JAMA | US Preventive Services Task Force | EVIDENCE REPORT

Screening for Thyroid Cancer
Updated Evidence Report and Systematic Review for the US Preventive Services Task Force

Jennifer S. Lin, MD, MCR; Erin J. Aiello Bowles, MPH; Selvi B. Williams, MD, MPH; Caitlin C. Morrison, MPH

IMPORTANCE The incidence of detected thyroid cancer cases has been increasing in the United States since 1975. The majority of thyroid cancers are differentiated cancers with excellent prognosis and long-term survival.

OBJECTIVE To systematically review the benefits and harms associated with thyroid cancer screening and treatment of early thyroid cancer in asymptomatic adults to inform the US Preventive Services Task Force.

DATA SOURCES Searches of MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials for relevant studies published from January 1966 through January 2016, with active surveillance through December 2016.

STUDY SELECTION English-language studies conducted in asymptomatic adult populations.

DATA EXTRACTION AND SYNTHESIS Two reviewers independently appraised the articles and extracted relevant study data from fair- or good-quality studies. Random-effects meta-analyses were conducted to pool surgical harms.

MAIN OUTCOMES AND MEASURES Thyroid cancer morbidity and mortality, test accuracy to detect thyroid nodules or thyroid cancer, and harms resulting from screening (including overdiagnosis) or treatment of thyroid cancer.

RESULTS Of 10 424 abstracts, 707 full-text articles were reviewed, and 67 studies were included for this review. No fair- to good-quality studies directly examined the benefit of thyroid cancer screening. In 2 studies (n = 354), neck palpation was not sensitive to detect thyroid nodules. In 2 methodologically limited studies (n = 243), a combination of selected high-risk sonographic features was specific for thyroid malignancy. Three studies (n = 5894) directly addressed the harms of thyroid cancer screening, none of which suggested any serious harms from screening or ultrasound-guided fine-needle aspiration. No screening studies directly examined the risk of overdiagnosis. Two observational studies (n = 39 211) included cohorts of persons treated for well-differentiated thyroid cancer and persons with no surgery or surveillance; however, these studies did not adjust for confounders and therefore were not designed to determine if earlier or immediate treatment vs delayed or no surgical treatment improves patient outcomes. Based on 36 studies (n = 43 295), the 95% CI for the rate of surgical harm was 2.12 to 5.93 cases of permanent hypoparathyroidism per 100 thyroidectomies and 0.99 to 2.13 cases of recurrent laryngeal nerve palsy per 100 operations. Based on 16 studies (n = 291 796), treatment of differentiated thyroid cancer with radioactive iodine is associated with a small increase in risk of second primary malignancies and with increased risk of permanent adverse effects on the salivary gland, such as dry mouth.

CONCLUSIONS AND RELEVANCE Although ultrasonography of the neck using high-risk sonographic characteristics plus follow-up cytology from fine-needle aspiration can identify thyroid cancers, it is unclear if population-based or targeted screening can decrease mortality rates or improve important patient health outcomes. Screening that results in the identification of indolent thyroid cancers, and treatment of these overdiagnosed cancers, may increase the risk of patient harms.

JAMA. 2017;317(18):1888-1903. doi:10.1001/jama.2017.0562

Editorial page 1840
Related article page 1882
Supplemental content
Related articles at
jamainternalmedicine.com
jamaotolaryngology.com
jamasurgery.com

Author Affiliations: Kaiser Permanente Center for Health Research, Kaiser Permanente Research Affiliates Evidence-based Practice Center, Portland, Oregon (Lin, Williams); Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Research Affiliates Evidence-based Practice Center, Seattle, Washington (Bowles, Morrison).

Corresponding Author: Jennifer S. Lin, MD, MCR, Kaiser Permanente Center for Health Research, 3800 N Interstate Ave, Portland, OR 97227-1098 (jennifer.s.lin@kpchr.org).

The incidence of detected thyroid cancer cases has been rising in the United States for both men and women, from 4.9 cases per 100 000 persons in 1975 to 14.3 cases per 100 000 persons in 2014.[1] However, mortality rates have remained stable at about 0.5 per 100 000 persons per year.[2] Differentiated thyroid cancer generally has a very good prognosis and accounts for about 90% of all cases of thyroid cancer.[3] Within this category, papillary thyroid cancer accounts for about 70% to 80% of thyroid cancer cases, and follicular cancer accounts for 10% to 15%. The 10-year overall survival rates for papillary and follicular thyroid cancer are 93% and 85%, respectively, for all stages of the disease.[4]

Screening for thyroid cancer can be performed with neck palpation, ultrasonography, or both. Screening may have the potential for early detection of malignant thyroid nodules that could make treatment more effective, with less harm, than if administered later. However, screening also may result in overdiagnosis (identification of a thyroid malignancy that likely would not have caused symptoms or death during a patient's lifetime), because it can detect very small or indolent tumors that might never affect a person's morbidity or mortality.[5,6]

No professional medical society recommends population-based screening for thyroid cancer. South Korea appears to be the only country that regularly practices screening for asymptomatic thyroid cancer using ultrasound; this practice arose opportunistically as an add-on option for persons undergoing sanctioned screening through an organized cancer screening program initiated in 1999.[7] This article reports the findings from a systematic review conducted to assist the US Preventive Services Task Force (USPSTF) in its process of updating its 1996 "D" recommendation (screening asymptomatic adults or children for thyroid cancer by neck palpation or ultrasound is not recommended).

Methods

Scope of Review

This review addressed 5 key questions (KQs) as shown in **Figure 1**. Additional methodological details regarding search strategies, detailed study inclusion criteria, quality assessment, excluded studies, and description of data analyses, as well as detailed results, are publicly available in the full evidence report available at https://www.uspreventiveservicestaskforce.org/Page/Document/final-evidence-review159/thyroid-cancer-screening1.

Data Sources and Searches

MEDLINE, PubMed, and the Cochrane Central Register of Controlled Trials were searched to locate primary studies that informed the KQs and that were published from January 1966 through January 2016 (eMethods in the Supplement). The database searches were supplemented with expert suggestions and by reviewing reference lists from existing relevant systematic reviews. ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform were searched for ongoing trials. Since January 2016, we continued to conduct ongoing surveillance through article alerts and targeted searches of high-impact journals to identify major studies published in the interim that may affect the conclusions or understanding of the evidence and therefore the related USPSTF recommendation. The last surveillance was conducted in December 2016. No studies were identified that would substantively change this review's interpretation of findings or conclusions.

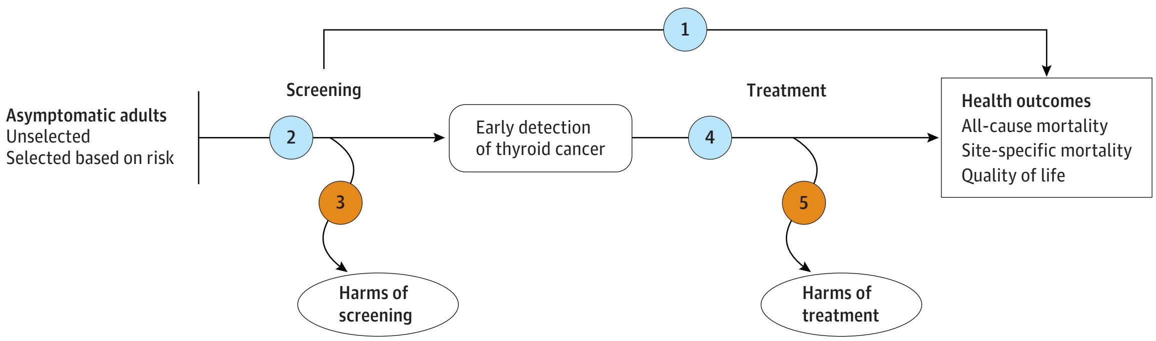
Study Selection

Two investigators independently reviewed titles, abstracts, and full-text articles against the specified inclusion criteria for studies of thyroid cancer screening, diagnostic accuracy, or treatment in screen-relevant or asymptomatic adults. Discrepancies were resolved through consensus and consultation with a third investigator.

For screening questions (KQ1 through KQ3), any studies of asymptomatic adult populations were included, either those at general risk (eg, unselected) or those with prior personal history of radiation exposure. Populations were excluded if they were selected based on high radiation exposure due to environmental disasters, inherited genetic syndromes associated with a high risk for developing thyroid cancer, or a personal history of thyroid cancer. Diagnostic accuracy studies of palpation or ultrasound had to include a reference standard (ultrasound for detection of nodules on palpation; histopathology results from fine-needle aspiration or surgery for detection of cancer on ultrasound), applied to both screen-positive and screen-negative persons (eg, all or a random subset of screen-negative persons). For screening effectiveness (KQ1), any patient health outcome of reduced morbidity or mortality associated with thyroid cancer was included. For test performance (KQ2), cancer detection rates and measures of diagnostic accuracy (eg, sensitivity, specificity, positive and negative predictive values) were included. For harms of screening (KQ3), direct harms of palpation and ultrasound, subsequent harms of diagnostic fine-needle aspiration, and measures of overdiagnosis were included. For overdiagnosis, studies that compared screened vs unscreened groups were sought. Studies that examined the increasing incidence of thyroid cancers, studies of the incidence and natural history of thyroid nodules and cancers, and autopsy studies were not included but are summarized in the Discussion section.

For treatment questions (KQ4 and KQ5), any studies of thyroid surgery (complete thyroidectomy, near-total thyroidectomy, lobectomy), with or without lymph node dissection or with or without radioactive iodine ablation, were included. Studies of chemotherapy, external beam radiation, and other nonsurgical ablative treatment other than radioactive iodine were excluded. To approximate the treatment of screen-detected cancers, treatment studies including persons with metastatic disease or anaplastic thyroid cancers were excluded. For treatment benefit (KQ4), studies had to have a control group (eg, untreated, surveillance, delayed treatment). To assess the benefit of treatment, the patient health outcomes of recurrence, mortality, and quality of life were considered. For treatment harms (KQ5), studies were not required to include a control group for direct procedural harms (eg, hypoparathyroidism, recurrent laryngeal nerve palsy) but needed a control group for other types of harms (eg, second primary malignancies from radioactive iodine therapy). The evolution of standard of care for the diagnostic workup (eg, use of ultrasound-guided fine-needle aspiration) and treatment of thyroid cancer over time has resulted in a change in the case mix of patients getting surgery with or without lymph node dissection or radioactive iodine therapy, as well as improvements in surgical techniques and radioactive iodine administered activity (doses) over time. To identify the most applicable

Figure 1. Analytic Framework and Key Questions



Key questions

1 Compared with not screening, does screening adults for thyroid cancer lead to a reduced risk of thyroid-specific mortality or morbidity, reduced all-cause mortality, and/or improved quality of life?

2 What are the test performance characteristics of screening tests for detecting malignant thyroid nodules in adults?

3 What are the harms of screening adults for thyroid cancer?

4 Does treatment of screen-detected thyroid cancer reduce thyroid-specific mortality or morbidity, reduce all-cause mortality, and/or improve quality of life?

5 What are the harms of treating screen-detected thyroid cancer?

Evidence reviews for the US Preventive Services Task Force (USPSTF) use an analytic framework to visually display the key questions that the review will address to allow the USPSTF to evaluate the effectiveness and safety of a preventive service. The questions are depicted by linkages that relate interventions and outcomes. Further details are available in the USPSTF procedure manual.[8]

evidence, studies conducted before 1990 and single-surgeon case series were excluded.

Data Extraction and Quality Assessment

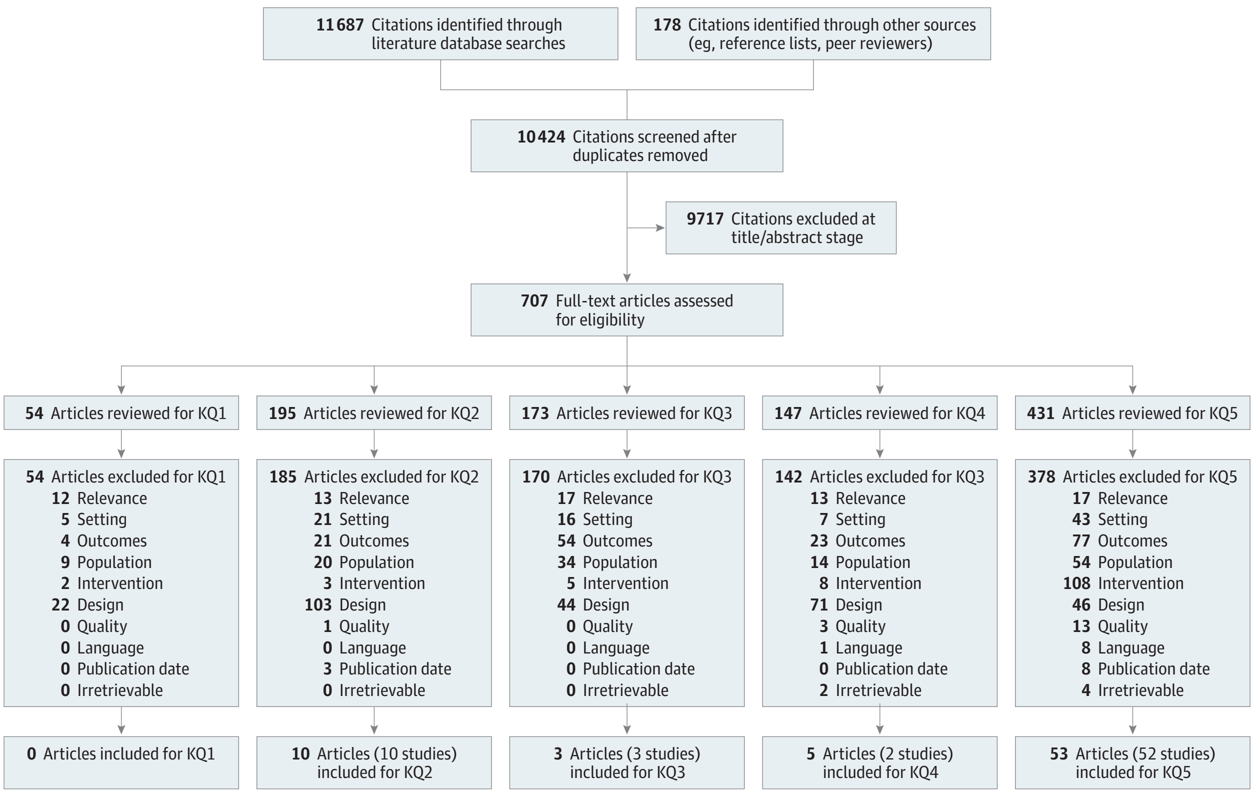
Two reviewers independently critically appraised all articles that met inclusion criteria using the USPSTF design-specific quality criteria[9] supplemented by the Newcastle Ottawa Scales for cohort and case-control studies[10] and by QUADAS (Quality Assessment of Diagnostic Accuracy Studies) and QUADAS II for studies of diagnostic accuracy[11,12] (eTable 1 in the Supplement). Poor-quality studies (those with a single fatal flaw or multiple important limitations that could invalidate results) were excluded from this review. Disagreements about critical appraisal were resolved by consensus and, if needed, consultation with a third independent reviewer. One reviewer extracted key data from included studies; a second reviewer checked the data for accuracy. Tables generally included details on study design and quality, setting and population (eg, country, inclusion criteria, age, sex, race/ethnicity, risk factors for thyroid cancer), screening and treatment details, reference standard or comparator details (if applicable), length of follow-up, and outcomes (eg, cancer yield, diagnostic accuracy, cancer morbidity, mortality, and harms).

Data Synthesis and Analysis

For each KQ, the number and design of included studies, summary of results, consistency and precision of results, reporting bias, summary of study quality, limitations of the body of evidence, and applicability of the findings were summarized. Findings were synthesized by KQ, screening test (eg, palpation, ultrasound) or treatment (eg, type of surgery, radioactive iodine therapy), and type of outcome. Because of the limited number of studies and the clinical heterogeneity of studies, the analyses were largely descriptive.

Random-effects meta-analyses were conducted using the restricted maximum likelihood estimation method to estimate the harms of surgical treatment of thyroid cancer (permanent hypoparathyroidism and permanent recurrent laryngeal nerve palsy). In subgroup analysis when the number of studies was less than 5, a fixed-effects model was used. The presence and magnitude of statistical heterogeneity were assessed among pooled studies using the I^2 statistic. Visual inspection of plots stratified or ordered by key study characteristics accounting for clinical heterogeneity among studies was conducted to see if these characteristics affected rates of surgical complications. Key study characteristics included the type of surgery (eg, partial or total thyroidectomy with or without lymph node dissection; type of lymph node dissection), case mix of patients (eg, histology of thyroid cancer, average tumor size, average age), setting (eg, country, year), and type and definition of outcome (eg, criteria for permanent harm). It was not possible to evaluate associations of surgical complications with study quality (because all studies were fair quality) or surgical experience (because experience and surgical volume were not reported in individual studies). Funnel plots and the Egger linear regression

Figure 2. Literature Search Flow Diagram



KQ indicates key question.

method were used to examine whether the distribution of the effect sizes was symmetric with respect to effect precision.

Significance threshold was 2-sided *P* = .05. All analyses were performed using R version 3.2.2 (R Project for Statistical Computing).

Results

A total of 10 424 unique abstracts and 707 full-text articles were reviewed (**Figure 2**). Of these, 67 unique studies were included: 10 studies of screening test performance (n = 203 718), 3 studies of screening harms (n = 5894), 2 studies of treatment benefits (n = 39 211), and 52 studies of treatment harms (n = 335 091).

Screening Effectiveness or Accuracy

Key Question 1. Compared with not screening, does screening adults for thyroid cancer lead to a reduced risk of thyroid-specific morbidity or mortality, reduced all-cause mortality, and/or improved quality of life?

No studies met the inclusion criteria for KQ1. No randomized clinical trials or controlled clinical trials evaluated the effect of thyroid cancer screening on patient morbidity or mortality compared with no screening. Two cohort studies that compared screened individuals vs a comparator group did not meet inclusion criteria for KQ1.[13,14]

Key Question 2. What are the test performance characteristics of screening tests for detecting malignant thyroid nodules in adults?

Ten fair-quality studies (n = 203 718) met the inclusion criteria for KQ2 (**Table 1**). Only 2 studies (n = 354) reported on diagnostic accuracy of palpation to detect nodules[16,17] and 2 (n = 243) on diagnostic accuracy of ultrasound to detect cancer.[18,19] The majority of studies that examined the diagnostic accuracy of ultrasound to detect thyroid cancer were not (or were not reported to be) conducted in screening populations and were excluded. Therefore, evidence to inform the true diagnostic accuracy of screening using neck palpation or ultrasound to detect thyroid cancer is limited. Among the included studies, 4 reported on cancer yield from screening for thyroid cancer using palpation plus follow-up ultrasound,[14-17] another 4 on cancer yield from screening using ultrasound only,[18-21] and 2 from the 1980s on cancer yield from screening of adults with a history of childhood irradiation.[22,23] Cancer yield results are not discussed in this manuscript but are included in the full evidence review.

Two studies (n = 354) conducted by the same investigator, evaluating a single examiner in Finland in the late 1980s, found that neck palpation was not sensitive to detect thyroid nodules in adults.[16,17] Only one of these studies reported the diagnostic accuracy of palpation for all screened persons.[16] In that study of randomly selected adults (n = 253), an abnormal result from neck palpation (thyroid nodule or diffuse enlargement of the thyroid) was found in 5.1% of participants, whereas an abnormal result from ultrasound was found in 27.3%. The sensitivity and specificity of palpation to detect thyroid nodules (size not reported) were 11.6% (95% CI, 5.1%-21.6%) and 97.3% (95% CI, 93.8%-99.1%), respectively.[16]

Table 1. Included Studies for Key Question 2—Test Performance Characteristics of Screening Tests for Detecting Malignant Thyroid Nodules in Adults

Source[a]	Screening Method	Country (Recruitment Years)	Study Design	Mean Age, y	Screened		Reported	
					Total No.	Women, No. (%)	Diagnostic Accuracy	Cancer Yield
Average-Risk Population								
Suehiro,[15] 2006	Palpation	Japan (1989-2005)	Retrospective	49	46 433	20 895 (45[b])		✓
Brander et al,[16] 1991	Palpation	Finland (1989-1990)	Prospective	35	253	129 (51)	✓	✓
Brander et al,[17] 1989	Palpation	Finland (1988)	Prospective	52	101	101 (100)	✓	✓
Ishida et al,[14] 1988	Palpation	Japan (1980-1986)	Prospective	NR	152 651	152 651 (100)		✓
Kim et al,[18] 2010	Ultrasonography	South Korea (2005-2007)	Prospective	43	2079	43 (100)	✓	
Kim et al,[19] 2008	Ultrasonography	South Korea (2004-2006)	Retrospective	53	16 352	10 956 (67)	✓	
Lee et al,[20] 2003	Ultrasonography	South Korea (2003)	Prospective	43	697	697 (100)		✓
Chung et al,[21] 2001	Ultrasonography	South Korea (1997-1998)	Prospective	47	1401	1401 (100)		✓
High-Risk Population								
Ron et al,[22] 1984	Palpation + diagnostic follow-up[c]	Israel (NR)	Prospective	29	443	217 (49)		✓
Shimaoka et al,[23] 1982	Palpation + diagnostic follow-up[d]	United States (1977-1980)	Prospective	39	1500	960 (64)		✓

Abbreviation: NR, not reported.

[a] All studies were of fair quality.

[b] Percentage of examination visits that were women (calculated).

[c] Diagnostic follow-up consisted of technitium-99m thyroid scan and thyroid function tests.

[d] Diagnostic follow-up consisted of iodine 123 thyroid scan, blood tests, and indirect laryngoscopy.

Table 2. Key Question 2 Results—Diagnostic Accuracy of Screening Ultrasonography for Thyroid Cancer

Source[a]	Analytic Sample, No.	Characteristic	% (95% CI)	
			Sensitivity	Specificity
Kim et al,[18] 2010	113 persons	≥1 of following characteristics	94.3	55.0
		Microcalcification	34.0 (21.5-48.3)[b]	83.3 (71.5-91.7)[b]
		Irregular shape	88.7 (77.0-95.7)[b]	63.3 (49.9-75.4)[b]
		Taller-than-wide shape	67.9 (53.7-80.1)[b]	80.0 (67.7-89.2)[b]
		Ill-defined or microlobulated margin	86.8 (74.7-94.5)[b]	68.3 (55.0-79.7)[b]
		Marked hypoechogenicity	52.8 (38.6-66.7)[b]	86.7 (75.4-94.1)[b]
Kim et al,[19] 2008	140 nodules[c]	≥2 of following characteristics	94.8	86.6
		Microcalcification	70.7 (57.3-81.9)	98.8 (93.4-99.8)
		Taller-than-wide or irregular shape	55.2 (41.5-68.3)	89.0 (80.2-94.8)
		Spiculated margin	48.3 (35.0-61.8)	97.6 (91.4-99.6)
		Marked hypoechogenicity	55.2 (41.5-68.3)	96.3 (89.7-99.2)
		Solid	93.1 (83.3-98.0)	51.2 (39.9-62.4)

[a] Both studies were of fair quality. Both studies report accuracy only among patients who had at least 1 study-defined malignant ultrasound characteristic, providing no follow-up on the majority (n = 18 188) of screened individuals who did not have these characteristics.

[b] Calculated confidence intervals.

[c] Reported in 130 persons.

In the other study of women presenting for screening mammography (n = 101), palpation results were reported for the 36 patients with abnormal ultrasound examination results; the sensitivity of palpation to detect nodules in persons with an abnormal ultrasound result was 27.8%.[17]

In 2 methodologically limited studies conducted in South Korea (n = 243), screening with ultrasound was very sensitive to detect thyroid malignancy and can be specific for thyroid malignancy when using selected high-risk sonographic features (Table 2).[18,19] Both studies were conducted by the same investigators from 2004 to 2007 but had different study designs. The better-quality study prospectively examined the diagnostic accuracy of screening for thyroid cancer by ultrasound in 113 women referred for fine-needle aspiration (among 2079 screened). Seventy-seven of the analyzed women had 1 or more high-risk sonographic characteristics (presence of microcalcifications, irregular shape, ill-defined or microlobulated margin, marked hypoechogenicity, taller-than-wide orientation), and 36 had probable benign ultrasound findings but were referred for fine-needle aspiration by the radiologist or by request of their outpatient clinician.[18] Among these 113 women, 53 were diagnosed with papillary thyroid cancer. The sensitivity and specificity of having 1 or more malignant features on screening ultrasound were 94.3% (95% CI, 84.3%-98.8%) and 55.0% (95% CI, 41.6%-67.9%), respectively.

Table 3. Included Studies and Results for Key Question 3—Harms of Screening for Thyroid Cancer and Diagnostic Fine-Needle Aspiration

Source[a]	Country (Recruitment Years)	No. of Women/Total (%)	Mean Age, y	Study Aim	Outcome
Hobbs et al,[24] 2014	United States (2010-2011)	332/400 (83)	55	To determine the proportion of thyroid nodules undergoing ultrasound-guided fine-needle aspiration that do not meet Society of Radiologists in Ultrasound recommendations from 2005[b]	Persons undergoing fine-needle aspiration not meeting Society of Radiologists in Ultrasound recommendations: 96/400 (24.0%)
Abu-Yousef et al,[25] 2011	United States (2006-2007)	413/582[c] (71)	56	To determine whether there is a significantly increased incidence of bleeding complications from ultrasound-guided fine-needle aspiration of neck masses in patients receiving antithrombotic or anticoagulant therapy (compared with patients not receiving therapy)	Major complications (hospitalization or intervention required): 0/582 (0%) Postprocedural hematoma: 5/582 (0.9%)[d]
Ito et al,[26] 2005	Japan (1990-2002)	NR/4912 (NR)[e]	NR[e]	To investigate the relationship between needle tract implantation of papillary thyroid cancer and clinicopathological characteristics	Tumor implantation: 7/4912 (0.14%)

Abbreviation: NR, not reported.

[a] All studies were of fair quality and retrospective design.

[b] Fine-needle aspiration is appropriate for nodules that have a maximum diameter of 1 cm or larger and have microcalcifications; nodules that are 1.5 cm or larger and are solid or have coarse calcifications; nodules that are 2 cm or larger and are mixed solid and cystic; and nodules with substantial growth since the prior ultrasound.

[c] For thyroid masses only.

[d] Difference in incidence of hematomas between persons who were receiving antiplatelet or anticoagulant therapy vs persons not receiving therapy not statistically significant.

[e] Data reported for 10 persons with outcomes: mean age 65 years and 90% women.

The other study was a retrospective analysis of 130 asymptomatic persons selected from 1009 persons who underwent fine-needle aspiration based on ultrasound findings (from 16 352 persons who referred themselves to thyroid cancer screening).[19] The study sample included 58 of 150 lesions (38.7%) classified by fine-needle aspiration results as malignant (all papillary thyroid cancer) and 82 of 823 (10.0%) classified as benign, for a total of 140 nodules in 130 persons. Among these 140 nodules, the sensitivity and specificity of having 2 or more high-risk sonographic characteristics (presence of microcalcifications, spiculated margin, marked hypoechogenicity, taller-than-wide orientation or irregular shape, solid) were 94.8% and 86.6%, respectively (95% CI values could not be calculated). The studies did not follow up on the majority (n = 18 188) of screened individuals who were not referred for fine-needle aspiration; therefore, the potential false-negative cases are unknown, and estimates of sensitivity are likely overestimated.

Harms of Screening

Key Question 3. What are the harms of screening adults for thyroid cancer?

Three studies (n = 5894) met inclusion criteria for KQ3 (**Table 3**). No studies examined the harms of thyroid cancer screening with palpation or ultrasound, and no studies directly examined the effect of overdiagnosis in a screened vs unscreened group. A number of other study designs may indirectly inform the clinical importance and magnitude of overdiagnosis in thyroid cancer screening; these studies are summarized in the Discussion section. Overall, there is limited evidence to evaluate the potential harms of screening for thyroid cancer, including harms of diagnostic follow-up fine-needle aspiration. One US study (n = 400) found that 24.0% of persons who had undergone fine-needle aspiration of a thyroid nodule did not meet the Society of Radiologists in Ultrasound recommendation for fine-needle aspiration.[24] Two fair-quality retrospective studies (n = 5494) evaluated the harms of fine-needle aspiration of thyroid nodules, including hospitalization, postprocedural hematoma, and needle tract implantation.[25,26] These studies did not suggest serious harms to patients from ultrasound-guided fine-needle aspiration.

Benefits of Treatment

Key Question 4. Does treatment of screen-detected thyroid cancer reduce thyroid-specific mortality or morbidity, reduce all-cause mortality, and/or improve quality of life?

Two unique observational studies (n = 39 211) reported in 5 articles[27-31] met inclusion criteria for KQ4 (**Table 4**). No trials were designed to evaluate if earlier treatment or treatment of screen-detected, well-differentiated thyroid cancer results in better patient outcomes compared with observation (ie, delayed or no treatment). Because of major limitations in the designs of included studies (eg, lack of adjustment for confounders), it is uncertain if earlier or immediate treatment vs delayed or no surgical treatment improves patient outcomes for papillary carcinoma or papillary microcarcinoma. One retrospective observational study using US Surveillance, Epidemiology, and End Results (SEER) data from 1973 to 2005 compared survival rates of persons treated (surgery with or without radioactive iodine therapy) or not treated for papillary thyroid cancer.[27] A total of 35 663 persons were analyzed; only 440 (1.2%) had not been treated. Overall, untreated persons had a slightly worse 20-year survival rate compared with treated persons (97% [95% CI, 96%-100%] vs 99% [95% CI, 93%-100%], $P < .001$). One prospective study conducted from 1993 to 2013 in Japan examined the recurrence of disease and the survival rate for persons with papillary microcarcinoma who opted for immediate surgery vs those who opted for observation or active surveillance.[28-31] From 1993 to 2004, 1395 persons were analyzed, 340 of whom opted for observation with surveillance ultrasound.[28] Thirty-two percent (n = 109) who opted for observation ultimately had surgery.

Table 4. Included Studies and Results for KQ4—Treatment Effectiveness of Screen-Detected Thyroid Cancer on Patient Health Outcomes

Source[a]	Country (Recruitment Years)	Study Design	Histology	Group[b]	No. of Women/Total (%)	Mean Age, y	Follow-up, Mean (Range), y	No. of Thyroid Cancer-Specific Deaths/Total (%)	Thyroid Cancer-Specific Survival, % (95% CI)
Davies and Welch,[27] 2010	United States (1973-2005)	Retrospective observational	Papillary	Intervention	27 122/35 223 (77)	46	7.6 (0-32)	161/35 223 (0.45)	**20-y survival:** 99 (93-100) **10-y survival:** Treatment recommended: 99.5 (99.4-99.6)[c]
				Comparator	356/440 (81)	51	5.9 (0-31)	6/440 (1.4) Treatment recommended: 4/216 (1.9)[c] Treatment not recommended: 1/165 (0.6)[c] (*P* = .10 for recommended vs not recommended)	**20-y survival:** 97 (96-100) **10-y survival:** Treatment recommended: 98.1 (95.9-100)[c] Treatment not recommended: 99.3 (97.8-100)[c]
				Intervention vs comparator	*P* = .06	*P* < .001	*P* < .001	*P* = .09	**20-y survival:** *P* < .001
Oda et al,[31] 2016 Ito et al,[30] 2014[d]	Japan (1993-2004)	Prospective observational	Papillary microcarcinomas	Intervention	960/1055 (91)	52	6.3 (0.08-15.3)	2/1055 (0.2)	NR
				Comparator	313/340 (92)	NR	6.2 (1.5-15.6)	0/340 (0)	NR
Ito et al,[28] 2010 Ito et al,[29] 2003	Japan (2005-2013)	Prospective observational	Papillary microcarcinomas	Intervention	857/974 (88)	55[e]	3.9 (1.0-9.7)[e]	0/974 (0)	NR
				Comparator	1038/1179 (88)	57[e]	3.9 (1.0-9.7)[e]	0/1179 (0)	NR

Abbreviation: NR, not reported.

[a] All studies were of fair quality.

[b] Intervention group received surgical treatment; comparator or control group received no (immediate) surgical treatment.

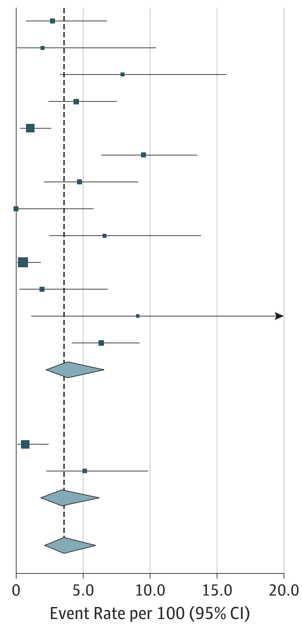
[c] Subset of population (1988-2005) that had treatment recommendation as a variable and refers to Surveillance, Epidemiology, and End Results classification: recommended indicates recommended to be treated; not recommended, not recommended to be treated.

[d] This study reports recurrence in the observation group.

[e] Median.

Figure 3. Key Question 5 Results—Permanent Hypoparathyroidism From Surgery (Event), Stratified by Type of Thyroidectomy

Source	Country	Tumor Size, cm	No. With Permanent Hypoparathyroidism From Surgery/Total No.	Event Rate per 100 (95% CI)
Total thyroidectomy				
Palestini et al,[51] 2008	Italy	<1.0	4/148	2.70 (0.74-6.78)
Kwan et al,[40] 2015	Hong Kong	1.1-2.0	1/51	1.96 (0.05-10.45)
Viola et al,[35] 2015	Italy	1.1-2.0	7/88	7.95 (3.26-15.70)
Ahn et al,[41] 2014	South Korea	1.1-2.0	13/291	4.47 (2.40-7.52)
Conzo et al,[36] 2014	Italy	1.1-2.0	4/390	1.03 (0.28-2.61)
Tartaglia et al,[32] 2014	Italy	1.1-2.0	27/284	9.51 (6.36-13.53)
Calò et al,[44] 2013	Italy	1.1-2.0	8/169	4.73 (2.07-9.11)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	0/62	0.00 (0.00-5.78)
Hartl et al,[48] 2013	France	2.1-4.0	6/91	6.59 (2.46-13.80)
Sywak et al,[58] 2006	Australia	2.1-4.0	2/391	0.51 (0.06-1.84)
Chaplin et al,[45] 1999	Australia	2.1-4.0	2/103	1.94 (0.24-6.84)
Boute et al,[42] 2013	France	NR	2/22	9.09 (1.12-29.16)
Giordano et al,[47] 2012	Italy	NR	25/394	6.35 (4.15-9.22)
Random-effects model for subgroup Heterogeneity: $I^2 = 71.8\%$			101/2484	3.85 (2.23-6.56)
Total or partial thyroidectomy				
Yassa et al,[59] 2007	United States	2.1-4.0	2/296	0.68 (0.08-2.42)
Shindo et al,[54] 1995	United States	NR	8/156	5.13 (2.24-9.85)
Fixed-effects model for subgroup Heterogeneity: $I^2 = 85.2\%$			10/452	3.40 (1.83-6.21)
Random-effects model for all studies Heterogeneity: $I^2 = 72.8\%$			111/2936	3.57 (2.12-5.93)



Tumor size indicates calculated mean tumor size. Size of the data markers indicates the weight used to calculate the pooled estimate. NR indicates not reported.

After approximately 6 years of follow-up, 2 persons in the immediate surgery group and no persons in the observation group had died. An additional 2153 persons were diagnosed with papillary microcarcinoma from 2005 to 2013; of these, 1179 opted for active surveillance and 974 opted for immediate surgery.[31] Only 8% (n = 94) who opted for observation ultimately had surgery. After approximately 4 years of follow-up, no patients in either group developed distant metastases or died from thyroid cancer. In both studies, there were several statistically significant differences between groups; known and potential confounders between treated patients and patients receiving delayed treatment or no treatment were not adjusted for, which limits the ability to compare the effect of treatment on patient outcomes.

Harms of treatment

Key Question 5. What are the harms of treating screen-detected thyroid cancer?

Fifty-two studies (n = 335 091) met inclusion criteria for KQ5. There were 36 studies (n = 43 295 [64 study groups]) of surgical harms, 32 studies (n = 15 811) of permanent hypoparathyroidism (hypocalcemia),[31-62] 28 studies (n = 20 125) of permanent recurrent laryngeal nerve palsy (vocal cord paralysis),[31,32,34,36-42,44-61,63] 2 studies (n = 19 438) of surgical mortality,[64,65] and 15 studies (n = 27 533) of other major surgical harms.[31,36,37,40,43,44,46,56-58,60,61,64-66] The majority of studies of surgical harms were retrospective observational studies, ranging from 76 to 13 854 persons. The main operations evaluated were total or partial thyroidectomy, with or without lymph node dissection (unilateral, bilateral, or not specified; and prophylactic, therapeutic, or not specified).

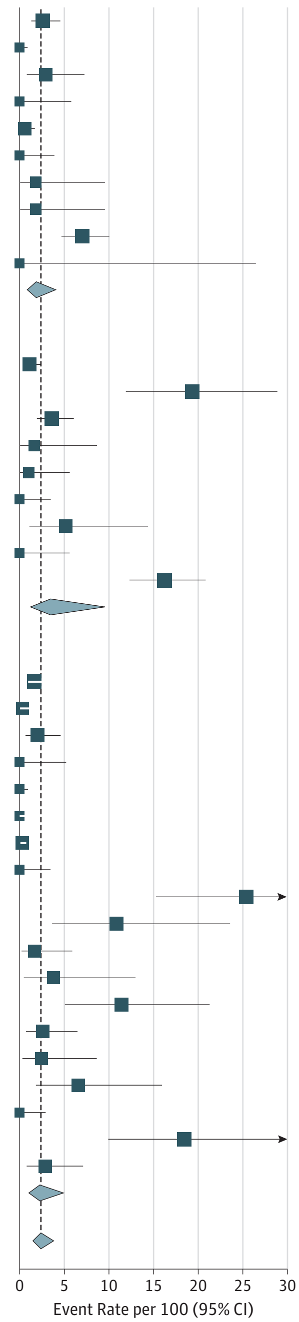
Permanent harm was generally defined as an adverse outcome persisting beyond 6 months. There was large variation in the rate of permanent hypoparathyroidism attributable to total or partial thyroidectomy without lymph node dissection: the 95% CI of the pooled estimate (15 study groups) was 2.12 to 5.93 events per 100 operations ($I^2 = 73\%$) (**Figure 3**). The rate of permanent hypoparathyroidism from thyroidectomy with lymph node dissection was more varied: the 95% CI for unilateral neck dissection (10 study groups) was 0.84 to 4.04 events per 100 operations ($I^2 = 73\%$), and the 95% CI for bilateral neck dissection (9 study groups) was 1.20 to 9.56 events per 100 operations ($I^2 = 91\%$) (**Figure 4**). However, the high degree of statistical heterogeneity may limit the validity of these estimates. The rate of hypoparathyroidism did not seem to vary by year, setting, country, study-level proxies for more advanced tumors, indication for lymph node dissection, or definition of permanent outcomes. Statistical testing suggested biased estimates due to smaller studies, such that smaller studies reported fewer events.

In contrast, there was little variation in the rates of permanent recurrent laryngeal nerve palsy due to thyroidectomy, with or without lymph node dissection. The 95% CI for recurrent laryngeal nerve palsy from thyroidectomy without lymph node dissection (14 study groups) was 0.99 to 2.13 events per 100 operations ($I^2 = 13\%$) (**Figure 5**). Estimates were similar for thyroidectomy with lymph node dissection (33 study groups) (**Figure 6**).

Sixteen studies (n = 291 796) reported harms of radioactive iodine therapy. Eight studies with overlapping populations addressed the risk of second primary malignancies,[67-74] 6 (n = 830) addressed the permanent adverse effects on salivary glands,[75-80]

Figure 4. Key Question 5 Results—Permanent Hypoparathyroidism From Surgery (Event), Stratified by Type of Lymph Node Dissection (Unilateral, Bilateral, Laterality Not Specified)

Source	Country	Tumor Size, cm	No. With Permanent Hypoparathyroidism From Surgery/Total No.	Event Rate per 100 (95% CI)
Unilateral lymph node dissection				
Caliskan et al,[43] 2012	South Korea	<1.0	11/428	2.57 (1.29-4.55)
Caliskan et al,[43] 2012	South Korea	<1.0	0/414	0.00 (0.00-0.89)
Kim et al,[61] 2011	South Korea	<1.0	4/138	2.90 (0.80-7.26)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	0/62	0.00 (0.00-5.78)
Lee et al,[49] 2010	South Korea	1.1-2.0	3/513	0.58 (0.12-1.70)
Palestini et al,[51] 2008	Italy	1.1-2.0	0/93	0.00 (0.00-3.89)
Son et al,[34] 2008	South Korea	1.1-2.0	1/56	1.79 (0.05-9.55)
Sywak et al,[58] 2006	Australia	1.1-2.0	1/56	1.79 (0.05-9.55)
Giordano et al,[47] 2012	Italy	NR	27/385	7.01 (4.67-10.04)
Moo et al,[50] 2009	United States	NR	0/12	0.00 (0.00-26.46)
Random-effects model for subgroup Heterogeneity: $I^2 = 73.4\%$			47/2157	1.86 (0.84-4.04)
Bilateral lymph node dissection				
So et al,[56] 2010	South Korea	<1.0	6/551	1.09 (0.40-2.35)
Viola et al,[35] 2015	Italy	1.1-2.0	18/93	19.35 (11.89-28.85)
Conzo et al,[35] 2014	Italy	1.1-2.0	13/362	3.59 (1.93-6.06)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	1/62	1.61 (0.04-8.66)
Lee et al,[49] 2010	South Korea	1.1-2.0	1/97	1.03 (0.03-5.61)
Moo et al,[50] 2009	United States	1.1-2.0	0/104	0.00 (0.00-3.48)
Son et al,[34] 2008	South Korea	1.1-2.0	3/58	5.17 (1.08-14.38)
Palestini et al,[51] 2008	Italy	2.1-4.0	0/64	0.00 (0.00-5.60)
Giordano et al,[47] 2012	Italy	NR	50/308	16.23 (12.30-20.84)
Random-effects model for subgroup Heterogeneity: $I^2 = 90.8\%$			92/1699	3.46 (1.20-9.56)
Lymph node dissection laterality not specified				
Oda et al,[31] 2016	Japan	<1.0	16/974	1.64 (0.94-2.65)
Chang et al,[37] 2015	South Korea	<1.0	2/613	0.33 (0.04-1.17)
Donatini et al,[39] 2015	France	<1.0	5/251	1.99 (0.65-4.59)
Donatini et al,[39] 2015	France	<1.0	0/69	0.00 (0.00-5.21)
Kim et al,[60] 2014	South Korea	<1.0	0/392	0.00 (0.00-0.94)
Lee et al,[49] 2010	South Korea	<1.0	0/636	0.00 (0.00-0.58)
Lee et al,[49] 2010	South Korea	<1.0	4/1390	0.29 (0.08-0.74)
Del Rio et al,[38] 2015	Italy	1.1-2.0	0/105	0.00 (0.00-3.45)
Tartaglia et al,[32] 2014	Italy	1.1-2.0	16/63	25.40 (15.27-37.94)
Calò et al,[44] 2013	Italy	1.1-2.0	5/46	10.87 (3.62-23.57)
Cirocchi et al,[46] 2012	Italy	1.1-2.0	2/120	1.67 (0.20-5.89)
Kwan et al,[40] 2015	Hong Kong	2.1-4.0	2/53	3.77 (0.46-12.98)
Ahn et al,[41] 2014	South Korea	2.1-4.0	8/70	11.43 (5.07-21.28)
Hartl et al,[48] 2013	France	2.1-4.0	4/155	2.58 (0.71-6.48)
Spear et al,[57] 2008	United States	2.1-4.0	2/81	2.47 (0.30-8.64)
Boute et al,[42] 2013	France	NR	4/61	6.56 (1.82-15.95)
Raj et al,[53] 2010	Australia	NR	0/125	0.00 (0.00-2.91)
Shah et al,[62] 2006	Canada	NR	12/65	18.46 (9.92-30.03)
Sim and Soo,[55] 1998	Singapore	NR	4/141	2.84 (0.78-7.10)
Random-effects model for subgroup Heterogeneity: $I^2 = 89.8\%$			86/5410	2.25 (1.02-4.90)
Random-effects model for all studies			225/9266	2.38 (1.48-3.81)

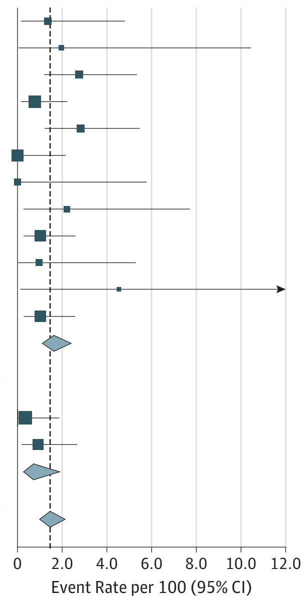


Tumor size indicates calculated mean tumor size. Size of the data markers indicates the weight used to calculate the pooled estimate. NR indicates not reported.

1 (n = 8946) focused on hyperparathyroidism,[81] and 1 (n = 18 850) examined reproductive harms[82] (**Table 5**). Three of the 8 studies examining risk of second primary malignancy due to radioactive iodine therapy used SEER data, none of which reported the indication for or the dose of radiation from radioactive iodine therapy.[67,69,74] Two SEER studies using similar study methods (ie, years studied, definition of second primary malignancy, number of years of follow-up, reference cohort, outcome measures) found that persons who received radioactive iodine therapy for papillary or follicular thyroid cancer had an excess absolute risk of 11.9 to 13.3 cancers per 10 000 person-years compared with a reference cohort.[67,69] The third SEER study had a different study aim

Figure 5. Key Question 5 Results—Permanent Recurrent Laryngeal Nerve Palsy From Surgery (Event), Stratified by Type of Thyroidectomy

Source	Country	Tumor Size, cm	No. With Permanent Recurrent Laryngeal Nerve Palsy From Surgery/Total No.	Event Rate per 100 (95% CI)
Total thyroidectomy				
Palestini et al,[51] 2008	Italy	<1.0	2/148	1.35 (0.16-4.80)
Kwan et al,[40] 2015	Hong Kong	1.1-2.0	1/51	1.96 (0.05-10.45)
Ahn et al,[41] 2014	South Korea	1.1-2.0	8/291	2.75 (1.19-5.34)
Conzo et al,[36] 2014	Italy	1.1-2.0	3/390	0.77 (0.16-2.23)
Tartaglia et al,[32] 2014	Italy	1.1-2.0	8/284	2.82 (1.22-5.47)
Calò et al,[44] 2013	Italy	1.1-2.0	0/169	0.00 (0.00-2.16)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	0/62	0.00 (0.00-5.78)
Hartl et al,[48] 2013	France	2.1-4.0	2/91	2.20 (0.27-7.71)
Sywak et al,[58] 2006	Australia	2.1-4.0	4/391	1.02 (0.28-2.60)
Chaplin et al,[45] 1999	Australia	2.1-4.0	1/103	0.97 (0.02-5.29)
Boute et al,[42] 2013	France	NR	1/22	4.55 (0.12-22.84)
Giordano et al,[47] 2012	Italy	NR	4/394	1.02 (0.28-2.58)
Random-effects model for subgroup Heterogeneity: $I^2 = 3.5\%$			34/2396	1.63 (1.11-2.40)
Total or partial thyroidectomy				
Yassa et al,[59] 2007	United States	2.1-4.0	1/296	0.34 (0.01-1.87)
Shindo et al,[54] 1995	United States	NR	3/326	0.92 (0.19-2.67)
Fixed-effects model for subgroup Heterogeneity: $I^2 = 0\%$			4/622	0.72 (0.27-1.89)
Random-effects model for all studies Heterogeneity: $I^2 = 13.4\%$			38/3018	1.46 (0.99-2.13)



Tumor size indicates calculated mean tumor size. Size of the data markers indicates the weight used to calculate the pooled estimate. NR indicates not reported.

and thus did not report the number of excess cancers by radioactive iodine exposure status.[74]

Nonetheless, this study did not find an association between exposure to radioactive iodine therapy and second primary malignancy using a standardized incidence ratio. However, this study was not limited to differentiated thyroid cancers, included thyroid cancer as a second primary malignancy, and had shorter follow-up for assessment of second primary malignancy. Five smaller studies not conducted in the United States also examined the incidence of second primary malignancies in persons with differentiated thyroid cancer being treated or not treated with radioactive iodine therapy.[68,70-73,83] These studies generally reported the cumulative radiation doses in GBq units. Radiation doses in clinical practice vary and generally correspond to the indication for radioactive iodine therapy, such that lower doses (1.11 GBq) are used for ablation and higher doses (up to 5.5 GBq) are used for adjuvant therapy for known or suspected residual disease.[84] Study results are difficult to compare, given differences in study design, populations, radiation doses, and outcomes. Overall, they demonstrate that use of radioactive iodine is generally associated with an excess risk of second primary malignancy across a range of doses used in clinical practice.

One retrospective study[79] and 5 prospective studies (n = 830) assessed the permanent harms of radioactive iodine therapy on the salivary glands.[75-78,80] The studies were generally small, and the mean radiation dose from radioactive iodine ranged from 1.1 to 5.3 GBq. The most common adverse effect of radioactive iodine on the salivary glands was xerostomia (dry mouth), which ranged from 2.3% to 35%. Dry mouth can adversely affect quality of life and vocal function and increase the risk of dental disease.

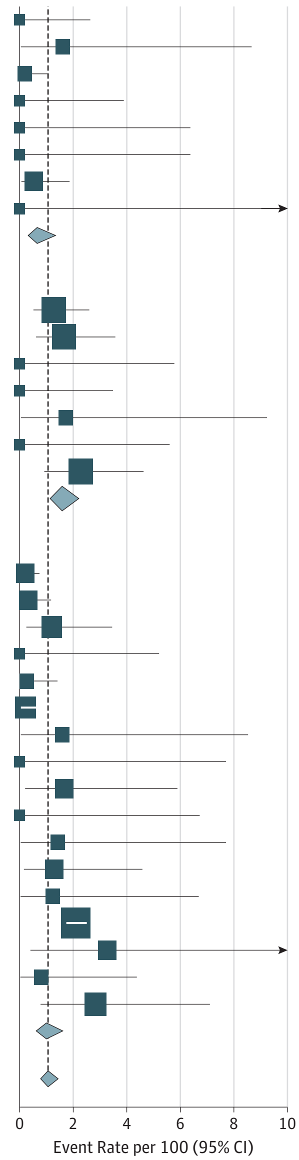
Discussion

A summary of evidence for all KQs is presented in **Table 6**. No trials or well-designed observational studies evaluated the net benefit of thyroid cancer screening. Very limited studies evaluated the true screening accuracy of palpation or neck ultrasound. Ultrasound is very sensitive to detect thyroid nodules, and studies in screening and nonscreening populations have demonstrated that specific high-risk sonographic characteristics can improve sensitivity and specificity for thyroid malignancy.[84-86] Although there was no evidence of serious direct harms from screening and diagnostic follow-up with fine-needle aspiration, screening can result in overdiagnosis; SEER data have demonstrated that almost all persons diagnosed with papillary thyroid cancer receive treatment, so there is the potential of adverse effects from unnecessary treatment. No studies evaluated if treatment of screen-detected cancers compared with symptomatic cancers improves patient health outcomes. It is unclear if immediate surgery, compared with active surveillance, improves patient health outcomes for small or well-differentiated thyroid cancers. Although thyroidectomy is considered a relatively benign operation, permanent hypoparathyroidism and recurrent laryngeal nerve palsy are not uncommon. Additionally, treatment with radioactive iodine is independently associated with a small increase in second primary malignancies, as well as permanent adverse effects on the salivary gland, such as dry mouth.

To accurately estimate the magnitude or effect of overdiagnosis, studies must compare screened and unscreened groups.[87] However, this review found no trials or observational studies

Figure 6. Key Question 5 Results—Permanent Recurrent Laryngeal Nerve Palsy From Surgery (Event), Stratified by Type of Lymph Node Dissection (Unilateral, Bilateral, Laterality Not Specified)

Source	Country	Tumor Size, cm	No. With Permanent Recurrent Laryngeal Nerve Palsy From Surgery/Total No.	Event Rate per 100 (95% CI)
Unilateral lymph node dissection				
Kim et al,[61] 2011	South Korea	<1.0	0/138	0.00 (0.00-2.64)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	1/62	1.61 (0.04-8.66)
Lee et al,[49] 2010	South Korea	1.1-2.0	1/513	0.19 (0.00-1.08)
Palestini et al,[51] 2008	Italy	1.1-2.0	0/93	0.00 (0.00-3.89)
Son et al,[34] 2008	South Korea	1.1-2.0	0/56	0.00 (0.00-6.38)
Sywak et al,[58] 2006	Australia	1.1-2.0	0/56	0.00 (0.00-6.38)
Giordano et al,[47] 2012	Italy	NR	2/385	0.52 (0.06-1.86)
Moo et al,[50] 2009	United States	NR	0/12	0.00 (0.00-26.46)
Random-effects model for subgroup Heterogeneity: $I^2 = 0\%$			4/1315	0.65 (0.32-1.34)
Bilateral lymph node dissection				
So et al,[56] 2010	South Korea	<1.0	7/551	1.27 (0.51-2.60)
Conzo et al,[36] 2014	Italy	1.1-2.0	6/362	1.66 (0.61-3.57)
Raffaelli et al,[52] 2012	Italy	1.1-2.0	0/62	0.00 (0.00-5.78)
Moo et al,[50] 2009	United States	1.1-2.0	0/104	0.00 (0.00-3.48)
Son et al,[34] 2008	South Korea	1.1-2.0	1/58	1.72 (0.04-9.24)
Palestini et al,[51] 2008	Italy	2.1-4.0	0/64	0.00 (0.00-5.60)
Giordano et al,[47] 2012	Italy	NR	7/308	2.27 (0.92-4.63)
Random-effects model for subgroup Heterogeneity: $I^2 = 0\%$			21/1509	1.59 (1.14-2.21)
Lymph node dissection laterality not specified				
Oda et al,[31] 2016	Japan	<1.0	2/974	0.21 (0.02-0.74)
Chang et al,[37] 2015	South Korea	<1.0	2/613	0.33 (0.04-1.17)
Donatini et al,[39] 2015	France	<1.0	3/251	1.20 (0.25-3.45)
Donatini et al,[39] 2015	France	<1.0	0/69	0.00 (0.00-5.21)
Kim et al,[60] 2014	South Korea	<1.0	1/392	0.26 (0.01-1.41)
Lee et al,[49] 2010	South Korea	<1.0	3/1390	0.22 (0.04-0.63)
Tartaglia et al,[32] 2014	Italy	1.1-2.0	1/63	1.59 (0.04-8.53)
Calò et al,[44] 2013	Italy	1.1-2.0	0/46	0.00 (0.00-7.71)
Cirocchi et al,[46] 2012	Italy	1.1-2.0	2/120	1.67 (0.20-5.89)
Kwan et al,[40] 2015	Hong Kong	2.1-4.0	0/53	0.00 (0.00-6.72)
Ahn et al,[41] 2014	South Korea	2.1-4.0	1/70	1.43 (0.04-7.70)
Hartl et al,[48] 2013	France	2.1-4.0	2/155	1.29 (0.16-4.58)
Spear et al,[57] 2008	United States	2.1-4.0	1/81	1.23 (0.03-6.69)
Francis et al,[63] 2014	United States	NR	119/5670	2.10 (1.74-2.51)
Boute et al,[42] 2013	France	NR	2/61	3.28 (0.40-11.35)
Raj et al,[53] 2010	Australia	NR	1/125	0.80 (0.02-4.38)
Sim and Soo,[55] 1998	Singapore	NR	4/141	2.84 (0.78-7.10)
Random-effects model for subgroup Heterogeneity: $I^2 = 59.9\%$			144/10274	1.00 (0.62-1.61)
Random-effects model for all studies			169/13098	1.06 (0.78-1.44)



Tumor size indicates calculated mean tumor size. Size of the data markers indicates the weight used to calculate the pooled estimate. NR indicates not reported.

comparing thyroid cancer screening with no screening. However, there are ecologic data on the trend of incidence and mortality of thyroid cancer, autopsy data, and limited natural history data to suggest that overdiagnosis of thyroid cancer is a problem. Multiple studies have shown an increase in the incidence in thyroid cancer detection over time, with no change in the mortality rate.[1,7,27,88-90] Several studies by Davies and Welch[1,27,90] have used SEER data to estimate the incidence of thyroid cancer and cancer-related mortality in the United States since the 1970s. The absolute increase in the incidence of thyroid cancer from 1975 to 2009 in the United States was 9.4 (95% CI, 8.9-9.9) cases per 100 000 persons, of which 9.1 (95% CI, 8.6-9.6) cases per 100 000 persons were papillary cancers.[1] Data from other countries have shown similar findings. Data from the Cancer Incidence in Five Continents database showed steady increases in thyroid cancer incidence in 12 selected countries from 1960 to 2007, primarily due to an increase in papillary carcinoma diagnoses.[88] South Korea has had an organized cancer screening program since 1999.[7] Although the program did not officially include thyroid cancer screening, physicians frequently offered thyroid

Table 5. Included Studies for Key Question 5—Harms of Radioactive Iodine Treatment of Screen-Detected Thyroid Cancer

Source[a]	Country (Recruitment Years)	Study Design	No.	Women, No. (%)	Mean Age, y	Reported	
						Second Primary Malignancy	Salivary Gland Harms
Hakala et al,[70] 2015	Finland (1981-2002)	Retrospective observational	910	746 (82)	49	✓	
Khang et al,[71] 2015	South Korea, (1976-2010)	Retrospective observational	2468	2073 (84)	46	✓	
Lin et al,[72] 2015	Taiwan (2000-2008)	Prospective observational	10 361	10 361 (100)	46	✓	
Seo et al,[73] 2015	South Korea (2008-2013)	Retrospective observational	211 360	173 315 (82)	48	✓	
Lang et al,[68] 2012 Lang and Wong,[83] 2011	Hong Kong (1971-2009)	Retrospective observational	895	725 (81)	47	✓	
Iyer et al,[69] 2011	United States (1973-2006)	Retrospective observational	37 176	NR	NR	✓	
Brown et al,[67] 2008	United States (1973-2002)	Retrospective observational	28 286[b]	21 497 (76)	42[c]	✓	
Ronckers et al,[74] 2005	United States (1973-2000)	Prospective observational	29 456	22 092 (75)	43[c]	✓	
Ryu et al,[80] 2015	South Korea (2010)	Prospective observational	160	130 (81)	49		✓
Jeong et al,[77] 2013	South Korea (2003-2006)	Prospective observational	213	194 (91)	47		✓
Grewal et al,[79] 2009	United States (1995-2003)	Retrospective observational	262	173 (66)	45		✓
Ish-Shalom et al,[76] 2008	Israel (NR)	Prospective observational	40	40 (100)	48		✓
Hyer et al,[75] 2007	United Kingdom (NR)	Prospective observational	76	57 (75)	51		✓
Solans et al,[78] 2001	Spain (1990-1995)	Prospective observational	79	68 (86)	46		✓[d]
Wu et al,[82] 2015[e]	United States (1999-2008)	Retrospective observational	18 850[f]	18 850 (100)	47		
Lin et al,[81] 2014[g]	Taiwan (1997-2008)	Retrospective observational	8946	7246 (81)	44		

Abbreviations: NR, not reported.

[a] Lin et al study[82] was of good quality; all other studies were of fair quality.

[b] Total number included in the radioactive iodine analysis, 28 286; total N for study, 30 278.

[c] Median.

[d] Study also reported dry eyes.

[e] Study reported adverse reproductive outcomes related to birthrate and median time to first delivery as serious harms.

[f] Total cohort included 25 333 persons; the reproductive outcomes subset included 18 850 women.

[g] Study reported hyperparathyroidism as a serious harm.

screening with ultrasound for a small additional cost. The rate of thyroid cancer diagnoses increased from 5 cases per 100 000 persons in 1993 to 70 cases per 100 000 persons in 2011.[7]

Autopsy studies have provided additional evidence on overdiagnosis of thyroid cancer. A 2014 review by Lee et al[91] summarized 15 studies published between 1969 and 2005 on latent thyroid cancer discovered at autopsy. Of 8619 thyroid glands obtained at autopsy, 989 (11.5%) were positive for papillary thyroid carcinoma. The proportion of papillary thyroid cancers varied widely, from 1.0% to 35.6%. Studies describing the natural history of thyroid nodules and malignancies also lend evidence to the problem of overdiagnosis of thyroid cancer. Durante et al[92] described a 5-year follow-up of 992 patients with benign thyroid nodules (0.4 cm to 4 cm). In 686 patients (69%), the size of the nodules remained stable; in 184 (18.5%), the size of 1 or more nodules decreased; and in 153 (15.4%), the size of 1 or more nodules increased by 20% or more (the groups were not mutually exclusive, because some persons had more than 1 nodule). No studies had follow-up of benign nodules beyond 5 years. The studies included in this review, as well as other studies, demonstrate the slow-growing nature of thyroid tumors and the low potential for recurrence or mortality due to papillary tumors and microcarcinomas.[27,29,93-95] However, data on the survival of patients who never receive treatment are very limited.

Limitations

This evidence review focused on screening practices relevant to general US practice in adults and therefore did not include studies primarily focused on cohorts exposed to high doses of radiation via environmental disasters or treated with radiation for childhood cancers. Additionally, it did not systematically review the diagnostic accuracy of ultrasound to detect thyroid cancer in nonscreening populations. The review of harms was limited to those directly related to surgery or radioactive iodine therapy (eg, excluded harms from suppressive doses of thyroxine). Older studies of harms were excluded because over time surgery, radioactive iodine doses, and the case mix of persons undergoing treatment have changed.

Although population-based screening trials for thyroid cancer are unlikely, trials or well-designed observational studies to address the benefit of screening in higher-risk populations (eg, those with a personal history of irradiation or a family history of differentiated

Table 6. Summary of Evidence, by Key Question

Test or Intervention	No. Studies (Design)	Summary of Findings[a]	Body of Evidence Limitations[b]	Quality	Applicability
KQ1: Effectiveness					
NA	0	No trials have evaluated effect of screening for thyroid cancer on patient morbidity or mortality.	NA	NA	NA
KQ2: Diagnostic Accuracy					
Palpation	2 (prospective diagnostic accuracy [n = 354]) 4 (prospective and retrospective cohort [n = 201 027])	Neck palpation not sensitive (11.6% to 27.8%) in detecting nodules compared with ultrasound. Yield of cancers ranged from 0 to 4.3 per 1000 persons. Yield of cancers in adults with history of irradiation ranged from 0 to 11.3 cancers per 1000 persons.	Two small studies reported diagnostic accuracy, 1 study did not follow up all persons with neck palpations. No evidence of reporting bias.	Fair	Poor: diagnostic accuracy studies are old and use a single examiner in Finland
Ultrasound	2 (prospective/retrospective diagnostic accuracy [n = 243]) 2 (prospective and retrospective cohort [n = 2094])	Using any 1 of several malignant sonographic characteristics can be highly sensitive (94.3%) for detecting cancers; using a combination (≥2) of these characteristics can be both highly sensitive (94.8%) and specific (86.6%). Yield of cancers ranged from 9.2 to 30.3 cancers per 1000 persons.	Two small studies reported diagnostic accuracy, neither of which followed up with the majority of screened individuals, such that the reported sensitivities are likely overestimates. No evidence of reporting bias.	Fair	Fair: both diagnostic accuracy studies conducted in South Korea by the same investigators, 1 of which included women only
KQ3: Screening Harms					
Ultrasound	1 (retrospective cohort [n = 400])	Twenty-four percent of persons underwent fine-needle aspiration of a nodule that did not meet the Society of Radiologists in Ultrasound criteria for fine-needle aspiration.	Only 1 study	Fair	Poor: single-institution; standards for referral to fine-needle aspiration have changed
Ultrasound-guided fine-needle aspiration	2 (retrospective cohort [n = 5494])	One study (n = 4912) observed 7 cases of needle tract implantation of papillary thyroid cancer with fine-needle aspiration. It is unclear what effect if any this had on patient outcomes. One study observed hematomas but no major bleeding complications requiring hospitalization from fine-needle aspiration.	One study for each type of harm; possible reporting bias	Fair	Fair: both single-institution studies
KQ4: Treatment Benefit					
Surgery	2 (prospective and retrospective cohort [n = 39 211])	US SEER data demonstrate that, overall, untreated persons with papillary thyroid cancer had a slightly worse 20-y survival rate (97%) than treated persons (99%) ($P < .001$). One Japanese study found no deaths in persons with papillary microcarcinoma who opted for ultrasound observation compared with 2 deaths in persons who opted for immediate surgery.	Studies were not designed to evaluate the comparative benefit of treatment vs no or delayed treatment. Lack of adjustment for confounders such that it is unclear if differences in survival are due to differences in treatment vs case mix of persons. No evidence of reporting bias.	Fair to poor	Fair: US study includes persons treated in 1970s and 1980s; Japanese study includes persons with papillary microcarcinoma
KQ5: Treatment Harms					
Surgery	36 (prospective and retrospective cohort [n = 43 295])	The rate of permanent hypoparathyroidism varied widely; best estimates were between 2 to 6 events per 100 thyroidectomies and were more variable with lymph node dissection. The rate of recurrent laryngeal nerve palsy was less variable, estimated at 1 to 2 events per 100 operations (with or without lymph node dissection).	Possible publication bias for hypoparathyroidism but not recurrent laryngeal nerve palsy outcomes. Reasons for the wide variation in estimates is unclear.	Fair	Fair: indication for type of surgery and case mix of patients going on to surgery have changed over time
Radioactive iodine	16 (prospective/retrospective cohort [n = 291 796])[c]	Treatment with radioactive iodine for differentiated thyroid cancer is associated with a small increase in primary second malignancies: approximately 12 to 13 excess cancers per 10 000 patients. Smaller studies demonstrate an association of excess cancers at clinically used doses. Other commonly reported permanent harms from radioactive iodine include dry mouth, ranging from 2.3% to 35% of persons.	Differences in study designs and variable reporting on radiation doses limits understanding of the magnitude and precision around risk of second primary malignancies. No evidence-reporting bias for commonly reported adverse outcomes.	Fair	Fair: indication and radiation dose of radioactive iodine have changed over time

Abbreviations: NA, not applicable; SEER, Surveillance, Epidemiology, and End Results.

[a] Includes consistency and precision.

[b] Includes reporting bias.

[c] Calculated sample size includes only the largest study using SEER data so as to avoid double-counting studies with overlapping populations.

thyroid cancers) would be helpful to understand if there is any role for screening for thyroid cancer. Given the indolent nature of many thyroid cancers and the risks associated with treatment, trials or well-designed observational studies examining the benefit of early treatment vs observation or surveillance for patients with (smaller) well-differentiated thyroid cancers are also needed. The net benefit of screening hinges on minimizing overdiagnosis and overtreatment; therefore, for screening to be of benefit, studies are needed to determine which set of prognostic indicators predict aggressive vs indolent disease.

Conclusions

Although ultrasonography of the neck using high-risk sonographic characteristics plus follow-up cytology from fine-needle aspiration can identify thyroid cancers, it is unclear if population-based or targeted screening can decrease mortality rates or improve important patient health outcomes. Screening that results in the identification of indolent thyroid cancers, and treatment of these overdiagnosed cancers, may increase the risk of patient harms.

ARTICLE INFORMATION

Accepted for Publication: January 23, 2017.

Author Contributions: Dr Lin had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: This research was funded under contract number HHSA-290-2012-00015-I-357 EPC4, Task Order 4, from the Agency for Healthcare Research and Quality (AHRQ) under a contract to support the USPSTF.

Role of the Funder/Sponsor: Investigators worked with USPSTF members and AHRQ staff to develop the scope, analytic framework, and key questions for this review. AHRQ had no role in study selection, quality assessment, or synthesis. AHRQ staff provided project oversight, reviewed the report to ensure that the analysis met methodological standards, and distributed the draft for peer review. Otherwise, AHRQ had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript findings. The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

Additional Contributions: We gratefully acknowledge the following individuals for their contributions to this project: Jennifer Croswell, MD, formerly at the Agency for Healthcare Research and Quality; current and former members of the US Preventive Services Task Force who contributed to topic deliberations; Gunjan Tykodi, MD (Kaiser Permanente Washington Health Research Institute), for expert consultation; Todd Hannon, MLS, and Smyth Lai, MLS (Kaiser Permanente Research Affiliates Evidence-based Practice Center), for creating and conducting the literature searches; Elizabeth L. Hess, MS (Kaiser Permanente Research Affiliates Evidence-based Practice Center), for editorial assistance; and Ning X. Smith, PhD, and Brittany Burda, DHSc (Kaiser Permanente Research Affiliates Evidence-based Practice Center), for assistance with meta-analyses. USPSTF members, expert consultants, and peer reviewers did not receive financial compensation for their contributions.

Additional Information: A draft version of this evidence report underwent external peer review from 4 invited content experts (Louise Davies, MD, School of Medicine at Dartmouth; Hyeong Sik Ahn, MD, Korea University School of Medicine; Edward G. Grant, MD, Keck School of Medicine, University of Southern California; and Mike Tuttle, MD, Memorial Sloan Kettering Cancer Center) and 3 federal partners (National Cancer Institute; Centers for Disease Control and Prevention; and the National Institute of Diabetes and Digestive and Kidney Diseases, Diabetes, Endocrinology and Obesity Branch). Comments were presented to the USPSTF during its deliberation of the evidence and were considered in preparing the final evidence review.

Editorial Disclaimer: This evidence report is presented as a document in support of the accompanying USPSTF Recommendation Statement. It did not undergo additional peer review after submission to *JAMA*.

REFERENCES

1. Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg*. 2014;140(4):317-322.

2. National Cancer Institute (NCI). Cancer Stat Facts: thyroid cancer. NCI website. https://seer.cancer.gov/statfacts/html/thyro.html. 2014. Accessed November 30, 2015.

3. Cooper DS, Doherty GM, Haugen BR, et al; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer [published corrections appear in *Thyroid*. 2010;20(6):674-675 and 2010;20(8):942]. *Thyroid*. 2009;19(11):1167-1214.

4. Xing MM. Thyroid carcinoma. ClinicalKey website. https://www.clinicalkey.com. 2012. Accessed August 5, 2014.

5. Cramer JD, Fu P, Harth KC, Margevicius S, Wilhelm SM. Analysis of the rising incidence of thyroid cancer using the Surveillance, Epidemiology, and End Results national cancer data registry. *Surgery*. 2010;148(6):1147-1152.

6. Hughes DT, Haymart MR, Miller BS, Gauger PG, Doherty GM. The most commonly occurring papillary thyroid cancer in the United States is now a microcarcinoma in a patient older than 45 years. *Thyroid*. 2011;21(3):231-236.

7. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic"—screening and overdiagnosis. *N Engl J Med*. 2014;371(19):1765-1767.

8. US Preventive Services Task Force. *US Preventive Services Task Force Procedure Manual*. Rockville, MD: Agency for Healthcare Research and Quality; 2008. AHRQ publication 08-05118 EF.

9. Harris RP, Helfand M, Woolf SH, et al; Methods Work Group, Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20(3)(suppl):21-35.

10. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute website. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. 2000. Accessed January 24, 2017.

11. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3(1):25.

12. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.

13. Bucci A, Shore-Freedman E, Gierlowski T, Mihailescu D, Ron E, Schneider AB. Behavior of small thyroid cancers found by screening radiation-exposed individuals. *J Clin Endocrinol Metab*. 2001;86(8):3711-3716.

14. Ishida T, Izuo M, Ogawa T, Kurebayashi J, Satoh K. Evaluation of mass screening for thyroid cancer. *Jpn J Clin Oncol*. 1988;18(4):289-295.

15. Suehiro F. Thyroid cancer detected by mass screening over a period of 16 years at a health care center in Japan. *Surg Today*. 2006;36(11):947-953.

16. Brander A, Viikinkoski P, Nickels J, Kivisaari L. Thyroid gland: US screening in a random adult population. *Radiology*. 1991;181(3):683-687.

17. Brander A, Viikinkoski P, Nickels J, Kivisaari L. Thyroid gland: US screening in middle-aged women with no previous thyroid disease. *Radiology*. 1989;173(2):507-510.

18. Kim SJ, Moon WK, Cho N. Sonographic criteria for fine-needle aspiration cytology in a Korean female population undergoing thyroid ultrasound screening. *Acta Radiol*. 2010;51(5):475-481.

19. Kim JY, Lee CH, Kim SY, et al. Radiologic and pathologic findings of nonpalpable thyroid carcinomas detected by ultrasonography in a medical screening center. *J Ultrasound Med*. 2008;27(2):215-223.

20. Lee HK, Hur MH, Ahn SM. Diagnosis of occult thyroid carcinoma by ultrasonography. *Yonsei Med J*. 2003;44(6):1040-1044.

21. Chung WY, Chang HS, Kim EK, Park CS. Ultrasonographic mass screening for thyroid carcinoma: a study in women scheduled to undergo a breast examination. *Surg Today*. 2001;31(9):763-767.

22. Ron E, Lubin E, Modan B. Screening for early detection of radiation-associated thyroid cancer: a pilot study. *Isr J Med Sci*. 1984;20(12):1164-1168.

23. Shimaoka K, Bakri K, Sciascia M, et al. Thyroid screening program; follow-up evaluation. *N Y State J Med*. 1982;82(8):1184-1187.

24. Hobbs HA, Bahl M, Nelson RC, Eastwood JD, Esclamado RM, Hoang JK. Applying the Society of Radiologists in Ultrasound recommendations for fine-needle aspiration of thyroid nodules: effect on workup and malignancy detection. *AJR Am J Roentgenol*. 2014;202(3):602-607.

25. Abu-Yousef MM, Larson JH, Kuehn DM, Wu AS, Laroia AT. Safety of ultrasound-guided fine needle aspiration biopsy of neck lesions in patients taking antithrombotic/anticoagulant medications. *Ultrasound Q*. 2011;27(3):157-159.

26. Ito Y, Tomoda C, Uruno T, et al. Needle tract implantation of papillary thyroid carcinoma after fine-needle aspiration biopsy. *World J Surg*. 2005;29(12):1544-1549.

27. Davies L, Welch HG. Thyroid cancer survival in the United States: observational data from 1973 to 2005. *Arch Otolaryngol Head Neck Surg*. 2010;136(5):440-444.

28. Ito Y, Miyauchi A, Inoue H, et al. An observational trial for papillary thyroid microcarcinoma in Japanese patients. *World J Surg*. 2010;34(1):28-35.

29. Ito Y, Uruno T, Nakano K, et al. An observation trial without surgical treatment in patients with papillary microcarcinoma of the thyroid. *Thyroid*. 2003;13(4):381-387.

30. Ito Y, Miyauchi A, Kihara M, Higashiyama T, Kobayashi K, Miya A. Patient age is significantly related to the progression of papillary microcarcinoma of the thyroid under observation. *Thyroid*. 2014;24(1):27-34.

31. Oda H, Miyauchi A, Ito Y, et al. Incidences of unfavorable events in the management of low-risk papillary microcarcinoma of the thyroid by active surveillance versus immediate surgery. *Thyroid*. 2016;26(1):150-155.

32. Tartaglia F, Blasi S, Giuliani A, et al. Central neck dissection in papillary thyroid carcinoma: results of a retrospective study. *Int J Surg*. 2014;12(suppl 1):S57-S62.

33. Hyun SM, Song HY, Kim SY, et al. Impact of combined prophylactic unilateral central neck dissection and hemithyroidectomy in patients with papillary thyroid microcarcinoma. *Ann Surg Oncol*. 2012;19(2):591-596.

34. Son YI, Jeong HS, Baek CH, et al. Extent of prophylactic lymph node dissection in the central neck area of the patients with papillary thyroid carcinoma: comparison of limited versus comprehensive lymph node dissection in a 2-year safety study. *Ann Surg Oncol*. 2008;15(7):2020-2026.

35. Viola D, Materazzi G, Valerio L, et al. Prophylactic central compartment lymph node dissection in papillary thyroid carcinoma: clinical implications derived from the first prospective randomized controlled single institution study. *J Clin Endocrinol Metab*. 2015;100(4):1316-1324.

36. Conzo G, Calò PG, Sinisi AA, et al. Impact of prophylactic central compartment neck dissection on locoregional recurrence of differentiated thyroid cancer in clinically node-negative patients: a retrospective study of a large clinical series. *Surgery*. 2014;155(6):998-1005.

37. Chang YW, Kim HS, Kim HY, Lee JB, Bae JW, Son GS. Should central lymph node dissection be considered for all papillary thyroid microcarcinoma? *Asian J Surg*. 2016;39(4):197-201.

38. Del Rio P, Maestroni U, Sianesi M, et al. Minimally invasive video-assisted thyroidectomy for papillary thyroid cancer: a prospective 5-year follow-up study. *Tumori*. 2015;101(2):144-147.

39. Donatini G, Castagnet M, Desurmont T, Rudolph N, Othman D, Kraimps JL. Partial thyroidectomy for papillary thyroid microcarcinoma: is completion total thyroidectomy indicated? *World J Surg*. 2016;40(3):510-515.

40. Kwan WY, Chow TL, Choi CY, Lam SH. Complication rates of central compartment dissection in papillary thyroid cancer. *ANZ J Surg*. 2015;85(4):274-278.

41. Ahn D, Sohn JH, Park JY. Surgical complications and recurrence after central neck dissection in cN0 papillary thyroid carcinoma. *Auris Nasus Larynx*. 2014;41(1):63-68.

42. Boute P, Merlin J, Biet A, Cuvelier P, Strunski V, Page C. Morbidity of central compartment dissection for differentiated thyroid carcinoma of the follicular epithelium. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2013;130(5):245-249.

43. Caliskan M, Park JH, Jeong JS, et al. Role of prophylactic ipsilateral central compartment lymph node dissection in papillary thyroid microcarcinoma. *Endocr J*. 2012;59(4):305-311.

44. Calò PG, Medas F, Pisano G, et al. Differentiated thyroid cancer: indications and extent of central neck dissection—our experience. *Int J Surg Oncol*. 2013;2013:625193.

45. Chaplin JM, O'Brien CJ, McNeil EB, Haghighi K. Application of prognostic scoring systems in differentiated thyroid carcinoma. *Aust N Z J Surg*. 1999;69(9):625-628.

46. Cirocchi R, Boselli C, Guarino S, et al. Total thyroidectomy with ultrasonic dissector for cancer: multicentric experience. *World J Surg Oncol*. 2012;10:70.

47. Giordano D, Valcavi R, Thompson GB, et al. Complications of central neck dissection in patients with papillary thyroid carcinoma: results of a study on 1087 patients and review of the literature. *Thyroid*. 2012;22(9):911-917.

48. Hartl DM, Mamelle E, Borget I, Leboulleux S, Mirghani H, Schlumberger M. Influence of prophylactic neck dissection on rate of retreatment for papillary thyroid carcinoma. *World J Surg*. 2013;37(8):1951-1958.

49. Lee YS, Nam KH, Chung WY, Chang HS, Park CS. Postoperative complications of thyroid cancer in a single center experience. *J Korean Med Sci*. 2010;25(4):541-545.

50. Moo TA, Umunna B, Kato M, et al. Ipsilateral versus bilateral central neck lymph node dissection in papillary thyroid carcinoma. *Ann Surg*. 2009;250(3):403-408.

51. Palestini N, Borasi A, Cestino L, Freddi M, Odasso C, Robecchi A. Is central neck dissection a safe procedure in the treatment of papillary thyroid cancer? our experience. *Langenbecks Arch Surg*. 2008;393(5):693-698.

52. Raffaelli M, De Crea C, Sessa L, et al. Prospective evaluation of total thyroidectomy versus ipsilateral versus bilateral central neck dissection in patients with clinically node-negative papillary thyroid carcinoma. *Surgery*. 2012;152(6):957-964.

53. Raj MD, Grodski S, Martin SA, Yeung M, Serpell JW. The role of fine-needle aspiration cytology in the surgical management of thyroid cancer. *ANZ J Surg*. 2010;80(11):827-830.

54. Shindo ML, Sinha UK, Rice DH. Safety of thyroidectomy in residency: a review of 186 consecutive cases. *Laryngoscope*. 1995;105(11):1173-1175.

55. Sim R, Soo KC. Surgical treatment of thyroid cancer: the Singapore General Hospital experience. *J R Coll Surg Edinb*. 1998;43(4):239-243.

56. So YK, Son YI, Hong SD, et al. Subclinical lymph node metastasis in papillary thyroid microcarcinoma: a study of 551 resections. *Surgery*. 2010;148(3):526-531.

57. Spear SA, Theler J, Sorensen DM. Complications after the surgical treatment of malignant thyroid disease. *Mil Med*. 2008;173(4):399-402.

58. Sywak M, Cornford L, Roach P, Stalberg P, Sidhu S, Delbridge L. Routine ipsilateral level VI lymphadenectomy reduces postoperative thyroglobulin levels in papillary thyroid cancer. *Surgery*. 2006;140(6):1000-1005.

59. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*. 2007;111(6):508-516.

60. Kim BS, Kang KH, Kang H, Park SJ. Central neck dissection using a bilateral axillo-breast approach for robotic thyroidectomy: comparison with conventional open procedure after propensity score matching. *Surg Laparosc Endosc Percutan Tech*. 2014;24(1):67-72.

61. Kim WW, Kim JS, Hur SM, et al. Is robotic surgery superior to endoscopic and open surgeries in thyroid cancer? *World J Surg*. 2011;35(4):779-784.

62. Shah MD, Witterick IJ, Eski SJ, Pinto R, Freeman JL. Quality of life in patients undergoing thyroid surgery. *J Otolaryngol*. 2006;35(4):209-215.

63. Francis DO, Pearce EC, Ni S, Garrett CG, Penson DF. Epidemiology of vocal fold paralyses after total thyroidectomy for well-differentiated thyroid cancer in a Medicare population. *Otolaryngol Head Neck Surg*. 2014;150(4):548-557.

64. Zerey M, Prabhu AS, Newcomb WL, Lincourt AE, Kercher KW, Heniford BT. Short-term outcomes after unilateral versus complete thyroidectomy for malignancy: a national perspective. *Am Surg*. 2009;75(1):20-24.

65. Hundahl SA, Cady B, Cunningham MP, et al. Initial results from a prospective cohort study of 5583 cases of thyroid carcinoma treated in the United States during 1996: U.S. and German Thyroid Cancer Study Group: an American College of Surgeons Commission on Cancer Patient Care Evaluation study. *Cancer*. 2000;89(1):202-217.

66. Hölzer S, Reiners C, Mann K, et al; U.S. and German Thyroid Cancer Group. Patterns of care for patients with primary differentiated carcinoma of the thyroid gland treated in Germany during 1996. *Cancer*. 2000;89(1):192-201.

67. Brown AP, Chen J, Hitchcock YJ, Szabo A, Shrieve DC, Tward JD. The risk of second primary malignancies up to three decades after the treatment of differentiated thyroid cancer. *J Clin Endocrinol Metab*. 2008;93(2):504-515.

68. Lang BH, Wong IO, Wong KP, Cowling BJ, Wan KY. Risk of second primary malignancy in differentiated thyroid carcinoma treated with radioactive iodine therapy. *Surgery*. 2012;151(6):844-850.

69. Iyer NG, Morris LG, Tuttle RM, Shaha AR, Ganly I. Rising incidence of second cancers in patients with low-risk (T1N0) thyroid cancer who receive radioactive iodine therapy. *Cancer*. 2011;117(19):4439-4446.

70. Hakala TT, Sand JA, Jukkola A, Huhtala HS, Metso S, Kellokumpu-Lehtinen PL. Increased risk of certain second primary malignancies in patients treated for well-differentiated thyroid cancer [published online September 26, 2015]. *Int J Clin Oncol*. doi:10.1007/s10147-015-0904-6

71. Khang AR, Cho SW, Choi HS, et al. The risk of second primary malignancy is increased in differentiated thyroid cancer patients with a cumulative (131)I dose over 37 GBq. *Clin Endocrinol (Oxf)*. 2015;83(1):117-123.

72. Lin CY, Lin CL, Huang WS, Kao CH. Risk of breast cancer in patients with thyroid cancer receiving or not receiving 131I treatment: a nationwide population-based cohort study [published online December 30, 2015]. *J Nucl Med*. doi:10.2967/jnumed.115.164830

73. Seo GH, Cho YY, Chung JH, Kim SW. Increased risk of leukemia after radioactive iodine therapy in patients with thyroid cancer: a nationwide, population-based study in Korea. *Thyroid*. 2015;25(8):927-934.

74. Ronckers CM, McCarron P, Ron E. Thyroid cancer and multiple primary tumors in the SEER cancer registries. *Int J Cancer*. 2005;117(2):281-288.

75. Hyer S, Kong A, Pratt B, Harmer C. Salivary gland toxicity after radioiodine therapy for thyroid cancer. *Clin Oncol (R Coll Radiol)*. 2007;19(1):83-86.

76. Ish-Shalom S, Durleshter L, Segal E, Nagler RM. Sialochemical and oxidative analyses in radioactive I131-treated patients with thyroid carcinoma. *Eur J Endocrinol*. 2008;158(5):677-681.

77. Jeong SY, Kim HW, Lee SW, Ahn BC, Lee J. Salivary gland function 5 years after radioactive iodine ablation in patients with differentiated thyroid cancer: direct comparison of pre- and postablation scintigraphies and their relation to xerostomia symptoms. *Thyroid*. 2013;23(5):609-616.

78. Solans R, Bosch JA, Galofré P, et al. Salivary and lacrimal gland dysfunction (sicca syndrome) after radioiodine therapy. *J Nucl Med*. 2001;42(5):738-743.

79. Grewal RK, Larson SM, Pentlow CE, et al. Salivary gland side effects commonly develop several weeks after initial radioactive iodine ablation. *J Nucl Med*. 2009;50(10):1605-1610.

80. Ryu CH, Ryu J, Ryu YM, et al. Administration of radioactive iodine therapy within 1 year after total thyroidectomy does not affect vocal function. *J Nucl Med*. 2015;56(10):1480-1486.

81. Lin CM, Doyle P, Tsan YT, Lee CH, Wang JD, Chen PC; Health Data Analysis in Taiwan (hDATa) Research Group. 131I treatment for thyroid cancer and risk of developing primary hyperparathyroidism: a cohort study. *Eur J Nucl Med Mol Imaging*. 2014;41(2):253-259.

82. Wu JX, Young S, Ro K, et al. Reproductive outcomes and nononcologic complications after radioactive iodine ablation for well-differentiated thyroid cancer. *Thyroid*. 2015;25(1):133-138.

83. Lang BH, Wong KP. Risk factors for nonsynchronous second primary malignancy and related death in patients with differentiated thyroid carcinoma. *Ann Surg Oncol*. 2011;18(13):3559-3565.

84. Haugen BRM, Alexander EK, Bible KC, et al; American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2016;26(1):1-133.

85. Brito JP, Gionfriddo MR, Al Nofal A, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2014;99(4):1253-1263.

86. Lin JS, Aiello Bowles EJ, Williams SB, Morrison CC. *Screening for Thyroid Cancer: A Systematic Review for the US Preventive Services Task Force*. Rockville, MD: Agency for Healthcare Research and Quality; 2016.

87. Carter SM, Rogers W, Heath I, Degeling C, Doust J, Barratt A. The challenge of overdiagnosis begins with its definition. *BMJ*. 2015;350:h869.

88. La Vecchia C, Bosetti C, Malvezzi M, et al. Author's reply to thyroid cancer: an epidemic of disease or an epidemic of diagnosis? *Int J Cancer*. 2015;136(11):2740.

89. Ho AS, Davies L, Nixon IJ, et al. Increasing diagnosis of subclinical thyroid cancers leads to spurious improvements in survival rates. *Cancer*. 2015;121(11):1793-1799.

90. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973-2002. *JAMA*. 2006;295(18):2164-2167.

91. Lee YS, Lim H, Chang H-S, Park CS. Papillary thyroid microcarcinomas are different from latent papillary thyroid carcinomas at autopsy. *J Korean Med Sci*. 2014;29(5):676-679.

92. Durante C, Costante G, Lucisano G, et al. The natural history of benign thyroid nodules. *JAMA*. 2015;313(9):926-935.

93. Noguchi S, Yamashita H, Uchino S, Watanabe S. Papillary microcarcinoma. *World J Surg*. 2008;32(5):747-753.

94. Yu XM, Wan Y, Sippel RS, Chen H. Should all papillary thyroid microcarcinomas be aggressively treated? an analysis of 18,445 cases. *Ann Surg*. 2011;254(4):653-660.

95. Tuttle M. Clinicopathologic staging of differentiated thyroid cancer. UpToDate website. 2011. http://cursoenarm.net/UPTODATE/contents/mobipreview.htm?10/39/10879. Accessed April 6, 2016.