatment for Graves disease. Such patients predictably and rapidly progress to symptomatic hypothyroidism if not treated. While the study demonstrated quite convincingly that early treatment can prevent symptoms in patients who have undergone thyroid ablation, patients who have no history of thyroid disease and no clinical findings or symptoms attributable to the thyroid are unlikely to progress as rapidly, so treatment has much less effect on symptoms in the short term. The vast majority of patients identified by screening in the clinic or population-based settings are in the latter group.

Third, there is no consensus on what TSH level should be used as the cut off to diagnose subclinical hypothyroidism or hyperthyroidism. Differences among assays make it difficult to establish a universal upper limit. Most studies define an abnormal TSH test result as the upper and lower limits of the assay’s 95 percent reference range, approximately 0.1 to 4.5 mIU/L. This method, although widely used in laboratory medicine, is not appropriate for identifying a threshold for diagnosing or treating subclinical hypothyroidism. As for other clinical measures, such as blood pressure, bone density, or serum lipid levels, the threshold should depend on the risk associated with a particular level as well as the balance of benefits and harms of treatment. Some have argued that the threshold for subclinical hypothyroidism should be raised above the upper limit of the reference range. The rationale is that otherwise healthy individuals who have a TSH in the range 4.5 mIU/L to 8 mIU/L or 9 mIU/L have not been shown to benefit from detection and treatment. Other experts argue that, because a TSH within the upper end of the usual reference range (2.5 to 4.5 mIU/L) confers some additional risk of
progressing to overt hypothyroidism over time, the threshold for diagnosing subclinical hypothyroidism should be lowered to 2.5 mIU/L.\textsuperscript{14-16} Approximately 9.7 percent of the U.S. adult population, representing 20.6 million Americans, has a TSH in this range and would be identified as having subclinical hypothyroidism if this change were made.\textsuperscript{4}

The appropriate level for decisionmaking can only be decided by better evidence about the adverse consequences in untreated individuals and the benefits and harms of treatment at different TSH thresholds.\textsuperscript{1,4,13,17} Because it depends on the risk of complications in a particular population, the appropriate TSH cutoff for defining subclinical hypothyroidism might vary with age, gender, and race.\textsuperscript{18}

Fourth, although the definition of subclinical thyroid dysfunction requires a “normal” free T4 level, the definition is not necessarily applicable to the individual patient. Some experts argue that if a patient’s TSH level is mildly elevated, then even if their thyroid hormone levels are within the normal range, they “are not truly normal for that individual.”\textsuperscript{19} Put differently, while higher levels of thyroid-stimulating hormone increase thyroid hormone production, the additional production does not fully compensate for the underlying deficiency. According to this view, subclinical hypothyroidism represents “early thyroid failure,” that is, less than full compensation for the diminished function of the thyroid gland.

**Prevalence and Course of Mild Thyroid Dysfunction**

Using the upper limit of the reference range as a cut off, approximately 5 percent of women and 3 percent of men have subclinical thyroid dysfunction (i.e., TSH > 4 mIU/L).\textsuperscript{1} Approximately one in four of these individuals has a markedly elevated TSH concentration (>10 mIU/L). Such patients are likely to progress to overt hypothyroidism over 20 years.\textsuperscript{20}

The other 75 percent of individuals with subclinical hypothyroidism have mildly elevated TSH levels (4 mIU/L < TSH ≤ 10 mIU/L). In this group, age, sex, geographic region, and the presence of thyroid auto-antibodies are strong predictors of the rate of progression to overt hypothyroidism. From one-third to two-thirds of these individuals have antithyroid antibodies. Depending on age, sex, and TSH level, 50 percent to 70 percent of these individuals will progress to overt disease over 20 years. In those who do not have antithyroid antibodies, the risk of progression is lower.\textsuperscript{20}

In general, prevalence increases with age, is higher among whites compared with blacks, and higher in women compared with men.\textsuperscript{21} Estimates of the prevalence of subclinical hypothyroidism vary based on demographic factors and differences in the defined upper normal limit for TSH. A systematic review and meta-analysis of good-quality cross-sectional studies\textsuperscript{1} estimated that the prevalence in women ranged from 1.2 percent among non-Hispanic, African-American women to 5.8 percent in non-Hispanic, white women and from 4 percent in women age 18-44 to over 17 percent in women over 75 years. In the NHANES sample, estimates ranged from 1.8 percent among non-Hispanic, African-American men to 2.4 percent among Mexican-American men.\textsuperscript{22} In a population-based epidemiological study in Whickham, England, the prevalence ranged from 1 percent among men 18-65 to 6.2 percent among men 65 or older.\textsuperscript{23}
Strategies for Detecting and Managing Subclinical Thyroid Dysfunction

Screening Strategies

Screening can be defined as “the application of a test to detect a potential disease or condition in a person who has no known signs or symptoms of that condition at the time the test is done.” In case-finding, testing for thyroid dysfunction is performed among patients who come to their clinicians for unrelated reasons. When the test is abnormal, the patient is called back for a detailed thyroid-directed history and confirmatory testing. Subclinical hypothyroidism is diagnosed if the TSH remains elevated and the free T4 remains normal for a period of 3-6 months. While hypothyroidism and hyperthyroidism are distinctly different disorders, with different symptoms and potential complications, screening for both subclinical hypo- and hyperthyroidism is accomplished through testing of serum TSH, with testing of serum free T4 if the TSH is high, and of T3 as well as free T4 if the TSH falls below the normal range.

Management Strategies

Subclinical Hypothyroidism

For patients with subclinical hypothyroidism, the alternative management strategies are treatment with thyroid replacement hormone versus watchful waiting. A detailed strategy for routine treatment is described elsewhere. Treatment strategies vary, but all begin with repeat testing to confirm that the TSH is still elevated. If the TSH is >10 mIU/L on repeat testing, treatment with levothyroxine is initiated. If the TSH is mildly elevated (above the reference range but below 10 mIU/L), some experts recommend routine treatment. Others recommend measurement of serum thyroid peroxidase antibodies and treatment with levothyroxine if antibody levels are high. All guidelines recommend levothyroxine rather than triiodothyronine or both as the drug of choice. The preference for levothyroxine is based on the results of randomized trials of levothyroxine alone versus combination therapy for patients with overt hypothyroidism that show no clear benefit of combination therapy over levothyroxine alone.

“Watchful waiting” encompasses two different strategies—“expectant management” or “active surveillance.” Expectant management means closely watching a patient’s condition and treating if symptoms develop or laboratory results change. Active surveillance means repeating thyroid function tests and taking a thyroid-directed history at a particular interval; then, if symptoms or laboratory results change, beginning treatment. The appropriate time-interval for retesting individuals with subclinical thyroid dysfunction is unknown.

A recent British guideline offered the following typical strategy for active surveillance:

- If screening is performed, and a high serum TSH concentration and normal free T4 is found, repeat measurement 3-6 months later after excluding nonthyroidal illness and drug interference
- If the TSH is mildly elevated (above the reference range but below 10 mIU/L), obtain serum thyroid peroxidase antibodies
- If antibody levels are high, repeat measurement of TSH annually. If they are low, repeat measurement of TSH every 3 years. Initiate treatment if the TSH level is greater than 10 mIU/L or the patient develops clinical findings of hypothyroidism.
Practice styles vary widely. A well-conducted chart review study of 500 patients with subclinical hypothyroidism seen at the Mayo Clinic in 1995-1996 found that clinicians treated 38.7 percent of patients who had a TSH between 5.1 and 10.0 mIU/L. Unfortunately, more recent data and primary care-based practice patterns are not available.

**Subclinical Hyperthyroidism**

For subclinical hyperthyroidism, some experts recommend repeating thyroid function tests after 3 months and, if the TSH remains suppressed, obtaining ultrasonography and a 24-hour Radioactive Iodine Uptake (RAIU) thyroid scan. These guidelines recommend antithyroid treatment if the patient has a persistent TSH level less than 0.1 mIU/L and is found to have Graves or nodular thyroid disease. Treatments include medications, such as propylthiouracil, or ablation with radioactive iodine or surgery. The guideline recommended against routine treatment in those whose TSH was between 0.1 and 0.45 mIU/L.

**Guidelines on Early Detection and Treatment**

A review of the history of guidelines for subclinical thyroid dysfunction provides insight into the reason for practice variation (see also Appendix E). In 1990 and again in 1998, the American College of Physicians found “it is reasonable to screen women older than 50 years of age for unsuspected but symptomatic thyroid disease.” The guideline specified that the goal of routine testing was to find overt, but overlooked, thyroid dysfunction, not subclinical hypothyroidism. Because the clinical significance of mildly elevated TSH test results was uncertain, the guideline recommended obtaining a free T4 test only when the TSH level was undetectable or 10 mIU/L or more. The ACP guideline panel used a systematic review of the literature to arrive at its recommendations. These guidelines expired in 2003.

In 1999, the American Association of Clinical Endocrinologists recommended screening asymptomatic women over the age of 60. In 2000, the American Thyroid Association recommended screening all patients over 35 years of age every 5 years (more frequently if the patient is at increased risk). These organizations used a consensus process to develop guidelines and did not use systematic reviews in arriving at their recommendations.

In 2003, the Institute of Medicine (IOM) published a volume entitled “Medicare Coverage of Routine Screening for Thyroid Disease,” which examined the issue of screening for thyroid dysfunction in the Medicare population and concluded that “there is insufficient evidence to recommend periodic, routine screening for thyroid dysfunction among asymptomatic persons using serum TSH levels.” In 2004, the USPSTF determined that the evidence was poor that treatment of screen-detected adults, in either the general population or in specific high-risk groups, improves clinically important outcomes and concluded that there was insufficient evidence to recommend for or against screening for thyroid disease in adults. These groups used essentially identical systematic reviews to arrive at their recommendations. The USPSTF issued an “I” recommendation, indicating that the evidence was insufficient to recommend for or against routine screening for thyroid disease in adults.

The conclusions about the evidence were:

- There is fair evidence that the thyroid stimulating hormone (TSH) test can detect subclinical thyroid dysfunction 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 (e.g., postpartum women, people with Down’s syndrome, and the elderly), there is 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 balance of benefits and harms of screening asymptomatic adults for thyroid disease could not be determined.

In 2004, a panel sponsored by the American Association of Clinical Endocrinologists, the American Thyroid Association, and the Endocrine Society evaluated data regarding the management of subclinical thyroid dysfunction. Unlike the older AACE and ATA guidelines, the panel used a systematic review of the evidence to arrive at its recommendations. The panel adopted the strength of evidence rating of the USPSTF in its assessment. The panel found insufficient evidence to support population-based screening and recommended against population-based screening for thyroid disease, though it did advocate aggressive case-finding in those considered high-risk, including pregnant women and women older than 60. They also recommended against routine treatment of patients with subclinical hypothyroidism with serum TSH levels of 4.5 - 10.0 mIU/L. Specifically, the panel found insufficient evidence to support routine treatment of individuals who have a mildly elevated TSH (Table 3). The findings about the evidence regarding the complications of subclinical hypothyroidism and the effects of treatment agreed with those of the IOM and USPSTF.

Table 3. Evidence for the association of subclinical hypothyroidism (TSH 4.5-10 mIU/L) and adverse health outcomes and quality of evidence for risks and benefits of treatment: Findings of the American Association of Clinical Endocrinologists, the American Thyroid Association, and the Endocrine Society Panel

<table>
<thead>
<tr>
<th>Complication</th>
<th>Strength of evidence for association with complications</th>
<th>Strength of evidence regarding benefits and harms of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progression to overt hypothyroidism</td>
<td>Good&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Not applicable&lt;sup&gt;g&lt;/sup&gt;</td>
</tr>
<tr>
<td>Adverse cardiac end points</td>
<td>Insufficient&lt;sup&gt;b&lt;/sup&gt;</td>
<td>No evidence</td>
</tr>
<tr>
<td>Elevation in serum cholesterol and LDL-C levels</td>
<td>Insufficient&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Cardiac dysfunction</td>
<td>Insufficient&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Symptoms of hypothyroidism</td>
<td>No evidence</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Psychiatric symptoms</td>
<td>No evidence</td>
<td>Insufficient</td>
</tr>
</tbody>
</table>

(Adapted from JAMA 2004<sup>21</sup>)

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

<sup>b</sup> Insufficient: 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.

<sup>c</sup> Data did not distinguish between serum TSH concentrations between 4.5-10 mIU/L and > 10 mIU/L.

<sup>d</sup> Thyroid hormone therapy normalizes serum TSH at any TSH concentration. In 2005, the three professional societies that sponsored the evidence-based panel issued a consensus statement rejecting its recommendations.<sup>34</sup> While acknowledging that the independent review panel found insufficient evidence to recommend treatment of patients with subclinical hypothyroidism in the range of
4.5-10.0 mIU/L, the counter argued that “lack of definitive evidence for a benefit does not equate to evidence for lack of benefit” and recommended that most patients with TSH levels between 4.5 and 10.0 mIU/L should be considered for treatment. Their rationale for recommending screening despite gaps in the evidence about treatment is discussed in detail in the next section of this report (“Rationale for Screening and Treatment”).

In 2006, three British professional associations (Association for Clinical Biochemistry, the British Thyroid Association, and the British Thyroid Foundation) published guidelines for using thyroid function tests. The guideline panel found that “screening for thyroid dysfunction in a healthy adult population is not warranted.” It recommended against routine treatment of patients who have a TSH concentration above the reference range but below 10 mIU/L but said that “a therapeutic trial of thyroxine [should be considered] on an individual patient basis.” The British panel noted that that there was growing evidence against the use of thyroxine in elderly patients, in whom an elevated TSH concentration “may reflect an adaptive mechanism to prevent excessive catabolism.” Like the ACP and AACE guidelines, the British panel recommended aggressive case-finding in women with non-specific symptoms. The panel used systematic reviews in their process for developing guidelines.

In 2007, The American College of Obstetricians and Gynecologists (ACOG) recommended against routine screening in pregnant women. It stated that there is no evidence that identifying and treating pregnant women with subclinical hypothyroidism improved either maternal or infant outcomes.

Rationale for Screening and Treatment

Subclinical Hypothyroidism

Although there is wide agreement that the long-term benefits of early treatment of subclinical thyroid dysfunction have not been proven, there is disagreement about what to do until better evidence is available. This disagreement reflects differing views of the clinical relevance of research about the complications of subclinical thyroid dysfunction.

Proponents argue that thyroid dysfunction is common and associated with significant morbidity. Additionally, a serum TSH test is relatively inexpensive, accurate, readily available, and generally a very acceptable test for patients to undergo. Symptoms of overt thyroid dysfunction also can be vague and at times difficult to diagnose, and therefore, thyroid screening may allow the diagnosis of overt disease earlier in the clinical course, thus reducing morbidity. For subclinical hypothyroidism, treatment with levothyroxine is noninvasive and inexpensive. Finally, proponents argue that the potential harms of screening are small in relation to the potential benefits: “Because the potential harm of early detection and treatment appear to be so minor and preventable, it seems prudent to err on the side of early detection and treatment until there is sufficient data to address these issues definitively.”

Since 2005, the most important, widely cited argument for early treatment of thyroid dysfunction is the association of untreated subclinical hypothyroidism with other risk factors for heart disease and with incident coronary disease or heart failure later in life. Subclinical hypothyroidism may be a risk factor for atherosclerosis and myocardial infarction, but epidemiologic studies of this question have had mixed results. Four recent meta-analyses have evaluated the association of subclinical thyroid disorder on cardiac mortality or all-cause mortality.
One meta-analysis included six cohort studies and found that those with subclinical hypothyroidism had a relative risk (RR) of 1.53 (1.31-1.79) of having coronary artery disease (CAD) at baseline. Subclinical hypothyroidism was associated with a RR of 1.188 (1.024-1.379) of developing CAD in followup (included studies followup ranged from 4 to 20 years), but was not found to be associated with all-cause mortality. The second review (of 14 cohort and two case-control studies) found no association with subclinical hyperthyroidism and circulatory or all-cause mortality. With regard to subclinical hypothyroidism, the association with circulatory mortality was not statistically significant, but a hazard ratio of 1.25 (1.3-1.53) was found for all-cause mortality. The third review, which included 10 population-based prospective studies and two studies of convenience samples of cardiac patients or patients with a history of stroke or hip replacement, found no overall statistically significant association between either subclinical hypothyroidism or subclinical hyperthyroidism and total mortality or coronary morbidity or mortality. A statistically significant, but small, increased risk for heart disease was found for individuals younger than 65. The last meta-analysis included 15 studies (six population-based cross-sectional studies and nine longitudinal cohort studies) with 2,531 subclinical hypothyroid participants and 26,491 euthyroid individuals. Five studies provided longitudinal data on the risk of coronary events; in these, overall there was no difference in incident ischemic heart disease (IHD) in subclinical hypothyroid participants and euthyroid participants (OR 1.27 [0.95–1.69]). In a subgroup analysis, in studies of subjects younger than 65 years, the odds ratio was 1.68 (1.27–2.23), but for older subjects, there was no relationship between subclinical hypothyroidism and incident IHD (OR 1.02 [0.85–1.22]).

The epidemiological studies included in these reviews had serious limitations. Many included individuals who had known thyroid disease, ischemic heart disease, or TSH levels within the reference range. Many included subjects who underwent treatment with levothyroxine during the followup period. In general, the studies did not adequately control for potential confounders, such as lipid levels and blood pressure.

The most recent research addresses some of the limitations of previous studies. One of these, from the Cardiovascular Health Study, found no relationship between an elevated TSH level and cardiovascular outcomes among 3,233 individuals aged 65 years or older enrolled in 1989-1990 and followed an average of 12.5 years. The other was an reanalysis of the Whickham Study, a survey of 2,779 adults sampled from a mixed urban and rural area of England conducted between July 1972 and June 1974. The goal of the study was to assess the prevalence of thyroid disease in a cross-section of the community. Personal and family history of thyroid disease and thyroid-related symptoms was collected along with history of diabetes and cardiovascular-related diseases, and fasting blood samples included assessment of levels of TSH, T4, T3, and cholesterol. Unlike the earlier report from the Whickham study, the new analysis included only subjects who had mildly elevated TSH levels, excluded subjects who had known thyroid disease or ischemic heart disease, stratified subjects according to whether they were treated with thyroxine during the followup period, and adjusted for several cardiovascular risk factors as well as socioeconomic status. After adjustment for baseline age, gender, social class, body weight, history of cerebrovascular disease, diabetes mellitus, smoking, systolic and diastolic blood pressure, serum cholesterol levels, and levothyroxine use during followup as covariates, subclinical hypothyroidism was associated with a higher risk of coronary events over 20 years (hazard ratio 1.76 [1.15–2.71]) When levothyroxine use during followup was not controlled for, the relationship was weaker (hazard ratio 1.53 [0.97–2.45]).
the study died, versus 22 of 71 who did not receive levothyroxine. Despite the low number of events, the authors reported a hazard ratio of 0.20 (0.05–0.89) for all-cause mortality after adjustment for age, gender, and total serum cholesterol.

While it is stronger than previous evidence, this analysis also had weaknesses. First, the description of the methods for ascertaining cardiac events is not clear. Specifically, it is not clear what events were considered in addition to documented myocardial infarction or death from coronary disease. In the context of risk factor epidemiology, for a broadly defined composite endpoint, the hazard ratios are low, making it more likely that confounding or methodological factors account for the observed differences. Second, the study did not control for the use of lipid-lowering medications during the followup period. It is possible that the better outcomes of subjects treated with levothyroxine could be due to other interventions introduced by their clinicians. Finally, the study was conducted between 1972 and 1974, before the era of statins and other aggressive cardiac risk management techniques. It is possible that the association between subclinical hypothyroidism and subsequent cardiovascular outcomes would become negligible in the context of current cardiac disease management.

The 2004 USPSTF review by Helfand found that evidence regarding mildly elevated TSH levels and hyperlipidemia and atherosclerosis was inconsistent. In the recent reanalysis of the Whickham study data, subjects with subclinical hypothyroidism had higher baseline systolic blood pressure (146.9 ±26.4 mm Hg vs. 139.5± 24.7 mm Hg) and total cholesterol levels (6.2 ±1.3 mmol/L [239.8 ± 50.3 mg/dL]) vs. 5.9 ±1.2 [228.2 ± 46.4 mg/dL]) and were slightly less likely to smoke than euthyroid subjects (see Appendix A for lipid conversion factors). After adjustment for age, gender, weight, smoking, and relevant medications, however, systolic blood pressure was associated with subclinical hypothyroidism, but total cholesterol was not. In other cross-sectional studies, subclinical hypothyroidism was weakly associated with total and LDL cholesterol, blood pressure, abnormalities of cardiac function, and subcutaneous fat.

Recently, two prospective studies have demonstrated a correlation between subclinical hypothyroidism and congestive heart failure (CHF). In the Health, Aging, and Body Composition Study cohort, 338 of 2,730 individuals, ages 70 to 79 were found to have subclinical hypothyroidism. After a 4 year followup, a total of 178 individuals had a CHF event. While those with a TSH of 4.5 to 6.9 did not have a statistically significant higher rate of CHF events compared to those without subclinical hypothyroidism, those with a TSH of 7 to 9.9 had a hazard ratio of 2.58 (95% CI, 1.19-5.60) and those with a TSH >10 had a hazard ratio of 3.26 (95% CI, 1.37-7.77). Additionally, of the 2,555 without a history of CHF, those with a TSH of >7 had a hazard ratio of 2.33 (95% CI, 1.1.0-4.96) of developing a incident CHF event. In the Cardiovascular Health Study cohort, 474 of 3044 subjects older than 65 years who had no history of CHF had subclinical hypothyroidism, 46 of whom had a TSH of >10. Of the 12 years of followup, those with a TSH >10 had a hazard ratio of 1.88 (95% CI,1.05-3.34) of developing a CHF event, while those with a TSH of 4.5 to 9.9 were not more likely to have a CHF event.

These two relatively large prospective cohort studies do indicate that an isolated TSH might be a risk factor for the development of CHF, particularly for individuals with a TSH of > 10. This association deserves further study. However, it is still unknown if thyroid replacement therapy would modify this potential risk.
Subclinical Hyperthyroidism

Cross-sectional studies have shown untreated subclinical hyperthyroidism to be associated with tachycardia, increased left ventricular mass leading to diastolic dysfunction, atrial arrhythmias, and a decline in bone mass density increasing the risk of fractures.\textsuperscript{1,2} Only atrial fibrillation and disease progression have been associated with subclinical hyperthyroidism in longitudinal studies.\textsuperscript{1} In a recent longitudinal study of 102 women who had a TSH between 0.1 to 0.4 mIU/L, three progressed to overt hyperthyroidism and 24 reverted to a normal TSH.\textsuperscript{52}

In summary, progression to overt disease is the best established complication of subclinical thyroid dysfunction. Epidemiologic data suggest that subclinical hypothyroidism is associated with cardiovascular disease in subjects younger than 65 years, but the magnitude of risk is low. Evidence about the relationship of a mildly elevated TSH to symptoms and to other cardiac risk factors, including the metabolic syndrome, is weak. Subclinical hyperthyroidism is less common and less studied, and thus, its natural history and effects are less clear.

Scope of Review

The main question addressed in this review was whether individuals who have mildly abnormal TSH values will either benefit from or be harmed by screening and potential subsequent treatment. We also addressed the question of whether the primary care clinician should screen for thyroid function in patients seen in general medical practice who have no specific indication for thyroid testing and who come to the clinician for other reasons. For the purposes of this review, we considered overt thyroid disease to be a well-defined clinical entity that has clear signs and symptoms, and thus, outside the scope of our review.

In order for a condition to be a good candidate for screening in the general population, several conditions need to be met. First, the condition needs to be relatively prevalent, having a significant impact on the health of the population or an easily identified special population. Second, there needs to be a test that is readily available to the general population that is of reasonable cost and accuracy and is acceptable to individuals to undergo. Finally, there needs to be an intervention that is of reasonable cost and tolerability that when administered in a timely fashion will alter the disease state to prevent morbidity and/or mortality.

The Helfand (2004) review established that subclinical thyroid disease is quite prevalent; may be responsible for morbidity; and that the serum TSH test is a readily available, reliable, and acceptable test to detect the condition with a sensitivity above 98 percent and specificity greater than 92 percent.\textsuperscript{1} However, in 2004, it remained unclear whether, if detected, treating patients with subclinical thyroid disease would reduce morbidity.

As evidence of prevalence, test yield, and test performance have already been adequately established,\textsuperscript{1} this current review focuses on whether new evidence demonstrates that treatment improves clinically important outcomes in adults with screen-detected thyroid disease.

Key Questions and Analytic Framework

We reviewed published studies to answer the following key questions:

Key Question 1

Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?
Key Question 2

What are the harms of screening? Specifically, how frequently and how severely do patients screened for subclinical thyroid dysfunction experience adverse psychological impacts or other harms of work-up from screening?

Key Question 3

Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?

Key Question 4

What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism? Specifically, what are the consequences of over-treatment, including effects on bone mineral density and incidence of atrial fibrillation, and how frequently do they occur?

We developed the final analytic framework shown in Figure 1 to guide the literature review. The analytic framework shows the populations, classification, intermediate outcomes, and health outcomes we examined in the review. We define the target population as community-living, non-pregnant adults, without a history of thyroid disease or symptoms of overt hypothyroidism or hyperthyroidism, who might be representative of those seen in primary care settings.

In the framework, Arrow 1 represents studies directly comparing the effects of screening versus not screening on important health outcomes (Key Question 1). Arrow 2, corresponding to Key Question 2, represents evidence about the potential harms of screening. Arrow 3 represents studies comparing different strategies (treatment vs. monitoring) for managing individuals who are found to have borderline high or low TSH levels (Key Question 3). Arrow 4 represents studies of the harms of treatment. Outcomes of interest for subclinical hypothyroidism were cardiovascular morbidity and mortality, measures of well-being including but not limited to cognition and memory, weight change, blood pressure changes, and changes in lipid levels. For subclinical hyperthyroidism, outcomes of interest were cardiovascular morbidity and mortality, osteoporotic fractures, measures of well-being including but not limited to cognition and memory, weight changes, blood pressure changes, changes in bone mineral density, and changes in lipid levels.
Figure 1. Analytic framework

Key Questions
1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?
2. What are the harms of screening?
3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?
4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism?

Note: A fibrillation=atrial fibrillation; CAD=coronary artery disease; CHF=congestive heart failure; DEXA=dual-energy x-ray absorptiometry; NNS=number needed to screen; NNT=number needed to treat; S=subclinical; TSH=thyroid stimulating hormone.


Methods

Topic Development

This topic was nominated by the public as part of AHRQ’s EHC Program. The USPSTF also expressed interest in updating its 2004 recommendations. We received initial key questions and an analytic framework from AHRQ. Using the methods of the EHC Program and the USPSTF and with members of the USPSTF, external technical experts, and AHRQ staff, we developed a revised analytic framework and key questions to guide our literature search and review. We sought input from a representative of the nominating organization, experts in thyroid disease, and members of the USPSTF to refine the clinical logic and PICOTS (Population, Intervention, Comparators, Outcome, Timing, and Setting) and to understand the rationale for screening. Draft key questions were posted for public comment on the EHC Program Web site. The final stage in topic development was the approval of the scope and key questions for the report by AHRQ personnel.

Search Strategy

A research librarian searched MEDLINE, the Cochrane Register Database of Systematic Reviews, and the Database of Abstracts of Reviews of Effects to identify systematic reviews regarding screening and treatment of subclinical hypothyroidism or subclinical hyperthyroidism, with no limits on dates (search strategies are described in Appendix B). As described in detail below, the systematic reviews we identified examined all relevant studies published prior to 2004. Therefore, we searched MEDLINE, AGELINE (AARP.org), EMBASE (embase.com) and the Cochrane Central Register of Controlled Trials to identify studies regarding screening and treatment of subclinical hypothyroidism or subclinical hyperthyroidism published from 2002 (thereby allowing an overlap with all previous reviews) to May 2010. In addition to searching bibliographic databases, we sought additional materials by searching for regulatory information, clinical trial registries, conference proceedings, and other sources of grey literature. Additional studies were identified from citations in relevant articles, discussions with experts, and requests to pharmaceutical companies.

We also performed a supplementary search of bibliographic databases that have a higher concentration of non-English language articles. For this search, we included CINAHL and the World Health Organization (WHO) Global Health Library, which includes African Index Medicus, Index Medicus for the Eastern Mediterranean Region, Latin American and Caribbean Health Sciences Literature, Pan American Health Organization Database, WHO Library Database, and WHO Western Pacific Region Database.

Selection of Systematic Reviews

The key questions and methods used in the 2004 systematic review by Helfand are similar to those for this report. A search for additional systematic reviews was conducted to ensure that the relevant literature included was comprehensive.

Two reviewers evaluated each review at the title/abstract and full-text article stages to determine eligibility for inclusion. Only reviews that (1) included studies of screening and treatment for subclinical thyroid dysfunction in asymptomatic individuals, (2) included studies...
that met the design and quality criteria of this current study, and (3) considered outcomes included in the analytic framework of the current study were included.

**Data Abstraction and Quality Rating of Systematic Reviews**

Data from the included systematic reviews (Table 4) were abstracted according to the guidance provided in the chapter on “Using Existing Systematic Reviews to Replace de Novo Processes in CERs” from the AHRQ EHC Program Methods Guide for Effectiveness and Comparative Effectiveness Reviews available at: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318.

**Table 4. Included systematic reviews**

<table>
<thead>
<tr>
<th>Author Year; N studies</th>
<th>Study types, N</th>
<th>Total participants</th>
<th>SR assessment of quality of primary literature</th>
<th>Overlapping studies N&lt;sup&gt;b,c&lt;/sup&gt;</th>
<th>Quality of SR</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helfand 2004; 8</td>
<td>RCT, 8; Metaanalysis of observational studies, 2</td>
<td>RCT, 291&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Screening Benefits: NA</td>
<td>Villar: 6 Surks: 7</td>
<td>AMSTAR 9/11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Screening Harms: Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yield: Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Treatment Benefits: Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Treatment Harms Osteoporosis: Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Overtreatment: Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Other: Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surks 2004; 23</td>
<td>RCT, 7; CCT, 15; Observational studies, 1; Meta-analysis, 1</td>
<td>RCT, 301; CCT, 442; Obs, 21; Meta, 13 studies, 248 patients</td>
<td>NA</td>
<td>Helfand: 7 Villar: 7</td>
<td>AMSTAR 3/11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Villar 2007; 12</td>
<td>RCT, 12</td>
<td>RCT, 461+</td>
<td>Maximum Jadad score of 5; median score of 12 studies was 4 with a range of 3 to 5</td>
<td>Helfand: 6 Surks: 7</td>
<td>AMSTAR 10/11</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** AMSTAR=Assessment of Multiple Systematic Reviews; CCT=controlled clinical trial; NA=not applicable; RCT=randomized controlled trial; SR=systematic review.

<sup>a</sup> One study did not report N.

<sup>b</sup> Two studies (Cooper 1984; Jaeschke 1996) were common to all three reviews.

<sup>c</sup> See Table 6 Study Concordance for details.
Abstracted data includes the author and year of the review; the number of included studies and their type (RCT, observational, etc.); total study participants; the reviews assessment of the quality of the body of evidence if available; the overlap of studies among the reviews; and the quality of the systematic review. The quality of each study was assessed using the AMSTAR measurement tool.55

**Individual Study Selection**

Two reviewers evaluated each study at the title/abstract and full-text article stages to determine eligibility for inclusion. Studies that included patients with Hashimoto’s disease were included if those patients were asymptomatic and likely were diagnosed as a result of screening for antibodies during the study. Because few studies included patients identified through screening in primary care settings, we included studies of patients from endocrine clinics and from laboratory database studies in which testing of thyroid function was done for a variety of reasons.

The types of studies included and excluded were different for the various key questions (Table 5). For Key Question 1 we included RCTs and observational studies of screening or case-finding for subclinical thyroid dysfunction. For Key Question 2 we did not restrict our search to any particular study design but searched for any harms, including those related to treatment as well as those related to patient issues and concerns about testing, screening, or case finding for subclinical thyroid dysfunction. For Key Question 3, we restricted our search to RCTs of treatment for subclinical hypothyroidism. However, because of the paucity of studies, in addition to RCTs, we included controlled trials, cohort studies, case-control, or other studies that looked at treatment of subclinical hyperthyroidism. For Key Question 4 we again did not restrict our search to any particular study design but searched for any articles reporting harms due to treatment for subclinical hypothyroidism or subclinical hyperthyroidism, including RCTs, controlled trials, cohort studies, and case-control studies.

A discussion of the costs or cost effectiveness of screening was not within the scope of this review. The target population was community living adults, representative of primary care settings. Hospitalized or recently hospitalized participants without history of thyroid disease or symptoms of overt hypothyroidism or hyperthyroidism were excluded as these individuals may have elevated TSH levels. Publications from 2002 to the present were included in the study. This start date was chosen to provide a small period of overlap with the publication date range used in the most recent AHRQ review (Helfand 2004). Non-English language studies were considered for inclusion in the review.

**Table 5. Criteria for inclusion/exclusion of studies in the review**

<table>
<thead>
<tr>
<th>Key Question</th>
<th>RCT</th>
<th>Controlled Trials</th>
<th>Cohort Studies</th>
<th>Case-control Studies</th>
<th>Observational Databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a and 1b</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>2</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>3</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>4</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

Abbreviation: RCT=randomized controlled trial.
Data Abstraction and Quality Rating of Individual Studies

We abstracted details about the patient population, study design, data analysis, and results. One author abstracted data and another author verified data abstraction for accuracy. We used predefined criteria developed by the USPSTF to assess the internal validity (quality) of studies (Appendix C). Two authors independently rated the internal validity of each study as “good,” “fair,” or “poor.” For all studies, we evaluated applicability to populations likely to be encountered in primary care screening settings. Discrepancies in quality ratings were resolved by discussion and consensus.

The quality of each study was assessed using criteria established by the USPSTF, and except for subclinical hyperthyroidism where no good or fair quality studies were found, poor quality studies were excluded from review.

Data Synthesis

We assessed overall strength for each body of evidence addressing a particular key question or part of a key question (for example, different treatments or screening tests) using guidance from the Strength of Evidence chapter of the AHRQ EHC Program Methods Guide for Effectiveness and Comparative Effectiveness Reviews available at: http://effectivehealthcare.ahrq.gov/ehc/products/60/294/2009_0805_grading.pdf.

To assign an overall strength of evidence (“high,” “moderate,” “low,” or “insufficient”) we considered the number, quality, and size of studies; consistency of results between studies; and directness of evidence.

External Review

We distributed a draft of the report for review by external experts not affiliated with AHRQ or the USPSTF, and have revised the report based on their comments as well as on comments received during the public review period.
Results

Systematic Reviews

Our search identified 86 unique reviews whose abstracts were reviewed for inclusion (Figure 2). After review, 18 were pulled for full-text review. Of those 18 reviews, four were not systematic reviews;\textsuperscript{56-59} six did not address the key questions of this review;\textsuperscript{12,43-45,60,61} two fell outside the timeframe for included reviews;\textsuperscript{62,63} and three were excluded for reasons discussed below. The studies included in the three remaining reviews are compared in Table 6.\textsuperscript{1,21,64}

Three of these reviews examined the effect of treating subclinical hypothyroidism on lipids.\textsuperscript{65-67} All of these demonstrated an improvement in lipids with treatment, but had serious flaws—they did not distinguish between studies of otherwise healthy patients from those who acquired hypothyroidism after treatment for hyperthyroidism, or they included poorly done observational studies.

A recent, good-quality Cochrane review provided the best information about several relevant outcomes, including cardiovascular mortality and morbidity, symptom improvement, and health-related quality of life.\textsuperscript{64} That review also examined all-cause mortality, lipid levels, systolic and diastolic heart function, and adverse effects of levothyroxine. The review concluded that while there might be some short-term improvement in lipid profiles and left ventricular function with treatment, trials of thyroid hormone therapy for subclinical hypothyroidism were not suited to assess survival or cardiovascular morbidity. These trials did not demonstrate that treatment improved health-related quality of life or symptoms.\textsuperscript{64}
Figure 2. Systematic review literature flow diagram

Abstracts of potentially relevant articles reviewed. Identified through bibliographical databases: **86**

- **Excluded abstracts:** **68**

  - Articles excluded: total: **15**
    - E1- Not systematic review: **4**
    - E2- Wrong population: **0**
    - E3- Wrong intervention: **0**
    - E4- No data by design: **0**
    - E5- No data by other parameter: **0**
    - E7- No relevance to KQ: **6**
    - E8- Related to topic but doesn’t meet inclusion criteria (wrong time frame or poor quality): **5**

- **Includes:** **3**

**Included Articles:**

- **Key Question 1. Screening**
  - 1 Helfand 2004

- **Key Question 2. Harms of screening**
  - 1 Helfand 2004

- **Key Question 3. Efficacy of treatment**
  - 3 Villar 2007
  - Helfand 2004
  - Surks 2004

- **Key Question 4. Harms of treatment**
  - 1 Helfand 2004
Table 6. Systematic review study concordance

<table>
<thead>
<tr>
<th>Study</th>
<th>Surks 2004²¹</th>
<th>Helfand 2004¹</th>
<th>Villar 2007²⁴</th>
<th>Rugge 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooper, D.S. et al., 1984¹¹</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Nystöm, E. et al., 1988²²</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ross, D.S., Daniels, G.H., Gouveia, D., 1990⁰⁹</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jaeschke, R. et al., 1996⁰⁰</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Michalopoulou, G. et al., 1998¹¹</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Meier, C. et al., 2001⁷²</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Monzani, F. et al., 2001³³</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pollock, M.A. et al., 2001⁷⁴</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Caraçcio, N., Ferrannini, E. &amp; Monzani, F., 2002⁷⁶</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kong, W.M. et al., 2002⁷⁶</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Yonem, O. et al., 2002²⁷</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Monzani, F. et al., 2004⁹⁶</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Yazici, M. et al., 2004⁹⁹</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Caraçcio, N. et al., 2005⁸⁰</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Jorde, R., Jorde, R., and Figenschau, Y., 2006⁸¹</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Buscemi, S. et al., 2007⁸³</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Razvi, S. et al., 2007²⁹⁴</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mikhail, G.S. et al., 2008⁹⁰</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Nagasaki, T. et al., 2009⁹⁰</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

The earlier reviews¹,²¹ addressed screening and treatment of both subclinical hypothyroidism and hyperthyroidism. As seen in Table 6, there was considerable overlap in the trials reviewed by these earlier reviews and the Cochrane review. A total of 14 trials were included in one or more of these reviews. The conclusions of the two earlier reviews largely agreed with those of the more recent Cochrane review. Because of the convergence of these disparate reviews on a finding of uncertainty, we considered the evidence from these three reviews to adequately reflect the state of the evidence through 2007 (the most recent search date for the Cochrane review). This report, therefore, only addressed the question of whether there is any new evidence from more recent studies.

**Individual Studies**

The current review includes eight trials, six of which were not examined in previous systematic reviews. Four were studies of treatment for subclinical hypothyroidism and two were studies of treatment for subclinical hyperthyroidism. The flow of studies from initial identification of titles and abstracts to final inclusion or exclusion using prespecified criteria (Appendix C) is diagrammed in Figure 3. Our original search found a total of 948 individual articles published from 2002 to May 2010. After selection of articles for review, we found that no foreign language articles had been included. To ensure that we had not erred in our selection of articles we rereviewed the abstracts of foreign language articles from the original search and expanded our search to include CINAHL and the WHO Global Health Library, databases with broader indexing of foreign language journals. After the abstracts were reviewed, 297 articles from our original search were found to meet criteria for full article review. The supplemental
review resulted in identifying 12 articles for full-text review. Of the 12 foreign-language articles reviewed, none met criteria for inclusion in this review. The flow of studies from this supplementary search is diagrammed in Figure 4. Excluded studies are described in Appendix D. No controlled trials of the benefits or harms of screening (Key Questions 1 and 2) were identified, but two observational studies relevant to the harms of screening were identified. Six trials assessed one or more relevant outcomes of treatment for subclinical hypothyroidism (Key Question 3), but none of them assessed harms systematically.

Several aspects of these trials make their applicability to a U.S. primary care population problematic. None of the studies was conducted in the United States. Most of the studies recruited subjects from specialty clinics, rather than from the primary care setting. Additionally, the studies were relatively short, the longest lasting 12 months. A longer duration would be needed to compare early treatment with annual retesting, the main alternative in practice.

A meta-analysis was not performed due to the methodological and clinical diversity among the included studies. The TSH value that was used to diagnose subclinical thyroid dysfunction varied among the studies. Different dosages of medications were used in the different trials.
Figure 3. Individual study literature flow diagram

Abstracts of potentially relevant articles reviewed. Identified through bibliographical databases: 948

Excluded abstracts: 651

Grey Literature: 42
Gleaned from references in retrieved articles: 108

Full text articles reviewed for relevance to key questions: 297

Included Articles:

Key Question 1. Screening
No studies

Key Question 2. Harms of screening
2 studies

Key Question 3. Efficacy of treatment
6 studies for hypothyroidism; 2 studies for hyperthyroidism

Key Question 4. Harms of treatment
No studies

Articles excluded: total: 287
E1. Background: 48
Complications of SCH – 3
Effect of elevated TSH - 6
Review – 11
E2. Wrong population: 78
E3. Wrong intervention: 0
E4. No data by design: 52
E5. No data by other parameter: 16
E7. No relevance to KQ: 98
E8. Related to topic but doesn’t meet inclusion criteria: 41

*The total exceeds the number in the corresponding box because articles could be excluded for more than one reason at this level
Figure 4. Foreign language literature flow diagram

Abstracts of potentially relevant articles reviewed. Identified through bibliographical databases: 173

Excluded abstracts: 161

Full text articles reviewed for relevance to key questions: 12

Articles excluded: total: 12
E1- Background: 0
E2- Wrong population: 2
E3- Wrong intervention: 0
E4- No data by design: 1
E5- No data by other parameter: 0
E7- No relevance to KQ: 1
E8- Related to topic but doesn’t meet inclusion criteria (poor quality): 8
E9 - Duplicate report of study: 0
E10- Translation not received: 0

Includes: 0

Included Articles:

Key Question 1.
Screening

Key Question 2.
Harms of
screening

Key Question 3.
Efficacy of
treatment

Key Question 4.
Harms of
treatment

0
0
0
0
Key Question 1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?

In 2004, Helfand1 did not find any studies that evaluated the impact of screening for thyroid disease on morbidity or mortality. We also identified no controlled trials of screening for subclinical hypothyroidism or subclinical hyperthyroidism in the general population.

Key Question 2. What are the harms of screening for subclinical thyroid dysfunction?

Information about the harms of screening is still sparse. We identified no RCTs or controlled observational studies that evaluated harms associated with screening for subclinical thyroid dysfunction. Overdiagnosis is a potential harm because TSH levels return to the reference range in a substantial proportion of individuals.

Potential harms of screening for subclinical hypothyroidism include false positive test results, anxiety related to test results, and overdiagnosis. There is little direct evidence about these potential adverse effects of screening.

In the context of screening for subclinical hypothyroidism, overdiagnosis can be defined as diagnosing subclinical hypothyroidism in a patient who cannot benefit from the diagnosis. Overdiagnosis is relevant to screening for subclinical hypothyroidism because many individuals who have a mildly elevated TSH level and a normal free T4 never develop complications and, in some, the TSH reverts spontaneously to a value below the upper reference limit.

Spontaneous reversion of mildly elevated TSH levels was observed in the Whickham study, other natural history studies, and in trials of levothyroxine therapy. Recent studies suggest that it is more frequent than previously thought and that it can occur after a long period of persistent elevation (Table 7). An observational, population-based study followed elderly patients living in Leiden, Netherlands, screening for thyroid disease at age 85 and then rechecking thyroid function at age 88.87 Twenty-one subjects were initially found to have subclinical hypothyroidism (all were 85 years old). Three years later, none of the subjects had progressed to overt hypothyroidism, eight continued to have subclinical hypothyroidism, two had developed subclinical hyperthyroidism, and 11 subjects now had normal thyroid function tests. In the same study, 12 subjects were initially found to have subclinical hyperthyroidism. Three years later, one had progressed to overt hyperthyroidism while five continued to have subclinical hyperthyroidism, five individuals had normal thyroid function, and one subject had developed subclinical hypothyroidism.

In a natural history study, 107 patients age 55 to 83 with newly diagnosed and untreated subclinical hypothyroidism referred from general practice and other specialty clinics to an endocrinology clinic were followed for a mean period of 32 months.88 All patients had two measurements of TSH above 5.0 mIU/L and free T4 in the normal range of 0.75-2.0 ng/dL one to three months prior to entering the study. During the 32 month followup period a total of 40 patients (37.4%) reverted back to normal TSH levels without treatment. While this is only a natural history study, it suggests that overdiagnosis may occur unless treatment is withheld until repeat testing confirms persistent and progressive elevations.
<table>
<thead>
<tr>
<th>Author, year</th>
<th>Duration</th>
<th>Country and population</th>
<th>Study design</th>
<th>Sample size</th>
<th>Begun/Completed</th>
<th>Results</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gussekloo, 2004</td>
<td>Born in 1912-1914, and living in Leiden, the Netherlands, between September 1997 and September 1999</td>
<td>Uncontrolled observational study</td>
<td>558</td>
<td>In followup group at age 88, 21/30 with subclinical hypothyroid were reassessed: 0 had overt hypothyroid, 8 had subclinical, 11 had normal thyroid, and 2 had overt hyperthyroid. 12/17 with subclinical hyperthyroid were reassessed: one had overt hyperthyroid, 5 had subclinical hyperthyroid and 5 had normal thyroid and 1 had subclinical hypothyroid</td>
<td>Not rated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diez, 2004</td>
<td>Patients with &quot;spontaneous&quot; subclinical hypothyroidism referred to an endocrinology clinic</td>
<td>Uncontrolled observational study</td>
<td>107</td>
<td>28 patients (26.2%) developed overt thyroid failure during followup (mean 31.7 months). 40 patients (37.4%) showed normal TSH at the end of followup period. Normal TSH levels more common in those with nonautoimmune hypothyroidism (61.5%) and with lower TSH level at start.</td>
<td>Not rated</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviation:** TSH=thyroid stimulating hormone.
Key Question 3. Does treatment of patients with subclinical hypothyroidism or subclinical hyperthyroidism detected by screening affect outcomes?

Subclinical Hypothyroidism

An earlier systematic review conducted for the USPSTF appraised eight trials of treatment for subclinical hypothyroidism. Three of the studies included individuals with known thyroid disease, making them less relevant to a population identified by screening. Only one of these studies was rated as a good-quality study. Cooper found that patients with a history of Graves disease had a modest improvement in hypothyroid symptoms following one year of therapy with levothyroxine, but no improvement in lipid levels. Another three studies included individuals without a history of thyroid disease, thus are likely more relevant to primary care. However, only one was considered of fair quality (the others were rated poor) and that study only found improvement in short-term memory, and failed to find improvement in the Sickness Impact Profile (SIP). The final two studies included individuals who were euthyroid but either had a TSH in the 4 to 4.5 mIU/L range or had at least three symptoms of hypothyroidism. One study was rated as poor (Michalopoulou, 1998). The fair-quality study, Pollock (2001), paradoxically, found a reduced SF-36 vitality score in euthyroid patients treated with levothyroxine.

A 2007 Cochrane review provided the most recent, good-quality review and meta-analysis of levothyroxine treatment for subclinical hypothyroidism. The review included 12 studies, five of which were not included in the reviews by Surks or Helfand (see Table 6). None of the trials assessed cardiovascular mortality or morbidity, and seven evaluated symptoms, mood, and quality of life, and found no statistically significant improvement. The authors also noted methodological deficiencies in previous meta-analyses of the effect of levothyroxine on cholesterol levels and other lipids. In their systematic review, six of the trials measured plasma cholesterol or LDL cholesterol. They noted baseline differences in cholesterol levels in several of the trials, and conducted two analyses finding: (1) the post-treatment difference of the means favored treatment, but (2) analysis of changes from baseline favored placebo.

Our review includes six trials that were not included in the 2004 USPSTF review, four of which were not included in the Cochrane review (See Table 6 and Tables 8-15.) These studies were not included because they were published after the literature searches were performed for the earlier reviews and/or due to differences in inclusion/exclusion criteria. The largest enrolled 120 individuals, and four enrolled fewer than 70 individuals. The followup period ranged from 5 to 12 months. Five of the trials did not report their method for allocation concealment, four did not report their method of randomization, and three reported the number of patients enrolled at baseline but not at study conclusion, and did not indicate that intention-to-treat analysis was followed. These studies were rated fair quality.
<table>
<thead>
<tr>
<th>Author, year</th>
<th>Duration</th>
<th>Country and Population</th>
<th>Study Design</th>
<th>Sample Size Begun/Completed</th>
<th>Results</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Caraccio, 2005*</td>
<td>6 months</td>
<td>Italy. Patients recruited from outpatient clinic. All suffered from Hashimoto’s thyroiditis and had positive antibodies TSH &gt; 3.6 mIU/L</td>
<td>Randomized, placebo-controlled trial. Randomization method not stated. Allocation concealment not described.</td>
<td>LT4*: 12/12 Placebo: 11/11</td>
<td>BMI ↔️ c</td>
<td>Fair</td>
</tr>
<tr>
<td>2. Iqbal, 2006*1</td>
<td>12 months</td>
<td>Norway. Subclinical hypothyroidism patients found from general health survey. 3.5 &lt; TSH &lt; 10 mIU/L</td>
<td>Randomized placebo-controlled trial. Randomization method not stated. Blinding and allocation concealment not described.</td>
<td>LT4: 32/NR Placebo: 32/NR</td>
<td>Total Cholesterol ↔️ LDL ↔️ HDL ↔️ Triglycerides ↔️ BMI ↔️</td>
<td>Fair</td>
</tr>
<tr>
<td>3. Jorde, 2006*2</td>
<td>12 months</td>
<td>Norway. Subclinical hypothyroidism patients found from general health survey. 3.5 &lt; TSH &lt; 10 mIU/L</td>
<td>Randomized placebo-controlled trial. Randomization method not stated. Allocation concealment not described.</td>
<td>LT4: 36/36 Placebo: 33/32</td>
<td>Cognitive ↔️ General Health ↔️ Depression ↔️</td>
<td>Fair</td>
</tr>
<tr>
<td>4. Mikhail, 2008*3</td>
<td>52 weeks</td>
<td>Kuwait. Patients, age 15-60, from endocrinology outpatient clinic. 4 &lt; TSH &lt; 10 mIU/L</td>
<td>Double-blind, Placebo-controlled trial. Blinding stated but not described; allocation concealment not stated.</td>
<td>LT4: 60/NR Placebo: 60/NR</td>
<td>Total Cholesterol ↑ p &lt; 0.03 LDL ↑ p &lt; 0.001 HDL ↔️ Triglycerides ↔️</td>
<td>Fair</td>
</tr>
<tr>
<td>5. Nagasaki, 2009*4</td>
<td>5 months</td>
<td>Japan. Newly diagnosed patients with subclinical hypothyroidism due to chronic thyroiditis with positive antibodies TSH &quot;above the normal upper limit&quot;</td>
<td>Randomized, placebo-controlled study. Randomization method not described. Blinding of patients but not caretakers or assessors described. Allocation concealment not described.</td>
<td>LT4: 48/NR Placebo: 47/NR</td>
<td>Total Cholesterol ↔️ LDL ↔️ HDL ↔️ Triglycerides ↔️ Blood Pressure ↔️ BMI ↔️</td>
<td>Fair</td>
</tr>
</tbody>
</table>
Table 8. Treatment evidence table (subclinical hypothyroidism) (continued)

<table>
<thead>
<tr>
<th>Author, year Duration</th>
<th>Country and Population</th>
<th>Study Design</th>
<th>Sample Size Begun/Completed</th>
<th>Results ( \oplus ): no improvement ( \uparrow ): improvement (^a)</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Razvi, 2007(^d) 24 weeks</td>
<td>United Kingdom. Patients from urban, general practice settings identified through laboratory database, not through screening. TSH &gt; 4 mIU/L</td>
<td>Randomized, placebo-controlled cross-over (12-week) trial</td>
<td>LT4:50/50 Placebo: 50/49</td>
<td>Total Cholesterol ( \uparrow ) ( p &lt; 0.001 ) LDL ( \uparrow ) ( p &lt; 0.001 ) HDL ( \perp ) Triglycerides ( \perp ) Blood pressure ( \perp ) Weight ( \perp ) QoL ( \perp )</td>
<td>Good</td>
</tr>
</tbody>
</table>

Abbreviations: BMI=body mass index; HDL=high-density lipoprotein; LDL=low-density lipoprotein; NR=not reported; QoL=quality of life; TSH=thyroid stimulating hormone.

\(^a\) LT4: Levothyroxine.

\(^b\) For details see separate results tables by outcome – Tables 9-15

\(^c\) \( \oplus \): No evidence of benefit; \( \uparrow \): Evidence of benefit.
<table>
<thead>
<tr>
<th></th>
<th>Thyroxine</th>
<th>Placebo</th>
<th>Delta Mean Difference*</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jorde, 2006** (Norway)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage</td>
<td>109.7 mcg/day (average)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive Before</td>
<td>1.8 +/- 3.4</td>
<td>-1.1 +/- 4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive After</td>
<td>1.5 +/- 3.7</td>
<td>-0.9 +/- 4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>GHQ-30 Before</td>
<td>1.5 +/- 2.3</td>
<td>0.7 +/- 1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHQ-30 After</td>
<td>1.9 +/- 3.3</td>
<td>1.2 +/- 2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>BDI Before</td>
<td>4.4 +/- 3.7</td>
<td>3.7 +/- 3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI After</td>
<td>4.3 +/- 3.6</td>
<td>3.3 +/- 4.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Razvi, 2007** (UK)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dosage</td>
<td>100 mcg/day for 12 weeks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-QoL After</td>
<td>-1.1 +/- 1</td>
<td>-1.2 +/- 0.9</td>
<td>0.2 (0.02 to 0.36)</td>
<td>NS</td>
</tr>
<tr>
<td>Sex life After</td>
<td>-2.3 +/- 2.7</td>
<td>-2.7 +/- 2.8</td>
<td>0.3 (0.02 to 0.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Motivation After</td>
<td>-3.6 +/- 2.7</td>
<td>-3.7 +/- 2.7</td>
<td>0.4 (-0.4 to 0.9)</td>
<td>NS</td>
</tr>
<tr>
<td>Worries After</td>
<td>-2.5 +/- 3</td>
<td>-2.8 +/- 2.9</td>
<td>0.2 (-0.2 to 0.7)</td>
<td>NS</td>
</tr>
<tr>
<td>AWI After</td>
<td>-2.7 +/- 2.4</td>
<td>-2.8 +/- 2.3</td>
<td>0.1 (-0.3 to 0.5)</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Abbreviations:** AWI=Average weighted impact of all 18 domains; NS=not significant; QoL=quality of life.

* Blank cells denote no data reported.
### Table 10. Weight (kg)/body mass index (BMI) (kg/m²) (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caraccio, 2005&lt;sup&gt;b&lt;/sup&gt; (Italy)</td>
<td>Thyroxine 65 mcg (average) for 6 months</td>
<td>22.8 +/- 0.9</td>
<td>22.3 +/- 0.9</td>
<td>-0.5</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>22.6 +/- 0.6</td>
<td>22.7 +/- 0.6</td>
<td>0.1</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NS</td>
</tr>
<tr>
<td>Iqbal, 2006&lt;sup&gt;c&lt;/sup&gt; (Norway)</td>
<td>Thyroxine 50 mcg/day 1&lt;sup&gt;st&lt;/sup&gt; 6 weeks 100 mcg/day final 46 weeks</td>
<td>28.7 +/- 5.7</td>
<td>28.4 +/- 5.8</td>
<td>-0.3</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>27.1 +/- 3.7</td>
<td>27.0 +/- 4.1</td>
<td>-0.1</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NS</td>
</tr>
<tr>
<td>Nagasaki, 2009&lt;sup&gt;d&lt;/sup&gt; (Japan)</td>
<td>Thyroxine 25.8 mcg/day (average) for 5 months</td>
<td>22.0 +/- 0.48</td>
<td>21.8 +/- 0.48</td>
<td>-0.2</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>22.2 +/- 0.51</td>
<td>22.1 +/- 0.50</td>
<td>-0.1</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NS</td>
</tr>
<tr>
<td>Razvi, 2007&lt;sup&gt;e&lt;/sup&gt; (UK)</td>
<td>Thyroxine 100 mcg/day for 12 weeks</td>
<td>75.9 +/- 15.9</td>
<td>75.8 +/- 16.5</td>
<td>-0.1</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>77.0 +/- 16.9</td>
<td>76.5 +/- 16.7</td>
<td>-0.5</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.6 (-1.1 to -0.1)</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI=body mass index; NS=not significant.

<sup>a</sup> Difference calculated from numbers provided.
Table 11. Blood pressure (mmHg) (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nagasaki, 2009&lt;sup&gt;86&lt;/sup&gt; (Japan)</td>
<td><strong>Thyroxine</strong> 25.8 mcg/day (average) for 5 months</td>
<td>132.8 +/- 3.9</td>
<td>128.8 +/- 3.8</td>
<td>-4.0</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>133.1 +/- 3.4</td>
<td>132.2 +/- 3.5</td>
<td>-0.9</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Diastolic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Thyroxine</strong></td>
<td>74.3 +/- 2.9</td>
<td>72.7 +/- 2.2</td>
<td>-1.6</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>75.7 +/- 1.9</td>
<td>72.8 +/- 2.0</td>
<td>-2.9</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Diastolic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quality: Fair

Razvi, 2007<sup>87</sup> (UK)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Thyroxine</strong> 100 mcg/day for 12 weeks</td>
<td>135.7 +/- 22.6</td>
<td>132.8 +/- 22.8</td>
<td>-2.9</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>129.4 +/- 20.1</td>
<td>134.6 +/- 22.9</td>
<td>5.2</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Diastolic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Thyroxine</strong></td>
<td>80.8 +/- 8.3</td>
<td>78.8 +/- 10.3</td>
<td>-2.0</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>79.1 +/- 10.0</td>
<td>79.9 +/- 9.6</td>
<td>0.8</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quality: Good

Abbreviation: NS = not significant.

<sup>a</sup> Difference calculated from numbers provided.
Table 12. Total cholesterol (mg/dL)\(^a\) (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference(^b)</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iqbal, 2006(^a) (Norway)</td>
<td>50 mcg/day 1(^st) 6 weeks 100 mcg/day final 46 weeks</td>
<td>227.7 +/- 42.5</td>
<td>220.0 +/- 42.5</td>
<td>-7.7</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine</td>
<td></td>
<td>Placebo</td>
<td>223.9 +/- 30.9</td>
<td>223.9 +/- 34.7</td>
<td>0</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-7.7(^o)</td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mikhail, 2008(^a) (Kuwait)</td>
<td>72 +/-3.8 mcg/day (average) for 52 weeks</td>
<td>194.9 +/- 37.8</td>
<td>183.0 +/- 33.6</td>
<td>-11.9</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine</td>
<td></td>
<td>Placebo</td>
<td>192.6 +/- 27.8</td>
<td>194.5 +/- 25.9</td>
<td>1.9</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-13.8(^o)</td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagasaki, 2009(^a) (Japan)</td>
<td>25.8 mcg/day (average) for 5 months</td>
<td>215.8 +/- 10.4</td>
<td>200.3 +/- 6.2</td>
<td>-15.5</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine</td>
<td></td>
<td>Placebo</td>
<td>213.5 +/- 9.7</td>
<td>205.7 +/- 9.3</td>
<td>-7.8</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-7.7(^o)</td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Razvi, 2007(^a) (UK)</td>
<td>100 mcg/day for 12 weeks</td>
<td>235.5 +/- 34.7</td>
<td>220.0 +/- 38.6</td>
<td>-15.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine</td>
<td></td>
<td>Placebo</td>
<td>231.6 +/- 54.0</td>
<td>231.6 +/- 38.6</td>
<td>0</td>
<td></td>
<td>-13.5 (-19.3 to -7.72)</td>
</tr>
<tr>
<td>Quality: Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NS=not significant.
\(^a\) Converted from mmol/liter by multiplying by 38.6.
\(^b\) Difference calculated from numbers provided.
Table 13. Low-density lipoprotein (LDL) cholesterol (mg/dL)\(^a\) (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference(^b)</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iqbal, 2006(^a) (Norway)</td>
<td>Thyroxine 50 mcg/day 1(^st) 6 weeks 100 mcg/day final 46 weeks</td>
<td>142.8 +/- 34.7</td>
<td>139.0 +/- 34.7</td>
<td>-3.8</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>139.0 +/- 30.9</td>
<td>139.0 +/- 38.6</td>
<td>0</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mikhail, 2008(^a) (Kuwait)</td>
<td>Thyroxine 72+/-3.8 mcg/day (average) for 52 weeks</td>
<td>127.4 +/- 34.7</td>
<td>111.6 +/- 22.8</td>
<td>-15.8</td>
<td>&lt; 0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>107.7 +/- 23.2</td>
<td>120.0 +/- 29.7</td>
<td>+12.3</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagasaki, 2009(^a) (Japan)</td>
<td>Thyroxine 25.8 mcg/day (average) for 5 months</td>
<td>138.2 +/- 8.5</td>
<td>121.2 +/- 11.2</td>
<td>-17.0</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>137.4 +/- 7.7</td>
<td>129.7 +/- 7.3</td>
<td>-7.7</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Razvi, 2007(^a) (UK)</td>
<td>Thyroxine 100 mcg/day for 12 weeks</td>
<td>139.0 +/- 30.9</td>
<td>131.2 +/- 30.9</td>
<td>-7.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>139.0 +/- 46.3</td>
<td>142.8 +/- 34.7</td>
<td>+3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NS=not significant.
\(^a\) Converted from mmol/liter by multiplying by 38.6.
\(^b\) Difference calculated from numbers provided.
Table 14. High-density lipoprotein (HDL) cholesterol (mg/dL)<sup>a</sup> (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;b&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iqbal, 2006&lt;sup&gt;ab&lt;/sup&gt; (Norway)</td>
<td>Thyroxine</td>
<td>50 mcg/day 1&lt;sup&gt;st&lt;/sup&gt; 6 weeks 100 mcg/day final 46 weeks</td>
<td>57.9 +/- 15.4</td>
<td>57.9 +/- 15.4</td>
<td>0.0</td>
<td>NS</td>
<td>0&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Placebo</td>
<td>57.9 +/- 23.2</td>
<td>57.9 +/- 19.3</td>
<td>0.0</td>
<td>NS</td>
<td>0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td>Mikhail, 2008&lt;sup&gt;ac&lt;/sup&gt; (Kuwait)</td>
<td>Thyroxine</td>
<td>72 +/-3.8 mcg/day (average) for 52 weeks</td>
<td>46.3 +/- 12.7</td>
<td>45.9 +/- 12.4</td>
<td>-0.4</td>
<td>NS</td>
</tr>
<tr>
<td>Placebo</td>
<td>44.4 +/- 8.9</td>
<td>42.5 +/- 9.7</td>
<td>-1.9</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td>Nagasaki, 2009&lt;sup&gt;ad&lt;/sup&gt; (Japan)</td>
<td>Thyroxine</td>
<td>25.8 mcg/day (average) for 5 months</td>
<td>54.4 +/- 1.9</td>
<td>54.4 +/- 3.1</td>
<td>0.0</td>
<td>NS</td>
</tr>
<tr>
<td>Placebo</td>
<td>53.3 +/- 2.3</td>
<td>53.7 +/- 2.3</td>
<td>0.4</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Fair</td>
<td>Razvi, 2007&lt;sup&gt;ae&lt;/sup&gt; (UK)</td>
<td>Thyroxine</td>
<td>100 mcg/day for 12 weeks</td>
<td>65.6 +/- 19.3</td>
<td>61.8 +/- 19.3</td>
<td>-3.8</td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>61.8 +/- 15.4</td>
<td>65.6 +/- 19.3</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: NS=not significant.
<sup>a</sup> Converted from mmol/liter by multiplying by 38.6.
<sup>b</sup> Difference calculated from numbers provided.
### Table 15. Triglycerides (mg/dL)\(^a\) (subclinical hypothyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference(^b)</th>
<th>p value</th>
<th>Treatment/No Treatment Change Difference</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iqbal, 2006(^m) (Norway)</td>
<td>Thyroxine 50 mcg/day 1(^st) 6 weeks 100 mcg/day final 46 weeks</td>
<td>132.9 +/- 79.7</td>
<td>132.9 +/- 88.6</td>
<td>0</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>141.8 +/- 79.7</td>
<td>141.8 +/- 62.0</td>
<td>0</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mikhail, 2008(^m) (Kuwait)</td>
<td>Thyroxine 72 +/- 3.8 mcg/day (average) for 52 weeks</td>
<td>104.5 +/- 62.9</td>
<td>84.2 +/- 47.0</td>
<td>-20.3</td>
<td>&lt; 0.002</td>
<td>-24.7(^m)</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>89.5 +/- 58.5</td>
<td>93.9 +/- 52.3</td>
<td>4.4</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagasaki, 2009(^m) (Japan)</td>
<td>Thyroxine 25.8 mcg/day (average) for 5 months</td>
<td>118.7 +/- 12.4</td>
<td>132.9 +/- 14.2</td>
<td>14.2</td>
<td>NS</td>
<td>13.3(^m)</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>121.4 +/- 11.5</td>
<td>122.3 +/- 12.4</td>
<td>0.9</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quality: Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Razvi, 2007(^m) (UK)</td>
<td>Thyroxine 100 mcg/day for 12 weeks</td>
<td>106.3 (62.0 to 327.8)</td>
<td>115.2 (44.3 to 363.3)</td>
<td>8.9</td>
<td>NS</td>
<td>-5.3 (-17.7 to -8.9)</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>106.3 (44.3 to 274.7)</td>
<td>115.2 (35.4 to 451.9)</td>
<td>8.9</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quality: Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviation:** NS=not significant.

\(^a\) Converted from mmol/liter by multiplying by 88.6.

\(^b\) Difference calculated from numbers provided.
Cardiac Events

None of the studies evaluated the effect on cardiovascular morbidity or mortality from treatment of subclinical hypothyroidism with levothyroxine.

Quality of Life

Two studies, with a total of 169 subjects, evaluated the effect of treatment on measures of quality of life. One good-quality, 12-week crossover study (Razvi) enrolled 100 participants from 322 patients from urban, general practices in the United Kingdom identified as eligible from thyroid function tests from a laboratory database. Of the 322 patients originally identified as eligible, 63 percent had been tested either for symptoms attributable to hypothyroidism (n=179) or because of familial history of thyroid disease (n=24). Fifty of the enrollees were given 100 mcg of levothyroxine without dose titration and 50 were given placebo for 12 weeks. Patients then switched treatment arms without a wash-out period. (The half-life of thyroxine is estimated to be 5 to 9 days.) While the absence of a washout period could introduce bias, it would bias the study towards the null, that is, decrease the likelihood of demonstrating a treatment effect. The primary endpoints of this study were improved brachial artery flow-mediated dilatation (as a marker of vascular endothelial function) and total cholesterol. Secondary endpoints assessed were changes in weight (as measured by BMI) and patient-reported outcomes including perceived health status, hypothyroidism-specific quality of life, and hypothyroid symptoms (as assessed by questionnaires). This trial found no improvement in overall quality of life measures; health status, as measured by the SF-36v2; or treatment satisfaction.

The other trial (Jorde, N=69) recruited subjects who had a TSH level between 3.5 and 10 mU/L from a population-based sample in Norway. Asymptomatic subjects who had no history of thyroid disease were invited to participate in a 12-month trial of levothyroxine treatment. After 12 months, there was no significant improvement in cognitive or emotional function or in hypothyroid symptoms.

Weight/Body Mass Index (BMI)

Four studies with a total of 305 subjects looked at either weight or BMI. Study periods ranged from 5 to 12 months. Three were rated fair quality; one was rated good quality. None found any significant change in either weight or BMI after 5 months, 24 weeks, 6 months, or 1 year.

Blood Pressure

Two studies, with a total of 195 subjects looked at blood pressure. One five-month study (n=95) was rated fair quality and the other (n=100), a 12-week cross-over study, was rated good. Neither found significant change in blood pressure.

Lipids

Four studies with a total of 379 subjects evaluated the impact of treatment on lipids. Two found improvements to lipids; two found no improvement.

The good-quality randomized, controlled 12-week, cross-over trial showed modest improvement in total cholesterol and LDL. The study also reported results for changes in LDL, HDL, and triglycerides. Razvi found a significant decrease in total cholesterol of 5.8 percent
(from 235.5 mg/dL to 220.0 mg/dL; p = <0.001) and a significant 5.6 percent decline in LDL (from 139.0 mg/dL to 131.2 mg/dL; p=<0.05).

We also found one fair-quality RCT that demonstrated improvement in total cholesterol and LDL from treatment for subclinical hypothyroidism. While this study described itself as a double-blind randomized, placebo-controlled trial, blinding was not described nor was allocation concealment. In addition, the number of enrolled patients was described at baseline, but not at the end of the study, and there was no mention of intention to treat analysis. This study of 120 patients from an endocrinology outpatient clinic in Kuwait found a significant decline in LDL of 12.4 percent (from 127.38 mg/dL to 111.55 mg/dL; p<0.01) and a significant 6.1 percent decline in total cholesterol (194.9 mg/dL to 183.0 mg/dL; p< 0.0001).

Iqbal 2006, a fair-quality study with 64 subjects did a post hoc analysis and found a significant decline in total cholesterol and LDL in the subgroup of 23 patients with serum TSH at the end of the 12-month study within the range 0.2 – 2.0 mIU/L but no difference in lipids when considering the full study sample. This study reported the number of patients enrolled at baseline but not at study conclusion and did not indicate that intention-to-treat analysis was followed. Nagasaki 2008, a fair-quality study of 95 subjects found no improvement in any lipids. Like Iqbal, Nagasaki reported the number of patients enrolled at baseline but not at study conclusion and did not indicate that intention-to-treat analysis was followed.

**Subclinical Hyperthyroidism**

No controlled trials for the treatment of subclinical hyperthyroidism were found in the Helfand 2004 systematic review for the USPSTF. We identified two poor-quality controlled trials that assessed the effect of treatment of subclinical hyperthyroidism. (See Tables 16 through 20) Buscemi (2007) was designed to assess the effects of treatment of subclinical hyperthyroidism on the heart as measured by blood pressure, basal heart rate, 24-hour heart rate, and atrial and ventricular premature beats; on bone turnover; and on bone density as measured by heel ultrasonometry (Stiffness Index). That study found a small, but significant weight gain (BMI 27.3 +/- 1.3 kg/m² to 27.8 +/- 1.4 kg/m² in treated patients vs. 27.9 +/- 1.2 kg/m² to 28.1 +/- 1.0 kg/m² in untreated controls; p < 0.05) after 12 months of treatment with methimazole. However, this study was weakened by failure to conceal allocation (alternating assignment) and lack of blinding (patients were given the option of changing their randomly assigned group; blinding of assessors was not described). In Yonem (2002) (N=20), 10 patients with subclinical hyperthyroidism randomized to treatment with propylthiouracil (nine patients) or radioiodine (one patient) were compared with 10 patients randomized to no treatment and found no change in bone mineral density or lipids, a small, but significant decrease in mean daytime systolic blood pressure (from 115.00 +/- 2.78 mmHg to 112.42 +/- 2.66 mmHg in treated patients vs. from 114.50 +/- 3.21 to 113.70 +/- 2.62 mmHg in untreated controls; p < 0.05), and inconsistent findings with regard to patient-reported outcomes. Of the 10 patients randomized to treatment, nine received propylthiouracil and one received radioactive iodine, and the analysis did not distinguish between treatments. In addition, patients randomized to control received no treatment and so the study was not adequately blinded.
### Table 16. Treatment evidence table (subclinical hyperthyroidism)

<table>
<thead>
<tr>
<th>Author, year Duration</th>
<th>Country and population</th>
<th>Study design</th>
<th>Sample size Begun/Completed</th>
<th>Results</th>
<th>Quality Score</th>
</tr>
</thead>
</table>
| 1. Buscemi, 2007\(^3\) 12 months | Italy. Newly diagnosed patients. TSH < 0.49 mIU/L | Controlled trial. Treatment allocation through alternate assignment. Patients given option of changing assigned group (none did). Not blinded. | Methimazole: 7/7 No treatment: 7/7 | BMI: \(\uparrow\)
Blood Pressure: \(\uparrow\) | Poor |
| 2. Yonem, 2002\(^7\) 6 months | Turkey. Patients with subclinical hyperthyroidism for 6 – 60 months. 2 patients with Graves disease and 8 with autonomous nodule. TSH < 0.4 mIU/L | RCT. Randomization method not stated. Not blinded. Allocation concealment not described. Multiple treatments not analyzed separately | Propylthiouracil: 9/NR Radioactive Iodine: 1/NR No Treatment: 10/NR Treatments not analyzed separately | Total Cholesterol: \(\uparrow\)
LDL: \(\uparrow\)
HDL: \(\uparrow\)
Triglycerides: \(\uparrow\)
Bone Mineral Density: \(\uparrow\)
Blood Pressure: \(\uparrow\) | Poor |

**Abbreviations:** BMI=body mass index; HDL=high-density lipoprotein; LDL=low-density lipoprotein; NR=not reported; RCT=randomized controlled trial; TSH=thyroid stimulating hormone.

\(^{a}\) For details see separate results tables by outcome – Tables 17-20

\(^{b}\) \(\uparrow\): No evidence of benefit; \(\uparrow\): Evidence of benefit.
Table 17. Body mass index (BMI) (kg/m²) (subclinical hyperthyroidism)

<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference(^a)</th>
<th>p value</th>
<th>Treatment/No Treatment Before-After Difference(^a)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buscemi, 2007(^b) (Italy)</td>
<td>Methimazole TSH &gt; 0.01: 10 mg/day TSH &lt;= 0.01: 15 mg/day for 12 months</td>
<td>27.3 +/- 1.3</td>
<td>27.8 +/- 1.4</td>
<td>0.5</td>
<td>&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No Treatment</td>
<td>27.9 +/- 1.2</td>
<td>28.1 +/- 1.0</td>
<td>0.2</td>
<td>NS</td>
<td></td>
<td>0.3</td>
</tr>
<tr>
<td>Quality: Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: NR=not reported; TSH=thyroid stimulating hormone.

\(^a\) Difference calculated from numbers provided.
<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buscemi, 2007&lt;sup&gt;12&lt;/sup&gt; (Italy)</td>
<td><strong>Systolic</strong></td>
<td>139 +/- 6</td>
<td>136 +/- 4</td>
<td>-3</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methimazole</td>
<td>TSH &gt; 0.01: 10 mg/day TSH &lt;= 0.01: 15 mg/day for 12 months</td>
<td>139 +/- 6</td>
<td>136 +/- 4</td>
<td>-3</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>136 +/- 8</td>
<td>126 +/- 11</td>
<td>-10</td>
<td>NS</td>
<td>7</td>
<td>NS</td>
</tr>
<tr>
<td>Methimazole</td>
<td></td>
<td>79 +/- 3</td>
<td>78 +/- 3</td>
<td>-1</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>78 +/- 3</td>
<td>80 +/- 3</td>
<td>+2</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yonem, 2002&lt;sup&gt;77&lt;/sup&gt; (Turkey)</td>
<td><strong>Daytime Systolic</strong></td>
<td>115.00 +/- 2.78</td>
<td>112.42 +/- 2.66</td>
<td>-2.58</td>
<td>&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propylthiouracil or Radioactive Iodine</td>
<td>Propylthiouracil 150 mg/day Radioactive Iodine not stated for 6 months</td>
<td>115.00 +/- 2.78</td>
<td>112.42 +/- 2.66</td>
<td>-2.58</td>
<td>&lt;0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>114.50 +/- 3.21</td>
<td>113.70 +/- 2.62</td>
<td>-0.80</td>
<td>NS</td>
<td>-1.78</td>
<td>NR</td>
</tr>
<tr>
<td>Propylthiouracil or Radioactive Iodine</td>
<td></td>
<td>71.10 +/- 2.37</td>
<td>69.40 +/- 1.78</td>
<td>-1.70</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>72.70 +/- 2.00</td>
<td>72.10 +/- 2.37</td>
<td>-0.60</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality: Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propylthiouracil or Radioactive Iodine</td>
<td>Nighttime Systolic</td>
<td>98.50 +/- 2.96</td>
<td>100.10 +/- 2.25</td>
<td>1.60</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>100.60 +/- 2.29</td>
<td>101.60 +/- 1.96</td>
<td>1.00</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propylthiouracil or Radioactive Iodine</td>
<td>Nighttime Diastolic</td>
<td>62.80 +/- 1.95</td>
<td>61.50 +/- 1.40</td>
<td>-1.30</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Treatment</td>
<td></td>
<td>62.30 +/- 1.71</td>
<td>61.80 +/- 2.80</td>
<td>-0.50</td>
<td>NS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: NS=not significant; NR=not reported. <sup>a</sup>Difference calculated from numbers provided.
<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yonem, 2002&lt;sup&gt;27&lt;/sup&gt; (Turkey)</td>
<td>Propylthiouracil or Radioactive Iodine</td>
<td>Femur Neck</td>
<td>Femur Neck</td>
<td>0.828 +/- 0.038</td>
<td>0.826 +/- 0.042</td>
<td>-0.002</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Propylthiouracil 150 mg/day Radioactive Iodine not stated for 6 months</td>
<td></td>
<td></td>
<td>0.848 +/- 0.017</td>
<td>0.868 +/- 0.019</td>
<td>0.020</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>No Treatment</td>
<td>Lumbar Vertebra</td>
<td>Lumbar Vertebra</td>
<td>0.991 +/- 0.046</td>
<td>0.998 +/- 0.048</td>
<td>0.007</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>No Treatment</td>
<td>Lumbar Vertebra</td>
<td>Lumbar Vertebra</td>
<td>0.968 +/- 0.030</td>
<td>0.968 +/- 0.030</td>
<td>0.000</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Quality: Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.007</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Abbreviation:** NS=not significant.

<sup>a</sup> Difference calculated from numbers provided.
<table>
<thead>
<tr>
<th>Study, Year (Country)</th>
<th>Dosage/Duration</th>
<th>Before</th>
<th>After</th>
<th>Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
<th>Treatment/No Treatment Before-After Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yonem, 2002&lt;sup&gt;2&lt;/sup&gt; (Turkey)</td>
<td>Propylthiouracil or Radioactive Iodine</td>
<td>Total Cholesterol</td>
<td>Total Cholesterol</td>
<td>183.0 +/- 12.0</td>
<td>182.6 +/- 8.9</td>
<td>-0.4</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Propylthiouracil 150 mg/day Radioactive Iodine not stated for 6 months</td>
<td>No Treatment</td>
<td>Total Cholesterol</td>
<td>161.7 +/- 10.0</td>
<td>157.1 +/- 6.6</td>
<td>-4.6</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LDL Cholesterol</td>
<td>LDL Cholesterol</td>
<td>110.0 +/- 8.8</td>
<td>105.8 +/- 6.6</td>
<td>-4.2</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No Treatment</td>
<td>LDL Cholesterol</td>
<td>86.9 +/- 6.9</td>
<td>91.1 +/- 5.8</td>
<td>4.2</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HDL Cholesterol</td>
<td>HDL Cholesterol</td>
<td>54.4 +/- 4.6</td>
<td>47.5 +/- 3.9</td>
<td>-6.9</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No Treatment</td>
<td>HDL Cholesterol</td>
<td>52.5 +/- 5.8</td>
<td>48.3 +/- 3.5</td>
<td>-4.2</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Triglycerides</td>
<td>Triglycerides</td>
<td>124.0 +/- 15.9</td>
<td>39.9 +/- 23.9</td>
<td>-84.1</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No Treatment</td>
<td>Triglycerides</td>
<td>93.9 +/- 10.6</td>
<td>76.2 +/- 14.2</td>
<td>-17.7</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Abbreviations:** HDL=high-density lipoprotein; LDL=low-density lipoprotein; NR=not reported; NS=not significant.

**Note:** Converted total cholesterol, LDL, and HDL from mmol/liter by multiplying by 38.6 and triglycerides by 88.6.

<sup>a</sup> Difference calculated from numbers provided.
Key Question 4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism?

Until August, 2000, levothyroxine sodium was an unapproved marketed drug. In August 1997, the U.S. Food and Drug Administration declared levothyroxine sodium tablets a “new drug,” requiring manufacturers to submit a new drug approval to continue manufacturing it. The first product was approved in August, 2000. The product label mentions adverse effects on bone mineral density and the cardiovascular system, such as provocation of angina and arrhythmias and increased cardiac wall thickness. The U.S. Food and Drug Administration medical reviewer cited the following evidence regarding the safety of levothyroxine:

- Elderly patients ≥ 60 years, with TSH suppressed to ≤0.1 mIU/L due to either subclinical hyperthyroidism or overtreatment with levothyroxine had approximately a 3-fold increased incidence of atrial fibrillation over a 10-year period compared with those with normal TSH levels.
- Two systematic reviews found that long-term suppression therapy with levothyroxine led to decreases in bone mass in post-menopausal women. No adverse effect was found in men.

The approval was conducted without requiring manufacturers to conduct studies to estimate the actual risks of short-term or long-term adverse effects. Consequently, scant information regarding the adverse effects of thyroid replacement is available. The 2004 Helfand review and the Cochrane review both mentioned that information is lacking about the frequency and severity of side effects when levothyroxine is used to treat subclinical hypothyroidism.

Three of the studies reviewed by Helfand (2004) provided some information on harms. In one study, out of 33 individuals, four treated with levothyroxine (and six with placebo) “felt worse.” Another study reported that out of 37 total individuals, one developed atrial fibrillation, and one developed angina. A third study found anxiety scores to be higher in the levothyroxine treated group. In another study of 20 individuals, two in the treatment arm dropped out, one for “nervousness” and another for “a sense of tachycardia.” The final study reported that within the treatment group a reduction in SF-36 vitality scores was found.

None of newer studies of treatment for either subclinical hypothyroidism or subclinical hyperthyroidism systematically evaluated harms. An assessment of harms was likely not a part of any of the studies’ protocols, nor does it appear that study participants were provided with a list of potential harms and asked to identify those that they experienced. Of the six studies of treatment of subclinical hypothyroidism included in the current review, only one reported on harms, stating that none of the patients reported side effects that would have required withdrawal or dose reduction. One of the other five reported that one participant withdrew from the study after reporting side effects from 12 days of placebo treatment; a second reported that one subject in the placebo group dropped out after six months because of serious disease unrelated to thyroid function.

The long-term adverse effects of levothyroxine therapy may depend on careful clinical and laboratory monitoring and adjustment of dosage accordingly. The previous systematic review demonstrated that overtreatment with levothyroxine leading to undetectable TSH levels is common in practice. We did not identify more recent data to estimate the frequency of overtreatment in current practice.
For subclinical hyperthyroidism, one of the two studies of treatment reported that they “did not find any increase in confusion, myopathy, atrial fibrillation, and deep tendon reflexes incidence” in either the treatment or observation group before or after treatment/observation. The second study did not discuss harms.

While none of the six included subclinical hypothyroidism studies discussed unnecessary treatment as a harm, three of the studies reported a decline in the TSH value in the placebo subjects. This suggests that at least some individuals who are classified with subclinical hypothyroidism will spontaneously improve, and therefore, unnecessary treatment can occur if a strategy of early treatment is undertaken.
Discussion

Tables 21 and 22 summarize our findings by key question.

Benefits of Screening/Treatment

No comparative studies have evaluated the benefits of screening for thyroid disease versus no screening (Key Question 1). Thus, it remains unknown if screening the general population with a serum TSH test will improve outcomes. The rationale for screening is based on the premise that subclinical hypothyroidism and hyperthyroidism have an inherent detrimental effect on health. We did not systematically review the literature concerning the relationship between mildly elevated TSH levels and long-term cardiovascular and cognitive effects. In previous systematic reviews, the relationship between subclinical hypothyroidism and risk factors for coronary disease was inconsistent. More recent studies, in particular the reanalysis of the Whickham data, suggest that subclinical hypothyroidism may be associated with increased cardiovascular morbidity over 20 years, but the validity and practical value of this observation is unclear. For example, whether a mildly elevated TSH confers additional risk for coronary events in the current era of widespread screening and treatment of hyperlipidemia is unclear, and the new data do not permit any estimate of the additional predictive ability of a TSH level above that of widely used risk assessment instruments based on the Framingham study.

Additionally, it is still unknown if treating individuals found to have subclinical thyroid dysfunction through population-based screening is beneficial compared with watchful waiting. In attempting to demonstrate the benefit of treatment, the literature contained a fair number of rigorous studies that looked at a variety of physiological parameters. These studies represent potentially important work that might someday help explain all the effects of thyroid hormone within the body and help explain why thyroid hormone is beneficial. However, we found only six trials of good or fair quality that reported outcomes that were in our analytic framework. No studies assessed effects on cardiac events. None of the studies found an improvement in patient-centered outcomes. Within the included studies, BMI/weight and measures of well-being were the only patient-centered outcomes evaluated, and improvement with treatment was not found. Four studies evaluated changes in intermediate outcomes that were included in our analytic framework. The findings regarding lipids were inconsistent, with two of the four studies finding some improvement and two finding no change. Neither of the two studies evaluating blood pressure found a significant change.
Table 21. Summary of evidence for subclinical hypothyroidism

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study Type: Number of studies Number of subjects</th>
<th>Risk of Bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect (strength of evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>KQ 2. What are the harms of screening for subclinical thyroid dysfunction?</td>
<td>2 studies</td>
<td>NA</td>
<td>No major inconsistency</td>
<td>Indirect; Low to moderate applicability to primary care</td>
<td>NA</td>
<td>No RCTs were identified; two natural history studies demonstrating indirect evidence of potential harm were included</td>
<td>(Insufficient)</td>
</tr>
<tr>
<td>KQ 3. Does treatment of patients with subclinical hypothyroidism detected by screening affect outcomes?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular events, coronary artery disease, and heart failure</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence (Insufficient)</td>
<td></td>
</tr>
<tr>
<td>Overall quality of life</td>
<td>RCT: 2 169</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; Moderate applicability to primary care setting</td>
<td>Imprecise: Small studies of 100 and 69 subjects</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Key Question</td>
<td>Study Type: Number of studies Number of subjects</td>
<td>Risk of Bias</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Comments</td>
<td>Magnitude of effect (strength of evidence)</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Changes in mood/cognition</td>
<td>RCT: 2 169</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; Moderate applicability to primary care setting</td>
<td>Imprecise; Small studies of 100 and 69 subjects</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Weight/BMI changes</td>
<td>RCT: 4 305</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect</td>
<td>Imprecise; The largest study had 100 subjects; the smallest had 23</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Blood pressure changes</td>
<td>RCT: 2 195</td>
<td>Medium</td>
<td>Consistent</td>
<td>Indirect; Low applicability to asymptomatic patients in primary care settings</td>
<td>Imprecise</td>
<td>No effect (Low)</td>
<td></td>
</tr>
<tr>
<td>Changes in lipid levels</td>
<td>RCT: 4 379</td>
<td>Medium</td>
<td>Inconsistent</td>
<td>Indirect; Low applicability to asymptomatic patients in U.S. primary care settings</td>
<td>Imprecise</td>
<td>Small effect for LDL and total cholesterol (Low)</td>
<td></td>
</tr>
<tr>
<td>KQ4. What are the harms of treatment of subclinical hypothyroidism and subclinical hyperthyroidism?</td>
<td>No Studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>(Insufficient)</td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI=body mass index; LDL=low-density lipoprotein; NA=not applicable; RCT=randomized controlled trial.
Table 22. Summary of evidence for subclinical hyperthyroidism

<table>
<thead>
<tr>
<th>Key Question</th>
<th>Study Type: Number of studies Number of subjects</th>
<th>Risk of Bias</th>
<th>Consistency</th>
<th>Directness</th>
<th>Precision</th>
<th>Comments</th>
<th>Magnitude of effect strength of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>KQ1. Does screening for subclinical thyroid dysfunction reduce morbidity or mortality?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>KQ 2. What are the harms of screening for subclinical thyroid dysfunction?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>KQ 3. Does treatment of patients with subclinical hyperthyroidism detected by screening affect outcomes?</td>
<td>Cardiovascular events including angina, atrial fibrillation, and other clinically significant arrhythmias</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td></td>
<td>Fractures</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td></td>
<td>Overall quality of life</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td></td>
<td>Changes in mood/cognition</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>No evidence</td>
<td>No evidence (Insufficient)</td>
</tr>
<tr>
<td>Key Question</td>
<td>Study Type: Number of studies</td>
<td>Number of subjects</td>
<td>Risk of Bias</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Comments</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------------------------</td>
<td>--------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
<td>-----------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Weight/BMI changes</td>
<td>Controlled trial: 1</td>
<td>14</td>
<td>High</td>
<td>NA</td>
<td>Direct</td>
<td>Imprecise</td>
<td>About 1% greater decrease in BMI in treated as compared to placebo group; absolute change in BMI in treated group of 0.5 kg/m² (Insufficient)</td>
</tr>
<tr>
<td>Blood pressure changes</td>
<td>RCT: 1</td>
<td>20</td>
<td>High</td>
<td>Inconsistent</td>
<td>Indirect</td>
<td>Imprecise</td>
<td>2.58 mmHg reduction daytime systolic blood pressure from 1 study; no change in 2nd study (Insufficient)</td>
</tr>
<tr>
<td>Changes in bone density (as measured by DEXA scan)</td>
<td>RCT: 1</td>
<td>20</td>
<td>High</td>
<td>NA</td>
<td>Indirect</td>
<td>Imprecise</td>
<td>No effect (Insufficient)</td>
</tr>
<tr>
<td>Key Question</td>
<td>Study Type: Number of studies</td>
<td>Risk of Bias</td>
<td>Consistency</td>
<td>Directness</td>
<td>Precision</td>
<td>Comments</td>
<td>Magnitude of effect strength of evidence</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>--------------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Changes in lipid levels</td>
<td>RCT: 1 20</td>
<td>High</td>
<td>NA</td>
<td>Indirect; Treated subjects included 2 patients with Graves disease and 8</td>
<td>Imprecise</td>
<td>No effect (Insufficient)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>with autonomous nodules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ4. What are the harms of treatment of subclinical hyperthyroidism and subclinical hyperthyroidism?</td>
<td>No studies</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Insufficient</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BMI=body mass index; DEXA=dual-energy X-ray absorptiometry; NA=not applicable; RCT=randomized controlled trial
The literature surrounding subclinical hyperthyroidism was particularly sparse; we only found two poor quality controlled trials. One found a slight improvement in weight gain after 12 months of treatment with methimazole, but no change in blood pressure. The other study found a small decrease in the mean daytime systolic blood pressure.\textsuperscript{77,83}

**Harms of Screening/Treatment (Key Questions 2 and 4)**

The harms of screening are still poorly studied. Indirect evidence appears to indicate that overdiagnosis, and thus, over-treatment could be a potential problem.

The harms of treatment also remain insufficiently studied. We found no studies that primarily focused on the harms of treatment. Of the studies that evaluated the benefits of treating subclinical hypothyroidism, only one reported that no subjects had side effects that required withdrawal or dose reduction.\textsuperscript{86} One study that evaluated the benefits of subclinical hyperthyroidism, reported simply that there was no “increase in confusion, myopathy, atrial fibrillation, and deep tendon reflexes” with treatment.\textsuperscript{77}

**Strength of Evidence**

The EPC strength of evidence assessment involves assessing the body of literature based on four domains: risk of bias, consistency, directness, and precision.

**Risk of Bias**

The possibility of bias or error within the studies that evaluated the benefits of treating subclinical hypothyroidism was moderate. Of the six trials that evaluated the benefits of treating subclinical hypothyroidism, one study was judged to be of good quality,\textsuperscript{84} the other five, fair quality. The fair-quality studies did not adequately describe their randomization, allocation, and/or blinding.\textsuperscript{80-82,85,86}

The possibility of bias within the studies that evaluated the benefits of treating subclinical hyperthyroidism was high. Both of these studies were of poor quality; assigning treatment or placebo via alternate assignment and offering the patients the options of switching arms once assigned\textsuperscript{83}, and not describing randomization, allocation concealment, or blinding.\textsuperscript{77} Additionally, in one study different treatments were used in the treatment arm.\textsuperscript{77}

**Consistency**

The four trials that evaluated the effect of thyroid replacement therapy on lipids in subjects with subclinical hypothyroidism were inconsistent. Two trials demonstrated improvement in lipids,\textsuperscript{84,85} while another two trials showed no improvement.\textsuperscript{81,86} Blood pressure changes with thyroid replacement were also evaluated, and two studies were consistent in demonstrating no improvement.\textsuperscript{84,86}

Of the two patient-centered outcomes evaluated in trials evaluating the benefits of treatment, four studies were consistent in showing no benefit of weight/BMI change,\textsuperscript{80,81,84,86} and two studies were consistent in failing to show improvement in measures of well-being.\textsuperscript{82,84}

Of the two studies evaluating the benefits of treating subclinical hyperthyroidism, both looked at blood pressure. One study found improvement in the mean, daytime systolic blood pressure\textsuperscript{77} while the other did not.\textsuperscript{83}
Directness

For the purposes of this report, directness refers to both the applicability of a study to the population specified for this review (i.e., asymptomatic, non-pregnant individuals seen in primary care settings) and to the degree to which one can directly associate the intervention to a patient-centered outcome. As defined by our analytic framework, these include: coronary artery disease, congestive heart failure, measures of well-being, weight change, progression to overt disease, osteoporotic fractures, and atrial fibrillation. None of the trials included in this review found improvement in these outcomes.

The only benefit found was for lipids, an intermediate outcome for which results were inconsistent.

Applicability

Because (1) the topic was nominated with regard to a particular practice setting (primary care); (2) the key questions address a specific population (asymptomatic, non-pregnant individuals); and (3) this review will be used by a body, the USPSTF, to make a recommendation regarding screening for subclinical thyroid dysfunction, this review considered directness to include the domain of applicability. Several aspects of these trials make their applicability to the general U.S. population a concern. First, none of the studies occurred in the U.S. Most of the studies recruited subjects from specialty clinics, rather than from the primary care setting. Additionally, the studies were relatively short, the longest lasting 12 months. The applicability of studies of this duration to a condition that is generally life-long was unclear.

Precision

The studies included in the review were small. Among studies of subclinical hypothyroidism, the largest contained 120 individuals, with the smallest having only 23 subjects. Of the two studies of subclinical hyperthyroidism, one had 20 subjects and one 14. Due to the small sample size, none of the included studies were judged to provide a precise estimation of the effect of treatment.

Limitations

This review was designed to address a very narrow clinical question that lies in the uncertainty of modern clinical practice: What are the benefits and harms of screening and subsequently treating subclinical thyroid dysfunction. Not only is this a critical question to answer for patient management, but also to determine if screening for subclinical thyroid dysfunction is warranted. In particular, not only would screening find individuals with subclinical thyroid dysfunction, but it would also identify individuals with overt thyroid disease. While it remains unclear just how many individuals will have undetected overt thyroid dysfunction in primary care practices, a large population-based prevalence study performed in Colorado estimates that about 0.5 percent of the population might have unrecognized overt thyroid dysfunction. While it was unclear how many of these individuals might benefit from screening in primary care practices, it is likely that some would; thus, the benefit of screening may be underestimated by this review. Moreover, unlike the previous AHRQ report, we did not systematically review studies of the natural history of subclinical hypothyroidism or reevaluate literature on this subject that was included in previous reviews.
Additionally, this review compared the benefits and harms of screening with not screening for subclinical thyroid dysfunction in primary care practices. The benefits and harms of a third approach, case-finding, was not evaluated. As described earlier in this review, case-finding is different from screening in that after a TSH is obtained on all individuals, those with a high TSH are brought back for a detailed thyroid focused history. The decision to treat is then based on both the lab value and on the history obtained. This approach may allow those who will most benefit from treatment to be identified in the most efficient manner. By not including case-finding as a diagnostic strategy, it is possible that this review failed to evaluate a valid method of determining thyroid dysfunction in primary care practices.

Most of the studies included in this review did not recruit patients from primary care practices. Additionally, one of the systematic reviews included contained three studies that had individuals with known thyroid dysfunction. Ideally, we would have included only studies that included screen-detected thyroid dysfunction from primary care practices; however, the paucity of studies compelled the use of less than ideal studies. This limits the applicability to primary care practices and to screened populations.

Finally, the individual studies that were included in this review used different lab values as the cutoff for an abnormal TSH and for normal FT4 and used different doses of levothyroxine, making comparisons between studies difficult. Because of this clinical heterogeneity, a meta-analysis was not conducted.

Summary of Review Findings

As in 2004, there remains insufficient evidence directly linking screening of asymptomatic individuals for subclinical thyroid dysfunction to improvements in morbidity or mortality. However, a growing body of evidence, which this review rates as low quality, continues to suggest that treating subclinical hypothyroidism is not any more beneficial than watchful waiting for overall quality of life, mood or cognition, weight/BMI, or blood pressure changes. Studies regarding lipids remain inconsistent, small, and of varying quality. We, therefore, regard the evidence concerning lipids as insufficient to determine if treatment is more beneficial than watchful waiting. Larger trials of treatment that are longer in duration would be helpful in improving the quality of evidence for all of these outcomes.

The following information from this report may be helpful to the practicing clinician:

- It is still unclear if subclinical thyroid dysfunction has any impact on patient outcomes; systematic reviews included in this report regarding all-cause mortality and coronary artery disease mortality were inconsistent, and indicated that even if treatment is effective, a clinician would need to screen and treat a large number of patients before seeing a benefit to any individual patient
- We know that in some natural history studies, as many as 37 percent of patients initially found to have an elevated TSH and normal free T4 may have a normal TSH within the three years following initial testing
- We know, therefore, that there is potential for serious harms, including atrial fibrillation and angina, from unnecessary treatment or overtreatment of subclinical thyroid dysfunction, although in most cases reported harms appear to be minor, such as increased anxiety and nervousness, and are likely to improve with discontinuation of therapy
- The strength of evidence is low that treating subclinical hypothyroidism will lead to improvement in quality of life, mood, cognition, weight/BMI, or blood pressure when compared with watchful waiting
Finally, we know that the current evidence is inconsistent, with a low overall strength of evidence with regards to the benefits of treatment of subclinical hypothyroidism for lipids, that improvement is likely to be modest at best, and is likely to be of little clinical significance.

As in all cases, this evidence should be considered, along with the clinician’s knowledge of the individual, in making clinical judgments about the appropriate care for each patient. For instance, one might consider levothyroxine therapy in an individual with both subclinical hypothyroidism and dyslipidemia who also reports several symptoms of thyroid dysfunction, but who is at low risk for atrial fibrillation and other possible side effects. However, an individual at higher risk of atrial fibrillation, who has no or few hypothyroid symptoms, or who is known to be prone to anxiety might not be a good candidate for treatment.

Emerging Issues/Next Steps

Currently, progression to overt disease is the best established complication of subclinical thyroid dysfunction. Other possible complications of subclinical hypothyroidism still need to be fully delineated. The Whickham study was recently re-analyzed and found that the original study might have underestimated the association between subclinical hypothyroidism and risk for coronary events over 20 years. Other analyses suggest that subclinical hypothyroidism may be associated with cardiovascular disease in subjects younger than 65 years, but that the magnitude of risk is low. These analyses, however, do not establish the prognostic value of a TSH when added to existing risk scoring systems for cardiovascular events, particularly in the age of aggressive lipid therapy. Continuing work in this area should improve our understanding of the risks of cardiovascular complications in people with subclinical thyroid dysfunction and how those risks vary among different subpopulations.

Additionally, reversion of mildly elevated TSH levels may be more common than previously thought. Recent studies suggest that reversion can occur even after a long period of persistent elevation. A recent longitudinal study found that of 102 women with TSH between 0.1 and 0.4 mIU/L 3 percent progressed to overt hyperthyroidism while 24 percent reverted to normal. In a cross-sectional study of individuals in their 80s, after three years, of 12 subjects initially found to have subclinical hyperthyroidism, only one had progressed to overt disease and five had reverted to normal TSH levels. In that same study, in 21 out of 30 individuals reassessed after initially found to have subclinical hypothyroidism, none had progressed to overt disease, and 11 had reverted to normal TSH levels. Studies to discover who might progress to overt disease versus who might revert back to a normal TSH and the time frame for which this might occur would be helpful in determining who should be considered for early treatment and for whom watchful waiting might be more appropriate.

Finally, three recent studies suggest that in the elderly population a slightly elevated TSH might have a protective effect, questioning the conventional wisdom surrounding treating subclinical hypothyroidism. Taken together, these findings suggest the need for a better understanding of the impact on health outcomes of mildly elevated and mildly depressed TSH levels and how those impacts vary for different subpopulations, particularly the elderly.

Future Research

Because the prevalence of subclinical thyroid dysfunction in the general population may be as high as 7 percent in women and 5 percent in men, determining the harms and benefits of
screening and treating thyroid disease has the potential to significantly affect the health of many. Many experts who favor early treatment acknowledge the need for appropriately powered RCTs comparing early treatment with thyroxine therapy to active surveillance or expectant management in patients who have a mildly increased serum TSH level.

Prior to conducting trials of screening for subclinical thyroid dysfunction, it may be important to conduct well-designed trials of treatment versus delayed treatment. Any condition that is considered appropriate for screening must have an intervention that can clearly improve a clinical outcome. It would be ideal if the benefits of treatment and an understanding of which individuals may benefit from treatment were understood prior to determining if screening the general population is warranted. For a condition to be considered appropriate for screening, there must be an intervention that can improve a clinical outcome. To date, it remains unclear if treating subclinical thyroid dysfunction leads to improvement in patient-centered outcomes. Furthermore, subclinical thyroid dysfunction may only pose a health risk to a particular subpopulation, and if there is a health risk, treatment may only be effective beyond a specific TSH threshold. Well designed trials of early versus delayed treatment should be conducted to determine this. If a treatment benefit is found, then well designed screening trials should be conducted to determine if screening in the general population is beneficial.

Any trial designed to answer the key questions in this review should follow subjects for at least three years, preferably longer, and should be powered to detect the harms of therapy. Such a trial should recruit subjects who were identified by screening and preferably from primary care practices. Outcomes of interest should include patient-centered outcomes, similar to the ones outlined in our analytic framework, and any tool used to assess quality of life should be well-recognized and have a proven track-record. The validity of such a trial can be increased by requiring a longer pre-randomization period to ensure that subjects have persistently high TSH levels, and by including and stratifying subjects who have serologic evidence of autoimmunity.


21. Surks MI, Ortiz E, Daniels GH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA 2004 Jan 14;291(2):228-238.


52. Rosario PW. Natural history of subclinical hyperthyroidism in elderly patients with TSH between 0·1 and 0·4 mIU/l: a prospective study. Clinical Endocrinology 2010;72(5):685-688.


Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOG</td>
<td>The American College of Obstetricians and Gynecologists</td>
</tr>
<tr>
<td>ACP</td>
<td>The American College of Physicians</td>
</tr>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>AMSTAR</td>
<td>Assessment of Multiple Systematic Reviews</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>EHCP</td>
<td>Effective Health Care Program</td>
</tr>
<tr>
<td>EPC</td>
<td>Evidence-based Practice Center</td>
</tr>
<tr>
<td>HDL</td>
<td>High-density Lipoprotein</td>
</tr>
<tr>
<td>kg/m²</td>
<td>kilograms per meter squared (a measure of body mass index)</td>
</tr>
<tr>
<td>KQ</td>
<td>Key question</td>
</tr>
<tr>
<td>LDL</td>
<td>Low-density Lipoprotein</td>
</tr>
<tr>
<td>mcg</td>
<td>microgram</td>
</tr>
<tr>
<td>mg/dL</td>
<td>milligrams per deciliter</td>
</tr>
<tr>
<td>mIU/L</td>
<td>milli-international units per liter</td>
</tr>
<tr>
<td>mIU/L²</td>
<td>milli-international units per liter squared</td>
</tr>
<tr>
<td>mmHg</td>
<td>millimeters of mercury</td>
</tr>
<tr>
<td>ng/dl</td>
<td>nanograms per deciliter</td>
</tr>
<tr>
<td>PICOTS</td>
<td>Population, Intervention, Comparators, Outcome, Timing, and Setting</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
<tr>
<td>SIP</td>
<td>Scientific Information Packet</td>
</tr>
<tr>
<td>TPO</td>
<td>Thyroid Peroxidase</td>
</tr>
<tr>
<td>TSH</td>
<td>Thyroid Stimulating Hormone</td>
</tr>
<tr>
<td>T3</td>
<td>Triiodothyronine</td>
</tr>
<tr>
<td>T4</td>
<td>Thyroxine</td>
</tr>
<tr>
<td>USPSTF</td>
<td>United States Preventive Services Task Force</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Appendix A. Lipid Conversion Factors

**To Convert From mmol/L to mg/dL**
For total, HDL, and LDL cholesterol multiply mmol/L by 38.67
e.g. 3.5 mmol/L = 3.5 mmol/L * 38.67 = 135 mg/dL

For triglycerides multiply mmol/L by 88.57
e.g. 1.9 mmol/L = 1.9 mmol/L * 88.57 = 168 mg/dL

**To Convert From mg/dL**
For total, HDL, and LDL cholesterol divide mg/dL by 38.67
e.g. 135 mg/dL = 135 mg/dL/38.67 = 3.5 mmol/L

For triglycerides divide mg/dL by 88.57
e.g. 168 mg/dL = 168 mg/dL/88.57 = 1.9 mmol/L
Appendix B. Search Strategies

### Search Strategy for Systematic Reviews and Observational Studies

<table>
<thead>
<tr>
<th>Set #</th>
<th>Concept</th>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE (PubMed) 3/3/10</td>
<td>#1 Sub-clinical Thyroid Disease</td>
<td>MEDLINE (PubMed) 3/3/10 Sub-clinical Thyroid Disease (Hyperthyroidism [Mesh] OR Hypothyroidism [Mesh] OR (hypothy* OR hypo-thyr* OR hyperthy* OR hyper-thyr*) [tiab] OR (thyroid deficien* OR thyroid insufficien* OR thyroid failure) [tiab]) AND (mild OR compens* OR subclinic* OR moderat* OR short-term) [tiab]) OR elevated tsh [tiab])</td>
</tr>
<tr>
<td></td>
<td>#2 Screening</td>
<td>Mass Screening [Mesh] OR Thyroid Function Tests [Mesh] OR (screening OR casefinding OR case finding) [tiab]</td>
</tr>
<tr>
<td></td>
<td>#5 Screening or treatment of sub-clinical thyroid disease</td>
<td>#1 AND (#2 OR #3 OR #4)</td>
</tr>
<tr>
<td></td>
<td>#6 Systematic reviews</td>
<td>systematic[sb]</td>
</tr>
<tr>
<td></td>
<td>#7 Observational Studies</td>
<td>(((((cohort studies[MeSH Terms]) OR comparative study[MeSH Terms]) OR follow-up studies[MeSH Terms]) OR prospective studies[MeSH Terms]) OR risk factors[MeSH Terms]) OR cohort[Title/Abstract]) OR compared[Title/Abstract]) OR groups[Title/Abstract]) OR multivariate[Title/Abstract]</td>
</tr>
<tr>
<td></td>
<td>Systematic Reviews of screening or treatment for subclinical thyroid disease</td>
<td>#6 AND #5 N=39</td>
</tr>
<tr>
<td></td>
<td>Observational Studies of Screening for subclinical thyroid disease</td>
<td>#1 AND #2 AND #7 limited to year 2002 and beyond N=236</td>
</tr>
<tr>
<td></td>
<td>[MeSH] = exploded term, all subterms and subtrees included [tiab] = in title and abstract * = truncation</td>
<td></td>
</tr>
<tr>
<td>Set #</td>
<td>Concept</td>
<td>Search Strategy</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>Cochrane Register Database of Systematic Reviews and Database of Abstracts of Reviews of Effects (OVID) 3/3/2010</td>
<td>#1 Sub-clinical Thyroid Disease</td>
<td>((hyperthyroidism.mp. OR hypothyroidism.mp. OR (hypothyroid* or hypo-thyr* or hyperthyroid* or hyper-thyr*).mp. OR (thyroid deficiency* or thyroid insufficiency* or thyroid failure).mp.) AND ((mild or compensate* or subclinical* or moderate* or short-term).mp.)) OR elevated adj3 tsh.mp.</td>
</tr>
<tr>
<td></td>
<td>#2 Screening</td>
<td>mass screening.mp. or OR thyroid function tests.mp. or OR screening.mp. OR case finding.mp.</td>
</tr>
<tr>
<td></td>
<td>#3 Treatment for hyperthyroidism</td>
<td>Antithyroid Agents.mp. OR anti-thyroid.mp. OR methimazole.mp. OR propylthiouracil.mp. OR radioiodine.mp. or radioactive.mp.</td>
</tr>
<tr>
<td></td>
<td>#4 Treatment for hypothyroidism</td>
<td>t3.mp. OR t4.mp. OR thyroxine.mp. OR levothyroxine.mp. OR triiodothyronine.mp. OR liothyronine.mp. OR thyrolar.mp. or liotrix.mp.</td>
</tr>
<tr>
<td>Systematic Reviews of screening or treatment for subclinical thyroid disease</td>
<td>#1 AND (#2 OR #3 OR #4) N=40 after deduplication 35</td>
<td></td>
</tr>
</tbody>
</table>
### Search Strategy for Individual Studies

<table>
<thead>
<tr>
<th>Set#</th>
<th>Concept</th>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Sub-clinical Thyroid Disease</td>
<td>(Hyperthyroidism [mesh] OR Hypothyroidism [mesh][hypothyr* OR hypo-thyr* OR hyperthyr* OR hyperthyr*] [tiab] OR (thyroid deficien* OR thyroid insufficient* OR thyroid failure) [tiab]) AND (mild OR compens* OR subclinic* OR moderat* OR short-term) [tiab]) OR elevated tsh [tiab]</td>
</tr>
<tr>
<td>#2</td>
<td>Screening</td>
<td>Mass Screening [mesh] OR Thyroid Function Tests [mesh] OR (screening OR casefinding OR case finding) [tiab]</td>
</tr>
<tr>
<td>#5</td>
<td>Treatment in general</td>
<td>(clinical[Title/Abstract] AND trial[Title/Abstract]) OR clinical trials[MeSH Terms] OR clinical trial[Publication Type] OR random*[Title/Abstract] OR random allocation[MeSH Terms] OR therapeutic use[MeSH Subheading])</td>
</tr>
<tr>
<td>#6</td>
<td>Harms</td>
<td>(adverse effects [sh] OR poisoning [sh] OR toxicity [sh] OR contraindications [sh] OR complications [sh] OR (safe OR safety OR side-effect$ OR undesirable effect$ OR treatment emergent OR tolerability OR toxicity) [tw] OR (adverse [tiab] AND (effect$ OR reaction$ OR event$ OR outcome$) [tiab])</td>
</tr>
</tbody>
</table>

**KQ1 Screening for Subclinical Thyroid Disease**

#1 AND #2 AND #8

**KQ2 Harms of Screening**

(#6 OR #7) AND (#1 AND #2)

**KQ3a treatment hyper**

#1 AND #3 AND #5

**KQ3b treatment hypo**

#1 AND #4 AND #8

**KQ4 harms of treatment**

(#3 OR #4) AND #6 AND #1

---

* [MeSH] = exploded term, all subterms and subtrees included
  * [mesh: noexp] = non-explored term, no subterms or subtrees included
  * [tiab] = in title and abstract
  * [tw] = word anywhere in the record
  * [pt] = publication type
  * * = truncation
  * all searches were limited to publication date 2002 or later
<table>
<thead>
<tr>
<th>Cochrane Register of Controlled Trials (OVID)</th>
<th>Set #</th>
<th>Concept</th>
<th>Search Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#1</td>
<td>Sub-clinical Thyroid Disease</td>
<td><em>((hyperthyroidism.mp. or exp Hyperthyroidism/) OR exp Hypothyroidism/ or hypothyroidism.mp. OR (hypothyrid</em> or hypo-thyr* or hyperthyrid* or hyper-thyr*).mp. OR (thyroid deficien* or thyroid insufficien* or thyroid failure).mp.) AND ((mild or compens* or subclinic* or moderat* or short-term).mp.)) OR elevated adj3 tsh.mp.</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>Screening</td>
<td>mass screening.mp. or exp Mass Screening/ OR thyroid function tests.mp. or exp Thyroid Function Tests/ OR screening.mp. OR casefinding.mp OR case finding.mp</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>Treatment for hyperthyroidism</td>
<td>exp Antithyroid Agents/ or anti-thyroid.mp. OR methimazole.mp. or exp Methimazole/ OR propylthiouracil.mp. or exp Propylthiouracil/ OR radioiodine.mp. or radioactive.mp.</td>
</tr>
<tr>
<td></td>
<td>#4</td>
<td>Treatment for hypothyroidism</td>
<td>t3.mp. OR t4.mp. OR exp Thyroxine/ or thyroxine.mp. OR levothyroxine.mp. OR exp Triiodothyronine/ or triiodothyronine.mp. OR liothyronine.mp. OR thyrolar.mp. or liotrix.mp.</td>
</tr>
</tbody>
</table>

KQ1/2 Screening #1 AND #2

KQ3/4 Treatment #2 AND (#3 OR #4)

all searches limited to publication date 2002 to current

| AGELINE (AARP.org) | hyperthyroidism ; hypothyroidism ; "thyroid deficien**" ; "thyroid insufficien**" ; "thyroid failure" AND "mild; compens* ; subclinic* ; moderat* ; "short-term" [ all content any position ; = OR] |
**EMBASE (EMBASE.com)**

| Screening Search (KQ1 and KQ2) | `('hyperthyroidism'/syn OR 'hypothyroidism'/syn OR 'thyroid deficien' OR 'thyroid insufficien' OR 'thyroid failure') AND (mild OR compens* OR subclinic* OR moderat* OR 'short-term') OR 'elevated tsh') AND ('mass screening' OR 'thyroid function tests' OR screening OR casefinding OR case AND finding) ` |
| Treatment KQ3 and KQ4 | `('hyperthyroidism'/syn OR 'hypothyroidism'/syn OR 'thyroid deficien' OR 'thyroid insufficien' OR 'thyroid failure') AND (mild OR compens* OR subclinic* OR moderat* OR 'short-term') OR 'elevated tsh') AND ('t3'/exp OR t4 OR 'thyroxine'/exp OR 'levothyroxine'/exp OR 'triiodothyronine'/exp OR 'liothyronine'/exp) AND ('therapeutic use' OR treatment OR 'therapy'/exp) OR ('anti thyroid' OR ('methimazole'/exp OR 'methimazole') OR propythiouracil)) AND (random OR ('placebo'/exp OR 'placebo') OR 'double blind')` |
| Harms of Treatment KQ4 | `('hyperthyroidism'/syn OR 'hypothyroidism'/syn OR 'thyroid deficien' OR 'thyroid insufficien' OR 'thyroid failure') AND (mild OR compens* OR subclinic* OR moderat* OR 'short-term') OR 'elevated tsh') AND ('t3'/exp OR t4 OR 'thyroxine'/exp OR 'levothyroxine'/exp OR 'triiodothyronine'/exp OR 'liothyronine'/exp) AND ('therapeutic use' OR treatment OR 'therapy'/exp) OR ('anti thyroid' OR ('methimazole'/exp OR 'methimazole') OR propythiouracil)) AND ('adverse effects' OR ('poisoning'/exp OR 'poisoning') OR ('toxicity'/exp OR 'toxicity') OR contraindications OR complications OR safe OR ('safety'/exp OR 'safety') OR 'side effect'$ OR undesirable AND effect$ OR treatment AND emergent OR tolerability OR ('toxicity'/exp OR 'toxicity')) OR (adverse AND (effect$ OR reaction$ OR event$ OR outcome$))` |

**Grey Literature**

<table>
<thead>
<tr>
<th>Type of information</th>
<th>Sources Searched</th>
</tr>
</thead>
</table>
| Clinical Trials     | ClinicalTrials.gov
| Regulatory Information | FDA
| Conference Proceedings | Conference Papers Index
| Other | NIH RePORTER (a searchable database of federally funded biomedical research projects conducted at universities, hospitals, and other research institutions)
| | HSRPROJ (a database providing access to ongoing grants and contracts in health services research)
| | Hayes, Inc. Health Technology Assessment
| | NY Academy of Medicine’s Grey Literature Index
## Foreign Language Search Strategy

<table>
<thead>
<tr>
<th>Set#</th>
<th>Concept</th>
<th>Search strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CINAHL (EBSCO) 1-15-2010</strong></td>
<td>#1 S17 N=266</td>
<td>Sub-clinical Thyroid Disease</td>
</tr>
<tr>
<td>#2 S22 N=53698</td>
<td>Screening</td>
<td>(MH &quot;Health Screening+&quot;) OR (MH &quot;Thyroid Function Tests+&quot;) OR (ti screening OR ti casefinding OR ti case finding) OR (ab screening OR ab casefinding OR ab case finding )</td>
</tr>
<tr>
<td>#3 S28 N=382</td>
<td>Treatment for hyperthyroidism</td>
<td>radioiodine OR ti radioactive OR (&quot;anti-thyroid&quot;) or (MH &quot;Thyroid Antagonists&quot;) OR propthyiouracil OR methimazole</td>
</tr>
<tr>
<td>#4 S41 N=1082</td>
<td>Treatment for hypothyroidism</td>
<td>(liotrix OR thyrolar OR liothyronine OR triiodothryonine OR &quot;levothryoxine&quot; OR (&quot;thyroxine&quot;) OR (MH &quot;Thyroxine&quot;) OR ti t4 OR ab t4 OR ti T3 OR ab T3) AND (ti treatment OR ab treatment OR ti therapy OR ab therapy OR (&quot;therapeutic use&quot;) or (MH &quot;Therapeutics&quot;))</td>
</tr>
<tr>
<td>#5 S48 N=647228</td>
<td>Treatment in general</td>
<td>(MH &quot;Therapeutics+&quot;) OR (MH &quot;Random Assignment&quot;) OR ti random* OR ab random* OR pt clinical trial OR (MH &quot;Clinical Trials+&quot;) OR (ti clinical OR ab clinical) AND (ti trial OR ab trial)</td>
</tr>
<tr>
<td>#6 S55 N=372918</td>
<td>Harms</td>
<td>(&quot;adverse effects&quot;) OR (MH &quot;Adverse Drug Event&quot;) OR (MH &quot;Adverse Health Care Event&quot;) OR (&quot;poisoning&quot;) or (MH &quot;Poisoning&quot;) OR (&quot;toxicity&quot;) or (MH &quot;Drug Toxicity&quot;) OR contraindication* or (&quot;complication&quot;) or (MH &quot;Treatment Complications, Delayed&quot;) or (&quot;safety&quot;) or (MH &quot;Safety&quot;)</td>
</tr>
<tr>
<td>#7 S63 N=149782</td>
<td>Patient issues of testing</td>
<td>patient preference* OR consequence* OR cost OR costs OR false positive* OR acceptability OR worry OR (MH &quot;False Positive Results&quot;) OR (MH &quot;False Negative Results&quot;) OR (MH &quot;Attitude to Health&quot;) OR (MH &quot;Stress&quot;) OR (MH &quot;Family&quot;) OR (MH &quot;Perception&quot;) OR (MH &quot;Patient Satisfaction&quot;)</td>
</tr>
<tr>
<td>#8 S68 N=80675</td>
<td>RCTs</td>
<td>(MH &quot;Clinical Trials&quot;) OR ti randomized OR ti placebo OR ti randomly</td>
</tr>
</tbody>
</table>

KQ1 Screening for Subclinical Thyroid Disease #1 AND #2 AND #8 (2 results both in English)  
KQ2 Harms of Screening (#6 OR #7) AND (#1 AND #2) (42 results, 2 non-English)  
KQ3a treatment hyper #1 AND #3 AND #5 (4 results all in English)  
KQ3b treatment hypo #1 AND #4 AND #8 (10 results, 1 non-English)  
KQ4 harms of treatment (#3 OR #4) AND #6 AND #1 (40 results, 2 non-English, one of these is duplicate)  
4 non-English articles found

1/15/2010
WHO Global Health Library (includes African Index Medicus, Index Medicus for the Eastern Mediterranean Region, Latin American and Caribbean Health Sciences Literature, Pan American Health Organization database, WHO Library Database, WHO Western Pacific Region database)

SEARCH: (Hyperthyroidism OR hypothyroidism ) AND subclinical AND YEAR>= 2002 AND non-English language
Appendix C. Criteria for Assessing Internal Validity of Individual Studies

The Methods Work Group for the US Preventive Services Task Force (USPSTF) developed a set of criteria by which the internal validity of individual studies could be evaluated. The USPSTF accepted the criteria, and the associated definitions of quality categories, that relate to internal validity at its September 1999 meeting.

This appendix describes the criteria relating to internal validity and the procedures that topic teams follow for all updates and new assessments in making these judgments.

All topic teams use initial “filters” to select studies for review that deal most directly with the question at issue and that are applicable to the population at issue. Thus, studies of any design that use outdated technology or that use technology that is not feasible for primary care practice may be filtered out before the abstraction stage, depending on the topic and the decisions of the topic team. The teams justify such exclusion decisions if there could be reasonable disagreement about this step. The criteria below are meant for those studies that pass this initial filter.

Presented below are a set of minimal criteria for each study design and then a general definition of three categories: “good,” “fair,” and “poor,” based on those criteria. These specifications are not meant to be rigid rules but rather are intended to be general guidelines, and individual exceptions, when explicitly explained and justified, can be made. In general, a “good” study is one that meets all criteria well. A “fair” study is one that does not meet (or it is not clear that it meets) at least one criterion but has no known “fatal flaw.” “Poor” studies have at least one fatal flaw.

Systematic Reviews

Criteria

• Comprehensiveness of sources considered/search strategy used
• Standard appraisal of included studies
• Validity of conclusions
• Recency and relevance are especially important for systematic reviews

Definition of Ratings From Above Criteria

Good. Recent, relevant review with comprehensive sources and search strategies; explicit and relevant selection criteria; standard appraisal of included studies; and valid conclusions.

Fair. Recent, relevant review that is not clearly biased but lacks comprehensive sources and search strategies.

Poor. Outdated, irrelevant, or biased review without systematic search for studies, explicit selection criteria, or standard appraisal of studies.
Case-Control Studies

Criteria

- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variables

Definition of Ratings Based on Criteria Above

**Good.** Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.

**Fair.** Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.

**Poor.** Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

Randomized Controlled Trials and Cohort Studies

Criteria

- Initial assembly of comparable groups
  - for RCTs: adequate randomization, including first concealment and whether potential confounders were distributed equally among groups
  - for cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- All important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention to treat analysis for RCTs.

Definition of Ratings Based on Above Criteria

**Good.** Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used
and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention to confounders in analysis. In addition, for RCTs, intention to treat analysis is used.

**Fair.** Studies will be graded “fair” if any or all of the following problems occur, without the fatal flaws noted in the “poor” category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention to treat analysis is done for RCTs.

**Poor.** Studies will be graded “poor” if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For RCTs, intention to treat analysis is lacking.

### Diagnostic Accuracy Studies

**Criteria**

- Screening test relevant, available for primary care, adequately described
- Study uses a credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Handles indeterminate results in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Administration of reliable screening test

**Definition of Ratings Based on Above Criteria**

**Good.** Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 100) broad-spectrum patients with and without disease.

**Fair.** Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a “medium” spectrum of patients.

**Poor.** Has fatal flaw such as: Uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size or very narrow selected spectrum of patients.
Appendix D. Excluded Studies

Excluded Systematic Reviews


**Excluded Studies**

1R01AG032317-01A1. Subclinical thyroid dysfunction in the elderly. **No data (e.g. description of study only)**

1R01MH080295-01A2. Subclinical Hypothyroidism: Mood, Cognition and the effect of L-thyroxine treatme. **No data (e.g. description of study only)**

5R01HL076645-03. Subclinical Thyroid Dysfunction & Risk of MI and Stroke. **No relevance to key questions**

5R21DK062787-02. Neurocognitive effects of subclinical thyroid disease. **Related to topic but does not meet criteria for study design or quality**


Abbott Laboratories. **No data (e.g. description of study only)**

Abdelrazek SS, Rogowski F, Szumowski P, Zonenberg A, Parfienczyk A, Szelachowska M, et al. The Outcome of Radioiodine Therapy in Patient with Subclinical Hyperthyroidism. 19th Annual Congress of the European Association of Nuclear Medicine (EANM 06), Megaron International Conference Center, Athens (Greece), 30 Sep-4 Oct 20062006. p. P534 (333). **Related to topic but does not meet criteria for study design or quality**

Abdelrazek SS, Rogowski F, Zonenberg A, Szelachowska M, Nikolajuk A, Parfienczyk A, et al. The Effect of Radioiodine Therapy on Some Parameters of Oxidant/Antioxidant Balance in Patients with Subclinical Hyperthyroidism. 20th Annual Congress of the European Association of Nuclear Medicine (EANM 2007), Copenhagen Congress Center, Copenhagen (Denmark), 13-17 Oct 20072007. p. P445 (S353). **Related to topic but does not meet criteria for study design or quality**


Adrees M, Gibney J, El-Saeity N, Boran G. Effects of 18 months of L-T4 replacement in women with subclinical hypothyroidism. Clin Endocrinol2009. p. 298-303. **Related to topic but does not meet criteria for study design or quality**

Al Sayed A, Al Ali N, Bo Abbas Y, Alfadli E. Subclinical hypothyroidism is associated with early insulin resistance in Kuwaiti women. Endocr J. 2006/08/12 ed2006. p. 653-7. **No data (e.g. description of study only); No relevance to key questions**


Alzoubi KH, Gerges NZ, Aleisa AM, Alkadhi KA. Levothyroxin restores hypothyroidism-induced impairment of hippocampus-dependent learning and memory: Behavioral, electrophysiological, and molecular studies. Hippocampus. 2008/08/06 ed2009. p. 66-78. **Wrong population**


Bakiner O, Erteror ME, Haydardeoglu FE, Bozkirli E, Tutuncu NB, Demirag NG. Subclinical hypothyroidism is characterized by increased QT interval dispersion among women. Med Princ Pract. 2008/08/08 ed2008. p. 390-4. **No relevance to key questions; Wrong population**


Batrinos ML. The problem of exogenous subclinical hyperthyroidism. Hormones (Athens, Greece)2006. p. 119-25. **Wrong population**


Beltran S, Lescure FX, El Esper I, Schmit JL, Desailloud R. Subclinical hypothyroidism in HIV-infected patients is not an autoimmune disease. Horm Res. 2006/05/11 ed2006. p. 21-6. **Wrong population**

Bensenor I. Screening for thyroid disorders in asymptomatic adults from Brazilian populations. Sao Paulo Med J. 2002/11/19 ed2002. p. 146-51. **No data by design (e.g. opinion, letter, editorial)**


Bhasin S, Enzlin P, Coviello A, Basson R. Sexual dysfunction in men and women with endocrine disorders. Lancet2007. p. 597-611. **No data by design (e.g. opinion, letter, editorial); Wrong population**


Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocrine Reviews2008. p. 76-131. **No data by design (e.g. opinion, letter, editorial)**


Boelaert K, Franklyn JA. Thyroid hormone in health and disease. Journal of Endocrinology 2005. p. 1-15. *No data by design (e.g. opinion, letter, editorial)*


Burlacu MC, Socin HV, Luyckx F, Beckers A. Subclinical Hypothyroidism: Does TRH Test Help to Narrow Current Reference Ranges? 32nd Annual Meeting of the European Thyroid Association, Carl Ludwig Institute, University Hospital Leipzig, Leipzig (Germany), 1-5 Sep 20072007. p. 67. *Related to topic but does not meet criteria for study design or quality*


Choi AR, Manning P. Overshooting the mark: subclinical hyperthyroidism secondary to excess thyroid hormone treatment may be more prevalent than we realise. N Z Med J. 2009/03/24 ed2009. p. 93-4.  *No data by design (e.g. opinion, letter, editorial)*


Christ-Crain M, Meier C, Huber PR, Staub JJ, Muller B. Effect of L-thyroxine replacement therapy on surrogate markers of skeletal and cardiac function in subclinical hypothyroidism. Endocrinologist2004. p. 161-6.  *No data by design (e.g. opinion, letter, editorial)*


Col NF, Surks MI, Daniels GH. Subclinical Thyroid Disease: Clinical Applications. Journal of the American Medical Association2004. p. 239-43. *Background - review*


Cooper DS. Approach to the patient with subclinical hyperthyroidism. J Clin Endocrinol Metab. 2007/01/09 ed2007. p. 3-9. *Background - review*


Dardano A, Monzani F. Hypothyroidism and endothelial function: A marker of early atherosclerosis? Recent Patents on Endocrine, Metabolic and Immune Drug Discovery2008. p. 79-96. *No data by design (e.g. opinion, letter, editorial)*

Dashe JF, Cunningham FG. Subclinical hypothyroidism. New England Journal of Medicine2001. p. 1855-6. *No data by design (e.g. opinion, letter, editorial)*

Davis JD, Tremont G. Neuropsychiatric aspects of hypothyroidism and treatment reversibility. Minerva Endocrinol. 2007/03/14 ed2007. p. 49-65. *No data by design (e.g. opinion, letter, editorial)*


Djezairi AH. Subclinical Hypothyroidism in Older Patients: An Analysis of Natural Course and Risk Factors for the Thyroid Failure. 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. **No relevance to key questions**


Duggal J, Singh S, Barsano CP, Arora R. Cardiovascular risk with subclinical hyperthyroidism and hypothyroidism: pathophysiology and management. J Cardiometab Syndr. 2007/09/06 ed2007. p. 198-206. **No data by design (e.g. opinion, letter, editorial)**


Duman D, Sahin S, Esertas K, Demirtunc R. Simvastatin improves endothelial function in patients with subclinical hypothyroidism. Heart Vessels. 2007/03/29 ed2007. p. 88-93. **Related to topic but does not meet criteria for study design or quality**

Duntas LH, Wartofsky L. Cardiovascular risk and subclinical hypothyroidism: focus on lipids and new emerging risk factors. What is the evidence? Thyroid. 2007/09/29 ed2007. p. 1075-84. **No relevance to key questions; No data by design (e.g. opinion, letter, editorial)**


Fatourechi V. Subclinical hypothyroidism: When to treat, when to watch? Consultant2004. p. 533-9. No data by design (e.g. opinion, letter, editorial)

Fatourechi V. Subclinical hypothyroidism: Why wait to treat? Consultant2004. p. 904%N 7. No data by design (e.g. opinion, letter, editorial)


Fazio S, Palmieri EA, Lombardi G, Biondi B. Effects of thyroid hormone on the cardiovascular system. Recent Prog Horm Res. 2004/01/30 ed2004. p. 31-50. No data by design (e.g. opinion, letter, editorial)

FDA. Medical Review for Levothyroxine Sodium (Unithroid). 2000. No data (e.g. description of study only)

FDA. Medical Review for Levothyroxine Sodium (Levoxyl). 2001. No data (e.g. description of study only)

FDA. Medical Review for Levothyroxine Sodium (Levo-T). 2002. No data (e.g. description of study only)

FDA. Medical Review for Levothyroxine Sodium (Novothyrox). 2002. No data (e.g. description of study only)
FDA. Medical Review for Levothyroxine Sodium (Synthroid). 2002. *No data (e.g. description of study only)*

FDA. Medical Review for Levothyroxine Sodium (Thyro-Tabs). 2002. *No data (e.g. description of study only)*

FDA. Medical Review for Levothyroxine Sodium (Levolet). 2003. *No data (e.g. description of study only)*

FDA. Medical Review for Levothyroxine Sodium (Triosint). 2006. *No data (e.g. description of study only)*

FDA. Medical Review for Levothyroxine Sodium (Tirosint). 2007. *No data (e.g. description of study only)*


Flynn RWV, MacDonald TM, Jung RT, Morris AD, Leese GP. Mortality and Vascular Outcomes in Patients Treated for Thyroid Dysfunction. J Clin Endocrinol Metab2006. p. 2159-64. *Wrong population*

Fowler PB. Mortality in elderly patients with subclinical hyperthyroidism. Lancet. 2002/03/13 ed2002. p. 798; author reply 9. *No data by design (e.g. opinion, letter, editorial)*


Freitas MCd, Campolina AG, Ribeiro RL, Kitadai FT. Comparação de função cognitiva e depressão em pacientes hipotiroides subclínicos com eutiroideos e hipotiroides em tratamento, acima de 65 anos;Comparison of cognitive function and depression among patients over 65 years in subclinical hypothyroidism, euthyroidism and hypothyroidism in treatment. Rev Soc Bras Clín Méd2009. *Not related to key questions*


Ghita S, Musat M, Simescu M, Balmes E, Ursu H, Badiu C, et al. Subclinical Hypothyroidism How Soon to Treat? 8th European Congress of Endocrinology (ECE 2006), SECC Congress Centre, Glasgow, Scotland (UK), 1-5 Apr 20062006. Related to topic but does not meet criteria for study design or quality


Gulseren S, Gulseren L, Hekimsoy Z, Cetinay P, Ozen C, Tokatlioglu B. Depression, Anxiety, Health-Related Quality of Life, and Disability in Patients with Overt and Subclinical Thyroid Dysfunction. Archives of Medical Research2006. p. 133-9. Related to topic but does not meet criteria for study design or quality


D-11
Health Canada Apotex Inc. Product Monography APO-Methimazole. 2004. *No data by design (e.g. opinion, letter, editorial)*

Health Canada EMD Serono Canada. Product Monograph Euthyrox. 2008. *No data by design (e.g. opinion, letter, editorial)*


Ho AC, Ho WK, Lam PK, Yuen AP, Wei WI. Thyroid dysfunction in laryngectomees-10 years after treatment. Head Neck. 2007/07/20 ed2008. p. 336-40. *No relevance to key questions*

Hoffmann CJ, Brown TT. Thyroid function abnormalities in HIV-infected patients. Clin Infect Dis. 2007/07/20 ed2007. p. 488-94. *No data by design (e.g. opinion, letter, editorial)*


Ineck BA, Ng TM. Effects of subclinical hypothyroidism and its treatment on serum lipids. Ann Pharmacother. 2003/04/24 ed2003. p. 725-30. **No data by design (e.g. opinion, letter, editorial)**


ISRCTN35570362. The cardiovascular risk in people with SCH, health status and quality of life can be improved by treatment with L-thyroxine. **No data (e.g. description of study only)**

ISRCTN52563354. The effects of a 12-week exercise intervention on respiratory exchange ratio, body composition, thyroid function and lipid metabolism in hypothyroid, obese women. **Wrong population**


Kalantari S, Heidarzadeh A. Thyroxine therapy improves serum lipoproteins and some clinical findings in patients with subclinical hypothyroidism. **Related to topic but does not meet criteria for study design or quality**


Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function tests in patients with stable untreated subclinical hypothyroidism. Thyroid 2008. p. 303-8. **No relevance to key questions**

Karmisholt J, Laurberg P. Serum-TSH and Serum-TPO-Ab Oscillate in Parallel and High Urinary Iodine Excretion (UIE) Predicts Subsequent Thyroid Failure. A One Year Study of Patients with Untreated Subclinical Hypothyroidism (SH). 32nd Annual Meeting of the European Thyroid Association, Carl Ludwig Institute, University Hospital Leipzig, Leipzig (Germany), 1-5 Sep 20072007. p. 69. **No relevance to key questions; Related to topic but does not meet criteria for study design or quality**

Kebapcilar L, Comlekci A, Tuncel P, Solak A, Secil M, Gencel O, et al. The Effects of Subclinical Hypothyroidism and Replacement Therapy on Paraoxonase-1 (PON-1) and Common Carotis Intima Media Thickness. 9th European Congress of Endocrinology (ECE 2007), Budapest (Hungary), 28 Apr-2 May 20072007. **No relevance to key questions; Related to topic but does not meet criteria for study design or quality**


Kim SK, Kim SH, Park KS, Park SW, Cho YW. Regression of the increased common carotid artery-intima media thickness in subclinical hypothyroidism after thyroid hormone replacement. Endocr J 2009. **Related to topic but does not meet criteria for study design or quality**

King Pharmaceuticals. No data (e.g. description of study only)


Koren Peleg R, al. e. The effect of levothyroxine on arterial stiffness and lipid profile in patients with subclinical hypothyroidism. Thyroid 2008. p. 825-30. **Related to topic but does not meet criteria for study design or quality**


Lazarus JH. Aspects of treatment of subclinical hypothyroidism. Thyroid. 2007/05/01 ed2007. p. 313-6. **No data by design (e.g. opinion, letter, editorial)**


Liu Y, Redetzke RA, Said S, Pottala JV, de Escobar GM, Gerdes AM. Serum thyroid hormone levels may not accurately reflect thyroid tissue levels and cardiac function in mild hypothyroidism. Am J Physiol Heart Circ Physiol. 2008/03/04 ed2008. p. H2137-43. **Wrong population**

Lock RJ, Marden NA, Kemp HJ, Thomas PH, Goldie DJ, Gompels MM. Subclinical hypothyroidism: a comparison of strategies to achieve adherence to treatment guidelines. Ann Clin Biochem. 2004/05/01 ed2004. p. 197-200. **No relevance to key questions; Wrong population and No relevance to key questions**

Luboshitzky R, Herer P. Cardiovascular risk factors in middle-aged women with subclinical hypothyroidism. Neuroendocrinology letters2004. p. 262-6. **B/ No relevance to key questions**

Macedo CG, Mamede FBV, Oliveira FvRd, Gomes JD, Moreira LA, Marino MÃnZ, et al. Influência do hipotireoidismo subclínico em pacientes com hipercolesterolemia; Influence of subclinical hypothyroidism in patients with hypercholesterolemia. Arq mÃ©d ABC2004. p. 37-43. **Related to topic but does not meet criteria for study design or quality**


Marina L, Vujovic S, Ivovic M, Tancic-Gajic M, Stojanovic M, Barac M, et al. Surprising: Hyperlipidemia in Subclinical Hyperthyroidism! 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. **No relevance to key questions**

Mariotti S, Cambuli VM. Cardiovascular risk in elderly hypothyroid patients. Thyroid. 2007/12/01 ed2007. p. 1067-73. **No data by design (e.g. opinion, letter, editorial)**


Martinez I, Giralt P. Subclinical Hyperthyroidism, a Retrospective Study for 11 Years. 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. **Related to topic but does not meet criteria for study design or quality**
Martin-Ezquerra G, Sanchez-Regana M, Massana-Gil J, Umbert-Millet P. Papular mucinosis associated with subclinical hypothyroidism: improvement with thyroxine therapy. J Eur Acad Dermatol Venereol. 2006/10/26 ed2006. p. 1340-1. No data by design (e.g. opinion, letter, editorial); No relevance to key questions

McCormack AI, Learoyd DL. A guide to thyroid dysfunction. Medicine Today2006. p. 28-35. No data by design (e.g. opinion, letter, editorial)


Meek S, Smallridge RC. Effect of thyroid hormone replacement on methionine-stimulated homocysteine levels in patients with subclinical hypothyroidism: a randomized, double-blind, placebo-controlled study. Endocr Pract. 2006/09/28 ed2006. p. 529-34. No relevance to key questions


Milionis H, Tambaki A, Kanioglou C, et al. Thyroid substitution therapy induces high density lipoprotein-associated platelet-activating factor acetylhydrolase in patients with sub-clinical hypothyroidism: a potential antiatherogenic effect. Thyroid2005. p. 455-60. Related to topic but does not meet criteria for study design or quality


Monzani F, Dardano A, Caracceo N. Does treating subclinical hypothyroidism improve markers of cardiovascular risk? Treat Endocrinol. 2006/03/18 ed2006. p. 65-81. No relevance to key questions

Moradi S, Bahreinian AM, Azizi F. The effects of L-thyroxine treatment on cognitive and psychiatric aspects of subclinical hypothyroidism: a randomized, double blind clinical trial [Farsi]. Journal of Iran University of Medical Sciences2005. p. 174-. Related to topic but does not meet criteria for study design or quality

Morocco M, Kloss RT. Subclinical hypothyroidism in women: who to treat. Dis Mon. 2003/02/04 ed2002. p. 659-70. No data by design (e.g. opinion, letter, editorial)


NCT00094237. Subclinical Thyroid Dysfunction and Risk of Myocardial Infarction and Stroke. Related to topic but does not meet criteria for study design or quality

NCT00497575. Diagnosis and Follow-up of Patients With Subclinical Hypothyroidism (CHRO1). Wrong population

NCT00498238. The Effect of Levothyroxine on Arterial Stiffness and Lipid Profile in Patients With Subclinical Hypothyroidism. No data (e.g. description of study only)


Pagano L, Pulcrano M, Ippolito S, Rossi A, r rico A, Sidiropulos M, et al. Coronary Flow Reserve (CFR) in Patients with Subclinical Hypothyroidism. 32nd Annual Meeting of the European Thyroid Association, Carl Ludwig Institute, University Hospital Leipzig, Leipzig (Germany), 1-5 Sep 20072007. p. 90. No relevance to key questions


Pearce EN. Hypothyroidism and dyslipidemia: Modern concepts and approaches. Current Cardiology Reports 2004. p. 451-6. No data by design (e.g. opinion, letter, editorial)

Pearce EN. Thyroid dysfunction in perimenopausal and postmenopausal women. Menopause Int. 2007/04/24 ed2007. p. 8-13. No data by design (e.g. opinion, letter, editorial)

Peng L, Gu M-J. [Influence of thyroxine treatment on serum lipid levels in patients with subclinical hypothyroidism: A meta-analysis]. Academic Journal of Second Military Medical University 2007. p. 519-23. Related to topic but does not meet criteria for study design or quality

Peppa M, Hadjidakis D, Alevizaki M, Dimitriadis G, Isari G, Economopoulos T, et al. Insulin Resistance and Oxidative Stress Induce Advanced Glycation End Products Formation in Patients with Clinical and Subclinical Hypothyroidism. 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. No relevance to key questions


Pietras SM, Safer JD. Diagnostic confusion attributable to spurious elevation of both total thyroid hormone and thyroid hormone uptake measurements in the setting of autoantibodies: case report and review of related literature. Endocr Pract. 2008/11/11 ed2008. p. 738-42. Related to topic but does not meet criteria for study design or quality

Pimentel L, Hansen KN. Thyroid disease in the emergency department: A clinical and laboratory review. Journal of Emergency Medicine 2005. p. 201-9. No data by design (e.g. opinion, letter, editorial); No relevance to key questions

Portella RB, Da Costa Silva JL, Wagman MB, De Oliveira FP, Buescu A, Vaisman M. Exercise performance in young and middle-aged female patients with subclinical hyperthyroidism. Thyroid2006. p. 731-5. No relevance to key questions

Prats Julia M. [Effect of treatment with levothyroxine in the lipid profile of the patients with subclinical hypothyroidism]. Endocrinol Nutr2009. p. 13-7. Related to topic but does not meet criteria for study design or quality

Premawardhana LDKE, Lazarus JH. Management of thyroid disorders. Postgraduate Medical Journal2006. p. 552-8. No data by design (e.g. opinion, letter, editorial)


Rao P, Pereira O, Weaver J. The Outcome of the Management of Subclinical Hyperthyroidism: Use of Fixed Dose of Radioiodine Therapy. 2007 BES Conference of the Society for Endocrinology, International Convention Centre, Birmingham (UK), 5-8 Mar 20072007. Related to topic but does not meet criteria for study design or quality

Rasic-Milutinovic Z, Milicevic D, Gluvic Z, Vujovic M, Tica J, Popovic-Radinovic V, et al. Subclinical Hypothyroidism and Bone Mineral Density. 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. No relevance to key questions

Razvi S, McMillan CV, Weaver JU. Quality of Life, Health Status, Symptoms and Treatment Satisfaction in Subclinical Hypothyroidism: A Double&Hyphen;Blind 12-Week Cross&Hyphen;Over Study of L-Thyroxine Versus Placebo. 8th European Congress of Endocrinology (ECE 2006), SECC Congress Centre, Glasgow, Scotland (UK), 1-5 Apr 20062006. Related to topic but does not meet criteria for study design or quality

Razvi S, Pearce S, Weaver J. Does L-Thyroxine Reduce Weight in People with Subclinical Hypothyroidism? 197th Meeting of the Society for Endocrinology, Kensington Town Hall, London (UK), 6-7 Nov 20062006. Related to topic but does not meet criteria for study design or quality

Razvi S, Weaver JU. Subclinical hypothyroidism and cardiovascular risk. Thyroid. 2004/12/14 ed2004. p. 870; author reply 1-2. No data by design (e.g. opinion, letter, editorial)

Razvi S IL, McMillan CV, Weaver JU. Health status in patients with sub-clinical hypothyroidism. Eur J Endocrinol 2005 May. p. 713-7. No relevance to key questions

Razvi Sea. Incident ischemic heart disease and thyroid status: The Whickham cohort reevaluated. Thyroid2007. p. S-72 #89. No data by design (e.g. opinion, letter, editorial) no data

*No relevance to key questions*


*Background - review*


*Background - review*


Saler T, Saglam ZA, Korkmaz K, Ataoglu HE, Yenigun M, Oguzhan B. The Effect of L-Thyroxin Theraphy on Left Ventricular Diastolic Dysfunction in Patients with Subclinical Hypothyroidism. 9th European Congress of Endocrinology (ECE 2007), Budapest (Hungary), 28 Apr-2 May 2007. *No relevance to key questions*

Samuels MH. Cognitive function in untreated hypothyroidism and hyperthyroidism. Curr Opin Endocrinol Diabetes Obes. 2008/09/05 ed2008. p. 429-33. *No relevance to key questions*

Samuels MH, Schuff KG, Carlson NE, Carello P, Janowsky JS. Health status, mood, and cognition in experimentally induced subclinical hypothyroidism. J Clin Endocrinol Metab. 2007/05/03 ed2007. p. 2545-51. *Wrong population*

Santhana Krishnan SG, Pathalapati R, Kaplan L, Cobbs RK. Falsely raised TSH levels due to human anti-mouse antibody interfering with thyrotropin assay. Postgrad Med J. 2006/11/14 ed2006. p. e27. *No relevance to key questions; No data by design (e.g. opinion, letter, editorial)*


Shinno H, Inami Y, Inagaki T, Kawamukai T, Utani E, Nakamura Y, et al. Successful treatment with levothyroxine for idiopathic hypersomnia patients with subclinical hypothyroidism. General Hospital Psychiatry2009. p. 190-3. No data by design (e.g. opinion, letter, editorial)


Surks MI. Primary hypothyroidism: new issues and controversies. Endocrinologist2006. p. 203-7. No data by design (e.g. opinion, letter, editorial)

Surks MI, Goswami G, Daniels GH. The thyrotropin reference range should remain unchanged. Journal of Clinical Endocrinology and Metabolism2005. p. 5489-96. No data by design (e.g. opinion, letter, editorial)

Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. Jama. 2004/01/15 ed2004. p. 228-38. **Related to topic but does not meet criteria for study design or quality**


Teixeira P, Cabral MD, Silva NA, Soares DV, Braulio V, Couto AP, et al. Serum leptin in overt and subclinical hypothyroidism: effect of levothyroxine treatment and relationship to menopausal status and body composition. Thyroid 2009. p. 443-50. **Related to topic but does not meet criteria for study design or quality**


Teixeira PF, Reuters VS, Ferreira MM, Almeida CP, Reis FA, Melo BA, et al. Treatment of subclinical hypothyroidism reduces atherogenic lipid levels in a placebo-controlled double-blind clinical trial. Horm Metab Res. 2007/12/19 ed2008. p. 50-5. **Wrong population**

Trokoudes KM, Skordis N, Picolos MK. Infertility and thyroid disorders. Curr Opin Obstet Gynecol. 2006/06/24 ed2006. p. 446-51. **No data by design (e.g. opinion, letter, editorial)**


Vaisman M, Teixeira P, Martins J, Soares D, Braulio V, Buescu A, et al. Effects of Levothyroxine Reposition on Insulin Sensitivity, Leptin Levels and Fat Mass in Patients with Subclinical Hypothyroidism. 32nd Annual Meeting of the European Thyroid Association, Carl Ludwig Institute, University Hospital Leipzig, Leipzig (Germany), 1-5 Sep 20072007. p. 75. **Related to topic but does not meet criteria for study design or quality**


D-22

van Harten AC, Leue C, Verhey FR. [Should depressive symptoms in patients with subclinical hypothyroidism be treated with thyroid hormone?]. Tijdschr Psychiatr. 2008/08/09 ed2008. p. 539-43. **Related to topic but does not meet criteria for study design or quality**

Velija-Asimi Z, Heljic B. Changes in Cystatin C during the Treatment of Subclinical Hypothyroidism. 10th European Congress of Endocrinology (ECE 2008), Berlin (Germany), 3-7 May 20082008. **No relevance to key questions**


Villar H, Gharib H. Review: Available evidence does not support a benefit for thyroid hormone replacement in adults with subclinical hypothyroidism: Commentary. Evidence-Based Medicine2008. p. 22%. N 1. **No data by design (e.g. opinion, letter, editorial**

Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. Cochrane Database Syst Rev. 2007/07/20 ed2007. p. CD003419. **Related to topic but does not meet criteria for study design or quality**

Volzke H, Dorr M, Wallaschofski H. Is Subclinical Hyperthyroidism Associated with Blood Pressure and Hypertension? 8th European Congress of Endocrinology (ECE 2006), SECC Congress Centre, Glasgow, Scotland (UK), 1-5 Apr 20062006. **No relevance to key questions**

Volzke H, Schwahn C, Wallaschofski H, Dorr M. Review: The association of thyroid dysfunction with all-cause and circulatory mortality: is there a causal relationship? J Clin Endocrinol Metab. 2007/05/03 ed2007. p. 2421-9. **No data by design (e.g. opinion, letter, editorial**


Wartofsky L, Dickey RA. The evidence for a narrower thyrotropin reference range Is compelling. J Clin Endocrinol Metab2005. p. 5483-8. **No data by design (e.g. opinion, letter, editorial**

Wartofsky L, Van Nostrand D, Burman KD. Overt and 'subclinical' hypothyroidism in women. Obstet Gynecol Surv. 2006/07/18 ed2006. p. 535-42. **No data by design (e.g. opinion, letter, editorial**

Weetman AP. Radioiodine treatment for benign thyroid diseases. Clin Endocrinol (Oxf). 2007/05/01 ed2007. p. 757-64. **No relevance to key questions**

Weissel M. Disturbances of thyroid function in the elderly. Wien Klin Wochenschr. 2006/02/21 ed2006. p. 16-20. **No data by design (e.g. opinion, letter, editorial)**


Woeber K. Observations concerning the natural history of subclinical hyperthyroidism. Thyroid 2005. p. 687. **No relevance to key questions**

Yandell SD, Harvey WC, Fernandes NJ, Barr PW, Feldman M. Radioiodine studies, low serum thyrotropin, and the influence of statin drugs. Thyroid. 2008/09/20 ed2008. p. 1039-42. **No relevance to key questions**


## Appendix E. Screening and Case-Finding Recommendations of Other Groups

<table>
<thead>
<tr>
<th>Organization</th>
<th>Guideline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ATA 2000 [Ladenson 2000]</strong></td>
<td>Adults should be screened for thyroid dysfunction by measurement of the serum thyrotropin concentration, beginning at age 35 years and every 5 years thereafter.</td>
</tr>
<tr>
<td><strong>IOM, 2003</strong></td>
<td>There is insufficient evidence to recommend periodic, routine screening for thyroid dysfunction among asymptomatic persons using serum TSH levels.</td>
</tr>
<tr>
<td><strong>AACE, ATA, Endocrine Society, 2004 [Surks 2004]</strong></td>
<td>The panel recommends against population-based screening for thyroid disease. Case ascertainment in certain high risk groups is encouraged. The panel finds the evidence insufficient to recommend for or against routine determination of TSH levels (screening) in pregnant women or women planning to become pregnant. It is reasonable to consider serum TSH measurement for women with a family history of thyroid disease, prior thyroid dysfunction symptoms or physical findings suggestive of hypothyroidism or hyperthyroidism, an abnormal thyroid gland on examination, type 1 diabetes mellitus, or a personal history of an autoimmune disorder.</td>
</tr>
<tr>
<td><strong>AACE, ATA, Endocrine Society, 2004 [Gharib 2005]</strong></td>
<td>Consensus statement favors routine screening for subclinical thyroid dysfunction in adults, including pregnant women and those contemplating pregnancy.</td>
</tr>
</tbody>
</table>
| **British Thyroid Association 2006** | - If screening is performed, and a high serum TSH concentration and normal FT4 is found, repeat measurement 3-6 months later after excluding non-thyroidal illness and drug interference.  
- If the TSH is mildly elevated (above the reference range but below 10 mU/L), obtain serum thyroid peroxidase antibodies.  
- If antibody levels are high, repeat measurement of TSH annually. If they are low, repeat measurement of TSH every 3 years. Initiate treatment if the TSH level is greater than 10 mU/L or the patient develops clinical findings of hypothyroidism. |
| **ACOG, 2007[ACOG Committee Opinion No. 381: Subclinical Hypothyroidism in Pregnancy]** | Based on current literature, thyroid testing in pregnancy should be performed on symptomatic women and those with a personal history of thyroid disease or other medical conditions associated with thyroid disease (eg, diabetes mellitus). Without evidence that identification and treatment of pregnant women with subclinical hypothyroidism improves maternal or infant outcomes, routine screening for subclinical hypothyroidism currently is not recommended. |
| **AAFP [AAFP website accessed 1/4/2010]** | The AAFP concludes that the evidence is insufficient to recommend for or against routine screening for thyroid disease in adults. |

**Abbreviations:** AACE=American Association of Clinical Endocrinologists; AAFP=American Academy of Family Physicians; ACOG=American Congress of Obstetricians and Gynecologists; ACP=American College of Physicians; ATA=American Thyroid Association; IOM=Institute of Medicine; USPSTF=United States Preventive Services Task Force.
Appendix F. Results of Search for Systematic Reviews


Dissemination CfRa. Radioiodine therapy (RAI) for Graves' disease (GD) and the effect on ophthalmopathy: a systematic review (Structured abstract). Database of Abstracts of Reviews of Effects2010.

Duley L, Gulmezoglu AM, Henderson-Smart DJ. Magnesium sulphate and other anticonvulsants for women with pre-eclampsia. Cochrane Database of Systematic Reviews2009.


Ma C, Kuang A, Xie J. Radioiodine therapy for differentiated thyroid carcinoma with thyroglobulin positive and radioactive iodine negative metastases. Cochrane Database of Systematic Reviews2009.


Ng SM, Anand D, Weindling AM. High versus low dose of initial thyroid hormone replacement for congenital hypothyroidism. Cochrane Database of Systematic Reviews 2009.


Osborn DA. Thyroid hormones for preventing neurodevelopmental impairment in preterm infants. Cochrane Database of Systematic Reviews 2009.


Shah VS, Ohlsson A. Venepuncture versus heel lance for blood sampling in term neonates. Cochrane Database of Systematic Reviews2009.


Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. JAMA. 2004/01/15 ed2004. p. 228-38.


Wiersinga WM. [The practice guideline 'Thyroid disorders' (first revision) from the Dutch College of General Practitioners; a response from the perspective of internal medicine]. Ned Tijdschr Geneeskd. 2008/02/02 ed2007. p. 2822-4.


