UK National Screening Committee

Screening for thyroid dysfunction in adults

External review against programme appraisal criteria for the UK National Screening Committee (UK NSC)

Version: Final

Solutions for Public Health

June 2017

The UK NSC advises Ministers and the NHS in all four UK countries about all aspects of screening policy. Its policies are reviewed on a 3 yearly cycle. Current policies can be found in the policy database at http://legacy.screening.nhs.uk/screening-recommendations.php and the policy review process is described in detail at https://www.gov.uk/guidance/evidence-and-recommendations-nhs-population-screening#evidence-review-process

Abbreviations List

CI	Confidence Interval
FT4	Free Thyroxine
FT3	Free Tri-iodothyronine
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NPV	Negative Predictive Value
OR	Odds Ratio
PICO	Population Intervention Comparator Outcomes
PPV	Positive Predictive Value
P value	Probability value
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
SH	Sub clinical hypothyroidism
RCT	Randomised Controlled Trial
TFTs	Thyroid Function Tests
TPO	Thyroid Peroxidase
TSH	Thyroid Stimulating Hormone
T4	Thyroxine
TT4	Total T4
T3	Tri-iodothyronine
TT3	Total Tri-iodothyronine
UK	United Kingdom
UK NSC	UK National Screening Committee

Competing Interest

All SPH authors have completed the ICMJE uniform disclosure form (www.icmje.org/coi_disclosure.pdf) and declare: grants from Public Health England to SPH to undertake the submitted work, no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work

Plain English Summary

The thyroid gland located at the front of the neck is responsible for the production of 2 hormones (T3 and T4) which are essential for regulating energy levels. A hormone from the pituitary gland located at the base of the brain (TSH) controls the level of these 2 hormones.

Thyroid dysfunction develops when these hormones are out of balance and can be divided into hyperthyroidism (over production of thyroid hormones) and hypothyroidism (under production of thyroid hormones). In sub clinical disease TSH is at an abnormal level and T3 and T4 are normal. In hypothyroidism if only TSH is abnormally high then this is diagnosed as sub clinical hypothyroidism but if T4 and or T3 are too low at the same time this is overt hypothyroidism. In hyperthyroidism if only TSH is too low then this is sub clinical hyperthyroidism but if the hormones T4 and or T3 are too high at the same time this is overt hyperthyroidism.

In the general population we would expect 3% to have some form of hypothyroidism (overt or sub clinical), with around 259 new cases occurring each year for every 100,000 people.

For hyperthyroidism we would expect 0.75% of the general population to have some from of the condition(either overt or sub clinical), with around 51 new cases occurring each year for every 100,000 population.

A wide range of consequences of thyroid dysfunction have been reported and include an increased risk of heart disease, decreased bone density, and stroke. Many people with thyroid dysfunction don't have any symptoms.

This document reviews new evidence about screening adults for thyroid dysfunction. It looks at evidence published between January 2011 and January 2017. The aim of a screening programme would be to identify people with either sub clinical disease or undiagnosed hyper and hypothyroidism. This would be followed by treatment to stop symptoms and reduce the risk of heart disease, stroke and other diseases linked to thyroid dysfunction.

The UK National Screening Committee (UK NSC) published its last review in 2013. This recommended against introducing a screening programme for thyroid dysfunction in the UK. The main points were:

- there is a lack of agreement on what the normal range of TSH, T4 and T3 should be from blood tests from patients.
- there isn't clear evidence on what the harms of treatment would be.
- some people may have an abnormal blood test for TSH, T4 and T3 at one point in time but this could change to normal at a later date. It's not clear who or how many people change back to normal without treatment.
- there is only limited evidence on the benefits of treatment of people who are diagnosed by a screening programme for sub clinical thyroid disease. More evidence is needed.

This review

The key questions considered in this review are:

- key question 1: what proportion of people with overt and sub clinical thyroid disease change back to normal without any treatment?

- key question 2: if the population were screened for thyroid disease is there an agreed blood test result that would show who should have treatment and who shouldn't?
- sub question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
- key question 3: how effective is the treatment for thyroid disease and what are the side effects?

Review findings

The review found the amount, quality and type of new evidence published since January 2011 does not show that systematic population screening for thyroid dysfunction in adults should be recommended in the UK. Several uncertainties remain.

Key question 1: what proportion of people with overt and sub clinical thyroid disease change back to normal without any treatment?

The studies available were limited to people with sub clinical hyperthyroidism and one study of older people for sub clinical hypothyroidism. No studies were found for the other groups of people (people under 65 with sub clinical hypothyroidism and anyone with overt hyper or hypothyroidism).

The evidence available showed that thyroid status can change back to normal in some people whilst others may change from normal to abnormal and back again over time. For people with sub clinical hyperthyroidism the proportion who reverted to normal reported in the studies described in this review were between 31.6% and 53% (over a period of time of between 2 months and 9 years).

The sub clinical hypothyroidism study focused on an older population and reported that at 2 years 35% of people with sub clinical hypothyroidism had reverted to normal however at 4 years this had dropped to 17%. This indicates that some people had fluctuated from sub clinical disease when they were diagnosed, to normal 2 years later and back again at 4 years.

The studies suggest there are a number of underlying reasons that can have an effect on the thyroid hormones such as age, sex, type of thyroid disease and the presence of antibodies that mistakenly attack the thyroid gland. However it is not possible to determine an obvious group of people whose thyroid status is more or less likely to revert to normal.

Key question 2: if the population were screened for thyroid disease is there an agreed blood test result that would show who should have treatment and who shouldn't.

Sub question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults

The studies published since the previous review did not show that there was agreement of what is considered normal or abnormal blood test results for TSH, FT4 and FT3 or who should have treatment based on the results of the test.

Key question3: how effective is the treatment for thyroid disease and what are the side effects?

One study showed that treatment (with levothyroxine) for people aged 40-70 years with sub clinical hypothyroidism may be beneficial in reducing the chances of developing cardiovascular disease which can be linked to thyroid disease. However the limitations of this type of study may lead to bias in the results as all the health information about all the participants isn't known. Another study in older people did not find that levothyroxine improved hypothyroid symptoms or tiredness.

Overall, there is a lack of high quality evidence that people who are shown to have an abnormal test result of TSH, T4 and T3 through a screening programme will benefit from treatment.

Recommendation

This update of the evidence found that no new evidence has been published to change the conclusions of the previous UK NSC review therefore population screening for thyroid disease is not recommended.

Executive Summary

This document reviews new evidence published between January 2011 and January 2017 on population screening for thyroid dysfunction.

Background

The thyroid gland located in the front of the neck is responsible for the production of 2 hormones, thyroxine (T4) and tri-iodothyronine (T3). These thyroid hormones are essential for the normal maturation and metabolism of all tissues in the body. In health these hormones are tightly regulated by the secretion of thyroid stimulating hormone (TSH) from the pituitary gland located at the base of the brain.

Thyroid dysfunction can be divided into hyperthyroidism (over production of thyroid hormones) and hypothyroidism (under production of thyroid hormones). Each of these can be further categorized into overt diagnosed, overt but un-diagnosed, and sub clinical disease.

Sub clinical thyroid dysfunction is defined biochemically, on the basis of an abnormal level of TSH in association with normal levels of thyroid hormones. Individuals may or may not experience symptoms with sub clinical thyroid dysfunction.

A 2014 meta-analysis of studies estimating the prevalence of thyroid dysfunction in Europe reported their findings in each of these categories (Table 1) [1].

Table 1: Estimated prevalence (%) of thyroid disease in Europe[2].

Thyroid dysfunction	Prevalence (%)(95% CI)
Total hypothyroidism	3.05(3.01-3.09)
Overt hypothyroidism	0.43(0.42-0.43)
Sub clinical hypothyroidism	2.60(2.56-2.63)
Total hyperthyroidism	0.75(0.73-0.77)
Overt hyperthyroidism	0.11(0.10-0.11)
Sub clinical hyperthyroidism	0.64(0.62-0.66)

The study also estimated incidence of combined overt and sub clinical hypothyroidism as 259.12 per 100,000 per year (95% CI 254.4-263.9) and overt and sub clinical hyperthyroidism as 51.04 per 100,000 per year (CI 95% 49.23-52.88).

Overt thyroid disease is associated with negative cardiovascular, musculoskeletal, dermatologic, gastrointestinal and other effects which are highly variable and depend on the level of abnormality.

Sub clinical hypothyroidism has in some studies been associated with increased risk of coronary artery disease whilst sub clinical hyperthyroid disease has been additionally associated with atrial fibrillation and decreased bone density.

Previous findings

The current UK NSC recommendation is that systematic population screening for thyroid dysfunction should not be offered. Clinical practice guidelines are covered by guidance from the National Institute of Health and Care.

The previous UK NSC external review of screening for thyroid dysfunction was published in 2013. The 2013 review concluded that there was insufficient information to recommend a population screening programme at that time. The gaps in the evidence related to:

- a lack of consensus about the TSH cut-off value or values for defining which patients should receive treatment, and what constitutes a normal TSH level is still a matter of debate (British Thyroid Association and the British Thyroid Foundation 2011).
- two systematic reviews of Randomised Controlled Trials(RCTs) on treatment of sub clinical thyroid disease have concluded that there is insufficient evidence of benefit and that the potential adverse effects from treatment have not been adequately studied.
- lack of clarity about the proportion of people with screen-detected sub clinical hypothyroidism who would revert to normal without treatment, (although it could be a substantial minority).
- the apparent short-term (4 month) benefits of treating screen-detected sub clinical hypothyroidism that were demonstrated in 15 subjects by Abu-Helalah et al (2010) need to be confirmed in a larger study with longer follow-up.
- a lack of RCTs of screening for thyroid disease.

The current review

The current review explores the volume, quality and direction of the literature published since 2011 and focuses on key questions relating to the conclusions of the previous review. The aim of the review is to inform discussion on whether the recent evidence provides a sufficient basis on which to recommend the introduction of a population screening programme for thyroid dysfunction in the UK.

The key questions considered in this review are:

- key question 1: what proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
- key question 2: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?
- sub question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
- key question3: what is the effectiveness of treatment of overt and sub clinical thyroid disease?

The review found the volume, quality and direction of new evidence published since January 2011 does not show that systematic population screening for thyroid dysfunction in adults should be recommended in the UK. Several uncertainties remain.

Key question 1: what proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?

The evidence identified as relevant for this question was limited to people with sub clinical hyperthyroidism and one study of an elderly cohort for sub clinical hypothyroidism. No

evidence for the proportion of people whose thyroid function normalises following overt disease or the wider adult population for sub clinical hypothyroidism was identified.

The evidence indicated that thyroid status can resolve to normal in some people whilst others may experience repeated fluctuations from normal to sub clinical status overtime.

The 5 studies tracking people with sub clinical hyperthyroidism and change in thyroid status reported between 31.6% and 53% patients with TSH reverting to within normal reference range values over a period of between 2 months to 9 years.

The one sub clinical hypothyroidism study focused on an older population (>65 years) and reported that at 2 years 35% of people with sub clinical hypothyroidism at baseline had reverted to normal however at 4 years this had dropped to 17%. This indicates that some people had fluctuated from sub clinical status at baseline to normal at 2 years and back again at 4 years at least once.

The evidence suggests there are multiple underlying factors that can impact thyroid status change such as age, disease aetiology and baseline TSH values. However it is not possible to determine an obvious cohort of people whose thyroid status is more likely to revert to normal.

Key question 2: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?

Sub-question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?

The literature search of publications since the previous review did not identify any studies that lead to a consensus of the reference ranges and thresholds of TSH, T4 or T3 as a whole population screening test for sub clinical or overt thyroid disease.

Key question3: what is the effectiveness of treatment of overt and sub clinical thyroid disease?

One retrospective cohort study provided some evidence that levothyroxine may confer cardiovascular benefits on younger patients (40-70 years) diagnosed with sub clinical hypothyroidism. However the limitations of this type of study lead to the possibility of residual confounding by participants with unknown cardiovascular disease status. One RCT in older people did not find that levothyroxine improved hypothyroid symptoms or tiredness.

Overall, there is a lack of evidence that people with screen detected sub clinical and overt thyroid disease will benefit from treatment and no new high quality evidence was identified considering treatment for undiagnosed overt thyroid dysfunction or sub clinical hyperthyroidism.

Recommendation

This update of the evidence found that no new evidence has been published to change the conclusions of the previous UK NSC review therefore population screening for thyroid disease is not recommended.

Introduction

The thyroid gland located at the front of the neck is responsible for the production of 2 hormones, thyroxine (T4) and tri-iodothyronine (T3). These thyroid hormones are essential for the normal maturation and metabolism of all tissues in the body. In health these hormones are tightly regulated by the secretion of thyroid stimulating hormone (TSH) from the pituitary gland located at the base of the brain. Thyroid hormones are present in the blood in either protein bound forms (the majority) or the free and active (unbound) forms of the hormones. The amount of unbound T3 is referred to as free T3 and unbound T4 as free T4. Currently, the majority of UK laboratories measure the free form of the hormones free T4 (FT4) or free T3 (FT3) to diagnose or rule out thyroid dysfunction in people presenting with symptoms of clinical disease.

Thyroid dysfunction can be divided into hyperthyroidism (over production of thyroid hormones) and hypothyroidism (under production of thyroid hormones). Each of these can be further categorized into overt and sub clinical disease. A 2014 meta-analysis of studies estimating the prevalence of thyroid dysfunction in Europe reported their findings in each of these categories (Table 1) [2].

Table 1: Estimated prevalence (%) of thyroid disease in Europe[2].

Thyroid dysfunction	Prevalence (%)(95% CI)
Total hypothyroidism	3.05(3.01-3.09)
Overt hypothyroidism	0.43(0.42-0.43)
Sub clinical hypothyroidism	2.60(2.56-2.63)
Total hyperthyroidism	0.75(0.73-0.77)
Overt hyperthyroidism	0.11(0.10-0.11)
Sub clinical hyperthyroidism	0.64(0.62-0.66)

The study also estimated incidence of combined overt and sub clinical hypothyroidism as 259.12 per 100,000 per year (95% CI 254.4-263.9) and overt and sub clinical hyperthyroidism as 51.04 per 100,000 per year (CI 95% 49.23-52.88).

The most common cause of hypothyroidism is chronic autoimmune (Hashimoto's) thyroiditis. Other causes include previously treated thyroid dysfunction, poor adherence to or under treatment with levothyroxine, external beam radiation in the head and neck area, and untreated adrenal insufficiency. Increased risk of hypothyroidism is reported in women, older age groups, people of Caucasian origin, people with type 1 diabetes mellitus, those with a family history of thyroid dysfunction, and people with Down syndrome[11].

Causes of hyperthyroidism include Graves' disease, autoimmune thyroiditis (Hashitoxicosis), functional thyroid nodules, overtreatment with levothyroxine and ingestion of iodine-containing drugs, such as amiodarone. Increased risk, of

hyperthyroidism is reported in women, older age groups, people of Caucasian origin people with a personal or family history of hyperthyroidism and those with low iodine intake[11].

Overt thyroid dysfunction is defined biochemically, on the basis of abnormal levels of both TSH and FT4 or FT3. Individuals may or may not experience symptoms with overt thyroid dysfunction irrespective of biochemical assay results of hormone levels. Overt undiagnosed thyroid disease represents a condition found in prevalence studies in which biochemical testing identifies people with TSH and FT4 beyond the laboratory determined reference ranges but who have not been previously diagnosed by their doctors. Sub clinical thyroid dysfunction is defined biochemically, on the basis of an abnormal level of TSH in association with normal levels of unbound thyroid hormones FT4 and FT3.

The euthyroid TSH reference range varies by testing platform and laboratory however the British Thyroid Association recommend a reference range for TSH of 0.4 -4.0mU/l[3]. FT4 also varies by laboratory and platform but the Association for Clinical Biochemistry and British Thyroid Foundation recommends a range of 9 to 25 pmol/l[4]). The recommended range for FT3 is 3.5-7.8 pmol/l[3].

Overt hypothyroidism

A diagnosis of overt hypothyroidism is indicated by a high serum TSH (≥4.0mU/l) and a low FT4 (<9mU/l) or T3 (<3.5pmol/l). Consequences of hypothyroidism include an association with a number of coronary heart disease risk factors including dyslipidaemia and hypertension[5]. Memory disturbance and dementia have also been described[6]. Additionally in rare cases coma can develop which can be fatal.

Overt hyperthyroidism

A diagnosis of overt hyperthyroidism is indicated by a low serum TSH (≤ 0.39 mU/l[7]) and a high FT4 (≥ 25.0 pmol/l) or FT3 outside of normal values (3.5-7.8pmol/l). Consequences of untreated hyperthyroidism include atrial fibrillation[8] and osteoporosis as a result of reduced bone mineral density[9].

Sub clinical hypothyroidism

Those with sub clinical hypothyroidism would be typically diagnosed with TSH concentrations above the normal reference range (≥4.0-20 mU/l) and where FT4 and FT3 are normal. In the UK the start of treatment is recommended once TSH levels reach >10mU/l or between 4.0-9.9mU/l in patients with symptoms who are under 65 years old[10]**Error! Hyperlink reference not valid.**. Sub clinical hypothyroidism has in some studies been associated with increased risk of coronary artery disease[11].

Sub clinical hyperthyroidism

Sub clinical hyperthyroidism is diagnosed where TSH values are below the normal reference range (≤0.39mU/l) and FT4 and FT3 are normal. Treatment of sub clinical hyperthyroidism is usually offered to people with a TSH level persistently equal to or less than 0.1 mU/l, if they are aged 65 years or older, are postmenopausal, are at risk of osteoporosis, have cardiac risk factors, or have symptoms of hyperthyroidism[12]. Sub clinical hyperthyroid disease has been associated with coronary heart disease and atrial fibrillation.

Basis for current recommendation

The current UK NSC recommendation made by the previous external review published in 2013 is that systematic population screening for thyroid dysfunction in adults is not recommended[13].The aim of a national screening programme for thyroid dysfunction would be to detect undiagnosed overt cases in order to start treatment and identify sub clinical cases with a high likelihood of progressing to overt disease. This requires high quality evidence about the screening and diagnostic test strategies including the cut off thresholds for TSH, FT4 and FT3 blood tests. In addition evidence of the best strategy to manage sub clinical cases detected by the programme would need to be evidenced and agreed.

The 2013 review concluded that:

- there is a lack of consensus about the TSH cut-off value or values for defining which patients should receive treatment, and what constitutes a normal TSH level is still a matter of debate.
- two systematic reviews of RCTs on treatment of sub clinical thyroid disease have concluded that there is insufficient evidence of benefit and that the potential adverse effects from treatment have not been adequately studied.
- it is unclear what proportion of people with screen-detected sub clinical hypothyroidism would revert to normal without treatment, but it could be a substantial minority.
- the apparent short-term (4 month) benefits of treating screen-detected sub clinical hypothyroidism that were demonstrated in 15 subjects by Abu-Helalah et al (2010)[14] need to be confirmed in a larger study with longer follow-up.
- there have been no RCTs of screening for thyroid disease.

Current update review and approach taken

The current review considers screening for thyroid dysfunction and was prepared by Solutions for Public Health, in discussion with the UK NSC evidence team.

The current evidence summary was developed using a rapid review methodology and assessed using the UK NSC reporting checklist for evidence summaries. The key questions addressed in the current review were developed by the UK NSC evidence team and are based on the key areas where thyroid dysfunction did not meet the criteria for a screening programme in the last UK NSC review. The aim of the current review is to update the evidence in these key areas, namely around:

- the proportion of people with sub clinical hypothyroidism that revert to normal levels of TSH.
- agreed cut-off values of TSH for defining which patients should receive treatment and a consensus of the healthy reference range of TSH, T4 and T3.
- the effectiveness of treatment of undiagnosed overt and sub clinical thyroid disease.

The key questions and the UK NSC criteria that they relate to are presented in Table 1 below.

A systematic literature search of 3 databases was conducted by the UK NSC evidence team in January 2017 for new evidence published since January 2011. The search was structured around the issues raised in the 2013 UK NSC external review. A total of 3463 unique references were identified and sifted by title and abstract by the UK NSC evidence team for potential relevance to the review. Details of the databases searched, search terms and a flow diagram summarising the references identified are presented in the Search Strategy section at the end of this report. A total of 282 references were sent to Solutions for Public Health for further appraisal and possible inclusion in the final review. Selection and appraisal of studies was undertaken by one reviewer. Any queries were resolved through discussion with a second reviewer or with the UK NSC evidence team.

Overall, 69 studies were identified as potentially relevant during title and abstract sifting and full text was further assessed. This includes papers where relevance could not be determined from the title or abstract alone. Reasons for excluding studies at the abstract stage included:

- non-systematic reviews
- clinical management
- guidelines (other than Europe, UK or US)
- primary care treatment strategies
- clinical performance and reference standards developed with specific equipment (eg; Roche cobas e411)

Each section below provides information on the evidence selection process and number of included studies for the given criterion.

The review was quality assured by a second senior reviewer who was not involved with the writing of the review in accordance with Solutions for Public Health's quality assurance process.

Table 1: Key questions for current review of screening for thyroid dysfunction

Criterion*	Key Questions	# Studies Included
1. The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?	6
5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?	5
9. There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered.	What is the effectiveness of treatment of overt and sub clinical thyroid disease?	3

* UK NSC evidence review criteria (January 2016)

Appraisal against UK NSC Criteria[†]

Each of the key questions and their associated criteria are considered in turn below.

Each criterion was summarised as 'met', 'not met' or 'uncertain' by considering the results of the included studies in light of the volume, quality and consistency of the body of evidence. Several factors were considered in determining the quality of the identified evidence, including study design and methodology, risk of bias and applicability of the evidence.

Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.

The 2013 UK NSC review reported that the condition was an important health problem and that the incidence and prevalence rates of the disease have been published. However there was insufficient understanding of the natural history of the disease. In particular it was unclear what proportion of people with screen-detected sub clinical hypothyroidism would revert to normal without treatment[13]. Key question 1 of this review update is concerned with identifying new evidence which might provide clarity about this area.

Key question1: what proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?

This key question includes both hyperthyroidism and hypothyroidism. The PICO developed by the UK NSC evidence team states that studies of screen detected populations should be prioritised if available and that results should be reported by duration of follow-up[15].

Description of the evidence

In the current review, of the studies assessed at full text, 6 were identified as potentially relevant to this question and were included (detailed in Appendices 1 to 6).

Of the 6 studies included 5 were concerned with the progression from sub clinical hyperthyroidism to overt hyperthyroidism or resolution to normal. The sixth study focused on progression, persistence or resolution to normal of sub clinical hypothyroidism in older people.

Proportion of overt and sub clinical hyperthyroidism normalising over time

The table below summarises the results of the 5 studies identified that focus on the changes in thyroid status of people with sub clinical hyperthyroidism (at baseline) over time. No studies using screen detected populations were identified and there were no studies that showed the normalisation of overt hyperthyroidism over time.

[†]These criteria are available online at UK NSC evidence review criteria (January 2016)

Table 2: Details of 5 studies concerned with normalisation or progression of sub clinical hyperthyroidism (SH) over time

	Population	Range of follow up	Proportion normalizing	Number(%) persisting; SH	Progression to overt hyperthyroidism	Factors influencing change in status
Zhyzhneuskaya et al 2016[16]	44 patients with sub clinical hyperthyroidism due to Graves disease‡	2-111 months	15(34%)	13(30%)	15(34%)	Increase in age by 1 year(Hazard ratio 1.06 – 95% CI1.02-1.10,p<0.01) Positive thyroid peroxidase antibody test (HR =10.15, 95% CI 1.83-56.23, p<0.01)
Das et al 2012[17]	323 adult patients with sub clinical hyperthyroidism	6-93 months (mean 32 months)	102(31.6%)	183(56.7%)	38(11.8%)	Patients with TSH level 0.1-0.39mU/l at baseline more likely to remain sub clinical than those with TSH levels <0.1mU/L (20.3% vs 6.8%, p<0.001).
Poola et al 2011[18]	116 patients with sub clinical hyperthyroidism	6 months-6.5 years	61(53%)	44(38%)	11(9%)	Presence or absence of nodules (p=<.001)
Schouten et al 2011[19]	96 patients with sub clinical hyperthyroidism (16-91years) from Graves disease (n=12), multi-nodule goiter (n=70) and autonomous nodule (n=14)‡	5.5 years	Not separately reported	Not separately reported	Overall =25 (17%) 1(9%) GD 15(21%) MNG 9(61%) (AN)	Change in status determined by underlying disease (p=0.003) Proportion of results normalizing or persisting as SH: Graves disease =11(91%), multi nodule goiter = 55(79%), autonomous goiter = 5(39%)
Vadiveloo et al 2011[20]	2024 cases with sub clinical hyperthyroidism from the general population aged ≥18	2yr(n=1044) 5 yr(n=693) 7 yr(n=495)	2 yr=168(16.1%) 5 yr=177(26.5%) 7 yr =134(27.1%)	2 yr=800(76.6%) 5yr=379(36.3%) 7yr=237(47.9%)	2yr=6(0.6%) 5yr=4(0.6%) 7yr=2(0.4%)	Reversion to normal was more common when TSH was between 0.1-0.4mU/L. Reversion to normal in the group <0.1mU/l was less likely in older people (, OR 0.97 CI 0.95-1.0, p=0.02).

‡ Graves disease(GD), multi nodal goitre(MNG) and autonomous nodule(AN) are all underlying causes of hyperthyroidism

A prospective cohort study of 44 people with sub clinical hyperthyroidism by Zhyzhneuskaya et al (2016)[16] found that people with a baseline TSH below 0.4mU/l with normal thyroid hormone levels changed status over time. A third each of the patients progressed to overt Graves disease[‡], normalised or remained in the sub clinical hyperthyroid state. Older people and those with positive thyroid peroxidase (TPO) antibodies (which suggest underlying autoimmune thyroid disease) had a higher risk of progression.

Das et al (2012)[17] reported a retrospective study of 323 patients with sub clinical hyperthyroidism with a mean follow up of 32 months. Two categories of TSH values were applied to the results, TSH between 0.10 to 0.39mU/l and TSH below 0.1mU/l. Overall most people remained sub clinically hyperthyroid (56.7%) or reverted to normal (31.6%). Those people with TSH lower values at baseline (<0.1mU/l) were more likely to progress to overt hyperthyroidism at follow up than those with higher (0.1-0.39mU/l) TSH values (20.3% vs 6.8%, p<0.001).

Poola et al (2011)[18] reported a retrospective study of 116 people with sub clinical hyperthyroidism (TSH <0.4mU/l) of which 53% normalised between baseline and follow up. A further 38% remained sub clinical and 11% progressed to overt hyperthyroidism. The study analysed presence or absence of thyroid nodules and of the 18(16%) with thyroid nodules 15(83%) did not revert to normal at follow up and 7(39%) required treatment. In comparison of those 98 patients without thyroid nodules 58(59%) reverted to normal thyroid status at follow up and 4(4%) patients received treatment.

Schouten et al (2011)[20] evaluated the progression to overt hyperthyroidism retrospectively in 96 consecutive patients diagnosed with sub clinical hyperthyroidism over a 6 year period. Progression to overt disease was seen in 8% of the total cohort at 1 year, 16% at 2 years, 21% at 3 years and 26% at 5 years. Progression to overt hyperthyroidism varied depending on disease aetiology (p=0.003). The cumulative percentage of people with sub clinical Graves disease requiring treatment for overt disease at 5 years was 9%, 21% for multinodular goitre and 61% for the autonomous nodule sub group.

Vadiveloo et al (2011)[20] as part of the Thyroid Epidemiology, Audit and Research Study (TEARS) retrospectively identified 2024 people with sub clinical hyperthyroidism from 8 databases capturing health information of the Tayside population. Follow up TSH values were extracted from the databases up to 7 years after baseline sub clinical hyperthyroid diagnosis. Overall the percentages of people who did not start any treatment and who reverted to normal were 17.2% after 2 years, 31.5% after 5 years and 35.6% after 7 years. The percentage of people who developed overt hyperthyroidism was 0.6% after 2 years 0.7% after 5 years and 0.5% after 7 years. Age at index date was an important factor effecting reversion to normality for people with TSH <0.1mU/l. With increasing age there was less likelihood of TSH levels returning to normal (odds ratio (OR) 0.97 CI 0.95 to 1.0, p=0.02).

Proportion of overt and sub clinical hypothyroidism normalising over time

There were no studies identified that showed the normalisation of overt hypothyroidism over time.

One longitudinal study by Somwaru et al (2012)[21] was identified that focussed on progression, persistence or resolution of sub clinical hypothyroidism of an elderly cohort (>65 years) who had been enrolled in the US Cardiovascular Health Study.

The authors[21] identified 3996 people who had been enrolled in the study and who had had a thyroid test.

From the baseline test 459(11.5%) people had sub clinical hypothyroidism. Blood samples from this group were tested for thyroid status at 2 years and 4 years after the baseline samples had been taken.

The 3 figures below illustrate how the thyroid function status changed as the cohort was tracked over 4 years.

Figure 1 shows the baseline cohort of 459 people with sub clinical hypothyroidism (SH) and the status of the cohort 2 years later. The largest group of 208(56%) individuals continued to have detectable sub clinical hypothyroidism, 128(35%) had results within the normal reference range and 33(7%) had either progressed to overt hypothyroidism or started treatment. There were no measurements for 90 (20%) people due to inadequate remaining sample, lack of follow up or death of the participant.

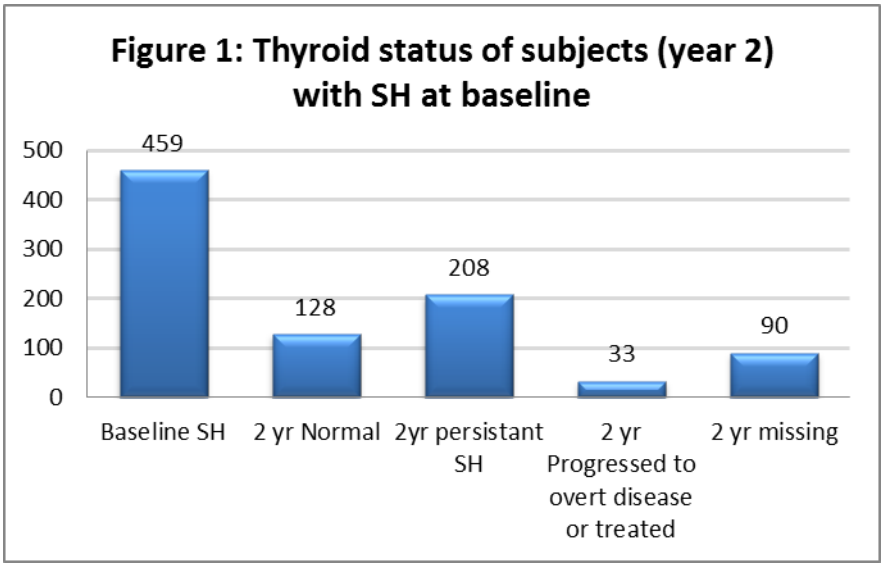


Figure 2 shows the thyroid status at 4 years for people who had had a baseline result of sub clinical hypothyroidism and a normal result at year 2. Of these 128 people, 62(48%) continued to have a normal thyroid function status at 4 years, 41(32%) had reverted back to sub clinical hypothyroid status and 4 (4%) had progressed or started treatment. For 20(16%) there was no measurement (due to inadequate sample, death of participant or no follow up).

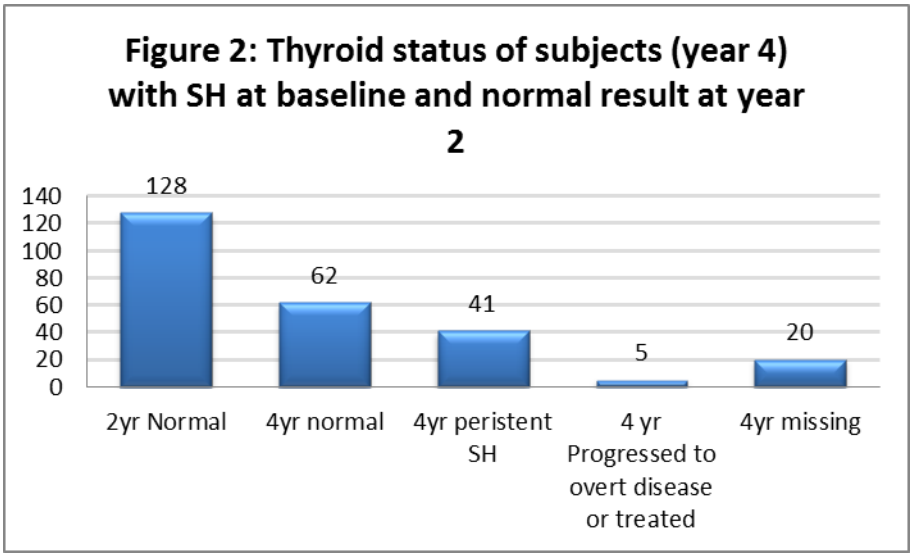
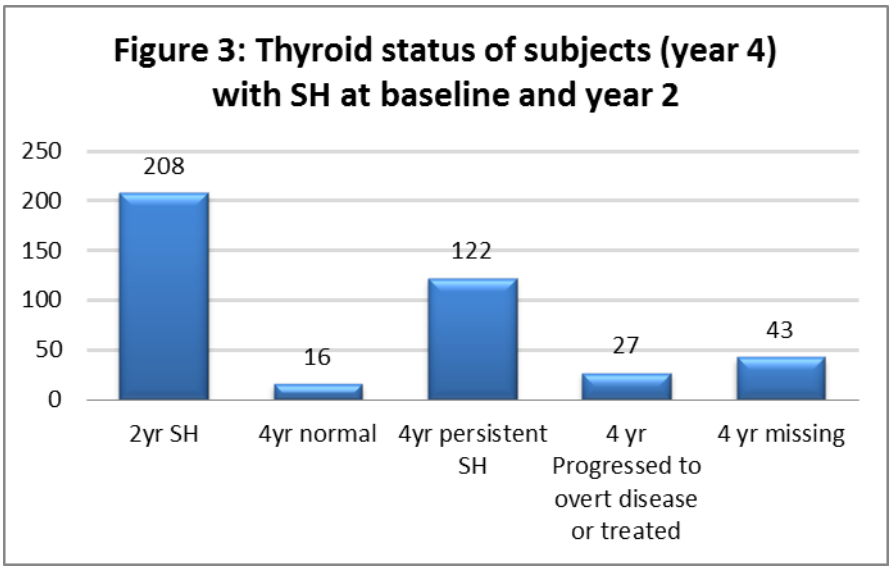


Figure 3 shows the thyroid status at 4 years for individuals who were sub clinical hypothyroid at baseline and at 2 years (n=208). The largest proportion of 122(59%) individuals remained sub clinically hypothyroid at 4 years, 27(13%) progressed or began treatment and 16(7%) resolved to normal thyroid function status. Measurements from 43(21%) people were not made (due to inadequate sample, no follow up or death of participant).



Overall of the 459 people at baseline with sub clinical hypothyroidism, at 4 years around a third (153) were missing through lack of follow up, 78(17%) had reverted to normal TSH levels, 65(14%) had progressed to overt hypothyroidism and 163(36%)continued to have sub clinical disease. By tracking peoples' TSH levels at 2 and 4 years it is clear that thyroid status can fluctuate from normal to sub clinical and back again over a period of time.

Discussion

Evidence was identified describing normalisation of thyroid status from sub clinical thyroid dysfunction but not from overt disease. No studies were identified that included people whose thyroid dysfunction had been screen detected. One large study concerning sub clinical

hypothyroidism included older people from aged 65 years and over whilst the other studies included adults of all ages.

Normalisation of sub clinical hyperthyroidism

The 5 studies tracking people with sub clinical hyperthyroidism and change in thyroid status have reported between 31.6% and 53% patients with TSH reverting to within normal reference range values. These changes have happened over varying periods of follow up from 2 months to over 9 years. Studies detected variations in the underlying factors which appeared to impact the likelihood of reversion to normal. This included presence or absence of thyroid nodules, baseline TSH level and age.

Four of the 5 studies were retrospective analyses of data from patients' records from first sub clinical diagnosis. Large numbers of subjects can be included in these types of study but analysis is limited to the accuracy and type of historical clinical data recorded at the time of diagnosis. The large population based retrospective study[20] analysing stored blood samples from the general population was the study most similar to a screen detected population but as with the other retrospective studies the limitations include a lack of full health information which could bias the results. In this case the study could not identify patients using oestrogens, androgens, corticosteroids, oral contraceptives or the presence of goitre or thyroid nodular disease.

The fifth study was a small prospective study of 44 patients with sub clinical hyperthyroidism due to Graves disease. These patients were symptomatic with one underlying condition and so do not represent the variation in severity and cause of sub clinical hypothyroidism likely to be encountered in a screen detected population.

Normalisation of sub clinical hypothyroidism

One longitudinal cohort study[21] was included that focused on sub clinical hypothyroidism in an older population (>65 years) who had enrolled in a cardiovascular health study. At 4 years 17% of people with sub clinical hypothyroidism at baseline reverted to normal. Individuals also changed thyroid status from normal to sub clinical and back again over a number of years without progressing to overt disease. This study is limited to an older population and the results may not apply to the whole adult population. However it does illustrate how TSH levels fluctuate across the reference range thresholds which may lead to over treatment.

Summary: Criterion 1: not met

The evidence identified was limited to people with sub clinical hyperthyroidism and 1 study of an elderly cohort for sub clinical hypothyroidism. No evidence for the proportion of people whose thyroid function normalises following overt disease or the wider adult population for sub clinical hypothyroidism was identified.

The evidence available indicates that thyroid status can resolve to normal in some people whilst others may experience repeated fluctuations from normal to sub clinical status over a period of years. There are a number of underlying factors which appear to contribute to whether someone changes thyroid status such as age, disease aetiology and TSH at baseline.

Overall it is still unclear what proportion of people with sub clinical or overt thyroid dysfunction will revert to normal. Therefore this criterion is not met.

Criterion 5: The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.

The 2013 UK NSC review found there was a lack of consensus about the TSH cut-off value or values for defining which patients should receive treatment, and what constitutes a normal TSH level. The review reported that considerable variation existed in published studies between the levels of thyroid hormones that were considered abnormal. A review of cohort studies evaluating sub clinical thyroid dysfunction and mortality (Haentjens et al 2008)[22] showed variation in the cut off levels used to define sub clinical thyroid disease. The TSH cut off levels for sub clinical hypothyroidism varied between >4.0mIU/L and >5.10mIU/L. For sub clinical hyperthyroidism, the TSH cut off varied between <0.3mIU/L and <0.5mIU/L. These variations mean that a test undertaken in one laboratory may indicate the presence of disease, while the same result may be considered normal in a different setting.

Key question 2 of this review update is concerned with identifying new evidence which might provide clarity about the agreed cut off thresholds for both overt and sub clinical thyroid disease

Key question 2: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?

Sub-question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?

Description of the evidence

In the current review, of the 9 studies relevant to this criterion assessed at full text 5 papers; 1 prospective study and 4 retrospective studies were identified as potentially relevant to this question and were included (detailed in Appendices 7 to 11).

Reasons for exclusion of 4 of the 9 studies include:

- prevalence study
- clinical review of variation in thyroid function
- study of association of TSH and cardiovascular disease
- study of association of TSH and blood pressure and lipid levels

The prospective study included here partially meets the PICO's criteria but did not list the outcomes of sensitivity, specificity, false positive rate, false negative rate or positive and negative predictive values listed in the PICO. None of the retrospective studies met the PICO criteria but are included as they were all attempting to define a TSH reference range from a thyroid healthy population. These are derived from test results in laboratories or from large epidemiological studies. Each study applies a set of exclusion criteria which have common themes but are not the same (Appendices 7-11).

The table below summarises some of the study characteristics with more detail given in Appendices 7-11.

Table 3: Details of 5 studies investigating TSH, T4 and T3 reference range and cut offs for overt and sub clinical thyroid disease

Study	Population	Test	Comparator	Outcomes
Asvold et al 2012[23] Prospective study	22,206 aged 20-70 with: TSH =0.2-4.5mU/l T4=8-20pmol/l TPO=<200U/ml	TSH values grouped into 8 categories: 0.2-0.49,0.5-1.4,1.5-1.9,2.0-2.4,3.0-3.4,3.5-3.9,4.0-4.5mU/l	15,106 people available to test TSH, T4 and TPO between 9.4-12.8 years follow up	355(3.5%) of 10,083 women and 63(1.3%) of 5023 men developed hypothyroidism 96 (1.1) women and 26 (0.6%) men developed hyperthyroidism At baseline women with TSH within reference range of 0.5-1.4mU/l were less likely to develop hypothyroidism at follow up compared to those with a higher baseline (1.5-4.5mU/l) TSH values (p<0.001). The association in men was similar (p<0.001) but at any given level the risk of hypothyroidism was lower than in women. At baseline TSH of 0.5-1.4mU/l probability of becoming hypothyroid in women was 1.1% (95% CI 0.8-1.4) and men 0.3% (95% CI 0.1-0.6). At baseline of 4.0-4.5mU/l probability of becoming hypothyroid in women was 31.5% (95% CI 24.6-39.3) and in men was 14.7 (95% CI 7.7-26.2). In people with baseline TSH of 4.1-4.5 mU/l TPO antibodies were measured. People with TPO antibodies were at higher risk of hypothyroidism (women with TPO, 43.3%, 95% CI 29.5-58.1% and men, 21.5%, 95% CI 7.1-49.6%) compared to without (women 20.3%, 95%CI 12.1-31.9% and men 12.9%, 95% CI 5.4-27.8). The risk of hyperthyroidism was higher in women with baseline TSH of 0.20-0.49mU/l (3.9% 95% CI1.8-8.4%) than baseline of 0.5-0.99 mU/l (1.4% 95% CI 0.9-2.1%).
Coene et al 2015[24] Retrospective analysis	All TSH and free T4 test results for one year from 9 labs in Netherlands	Reference range used by laboratories not reported	5 platforms used	There was inter-laboratory variation in TSH/FT4 both within and between platforms. Beckman Coulter Access and Siemans Immulite users reported a lower prevalence of sub clinical hypothyroidism than Roche Cobas and Abbott Architect users (4.6%, 7.6%, 14.3% and 15.5% respectively). Differences between laboratories using the same platform were apparent eg: Roche Cobas in one laboratory reported 11.8% sub clinical hypothyroidism whilst another reported 17.8% and sub clinical hyperthyroidism ranged from 3.4% to 11.2%. Four labs switched platforms part way through the year and changes in prevalence were reported. Three labs moved to the Roche Cobas platform and there was an

				average increase in prevalence of sub clinical hypothyroidism of 6.2%.
Langen et al 2014[26]	Nationwide two stage cluster sample of 5709 thyroid healthy adults selected randomly from Finnish adult population.	TSH with 2.5 and 97.5 percentile reference range applied	3 sub groups of healthy thyroid population comprised of individuals with increasingly rigorous exclusion criteria	Thyroid healthy group n=5709 reference range: 0.41-4.43mU/l. Group 1:Risk factor free group n=4586 (excluded those with TPO antibodies) reference range -0.41-3.71mU/l. Group 2: No thyroid affecting medications n=3453 (excluded all those with any thyroid affecting medications and TPO anti bodies) reference range 0.41-3.60mU/l. Group 3: Reference group n=1849(excludes anyone taking any medication and TPO antibodies) 0.43-3.37mU/l.
Milinkovic et al 2013[25] Retrospective cross sectional analysis	22,860 lab results from outpatient adults>18 years	TSH, free T4 and T3 using Abbott architect and applying 2.5 and 97.5 percentiles to determine reference range	Gender Age in decades	Mean TSH values lower in 18-30 year old males compared to males >70 (95% CI 0.31-0.69p<0.0001). Mean FT4 concentration lower in males 18-30 years compared to 41-50 years (95% CI 0.29-0.74, p<0.0001). Mean TSH higher in females than males aged>70 years(p<0.05). Mean FT4 higher in males than females in 3rd and 4th decade (p<0.05). No difference in mean T3 by gender or age.
Vadiveloo et al 2013[28]	153,127 thyroid healthy population(based on medical history) from general population of Tayside	TSH	Age decade, gender and TSH group (group 1=<0.4mU/l, group 2=0.4-4.0mU/l and group 3 >4.0mU/l)	Increasing median TSH with increasing age(p<0.001) At 31-40 years proportion of people in TSH groups are 1 = 1.1%, group 2=96.4%, group 3=2.4% At >90 years proportion of people in TSH groups are 1=3.5%, group 2=86.2% and group 3=10.2%. Those with positive TPO antibodies had a 39% increase in TSH concentration compared to those without TPO antibodies and so a higher proportion were found in the TPO positive group 3 (26.2%) compared to the TPO negative group 3(15.9%). Median TSH different in men and women (p<0.001) but pattern of increase with age was similar.

Prospective study

Asvold et al (2012)[23] reported details of a longitudinal population based study in Norway (Nord-Trondelag Health Study or HUNT study). There were 15,106 participants (10,083 women and 5,023 men) who had a baseline serum TSH recorded in 1995-1997 with a follow up test in 2006-2008. People aged 20-70 with a self-report declaring no thyroid disease at baseline, supported by an absence of prescriptions or interventions (from health data) were included.

Results showed that higher serum TSH concentrations within the reference range at baseline were positively and strongly associated with the risk of developing hypothyroidism ($p<0.001$). Compared with women with a baseline TSH of 0.5-1.4mU/l the risk of sub clinical or overt hypothyroidism at follow up (mean 11 years, range 9.4-12.8) was 5 fold higher in women with TSH of 1.5-1.9mU/l at baseline, 18 fold higher in women with TSH of 2.5-2.9mU/l at baseline and 42 fold higher in women with TSH of 4-4.5mU/l. In men there was a similar association but at any given level the risk of hypothyroidism was lower in men with a 21 fold risk at 4 to 4.5mU/l than at baseline.

TSH at lower levels of the reference range at baseline may be associated with higher risk of hyperthyroidism at follow up. This association was detected for women but there were too few men to carry out a similar analysis. Women with TSH between 0.20-0.49 at baseline had a four fold chance of being diagnosed with hyperthyroidism at 11 years follow up.

There was no cut-off that distinctly separated TSH levels that are associated with increased risk of sub clinical and overt hypothyroidism and those that were not. A substantial proportion (34-50% women and 13- 34% men) of people with baseline TSH in the upper end of the reference range (TSH 4.0-4.5 mU/l) are predicted to develop sub clinical or overt hypothyroidism (Appendix 7).

Retrospective studies

Three papers (Langen et al 2014, Milinkovic et al 2014, Vadiveloo et al 2013) reported retrospective studies of large numbers of test results or carried out tests on blood samples previously taken and stored [24;25;26]. Details of these studies are in appendices 9-11. These retrospective studies all attempted to derive a set of reference ranges from a thyroid healthy sub group. Two studies showed significant variation in TSH upper reference limits with age and gender. One study showed variation in TSH and FT4 reference ranges by age and gender but not FT3.

One study split their thyroid healthy population into 4 groups with increasingly rigorous exclusion criteria then applied 2.5 and 97.5 percentiles to derive a reference range. The table below shows the groups, exclusion criteria and reference range for each population. As more exclusion criteria are applied to the groups the TSH upper reference level drops from 4.43mU/l to 3.37mU/l (25%). The same effect is not observed in the TSH lower reference level that remains the same in 3 of the groups and only increases by 0.02mU/l (5%) from 0.41mU/l to 0.43mU/l in the reference group. The authors did apply the majority of the National Academy of Clinical Biochemistry guidelines[27] for defining TSH reference ranges in the thyroid healthy population but were unable to determine a family history of thyroid dysfunction or the thyroglobulin antibody concentrations (whose presence can indicate thyroid dysfunction).

Table 4: TSH reference ranges in thyroid healthy populations with increasingly rigorous exclusion criteria

Groups	Exclusions	TSH Reference range-at 2.5and 97.5 percentiles
Thyroid healthy group (n=5709)	Non ambulatory, history of thyroid disease or goitre, blood sample drawn outside 8am-6pm, using thyroid hormones or anti thyroid agents, pregnant or breast feeding, those with extreme TSH levels	0.41-4.43mU/l
Risk factor free group(n=4586)	All the above exclusions plus individuals testing positive for TPO antibodies	0.41-3.71mU/l
No thyroid medication group(n=3453)	All the above exclusions plus any individuals taking any thyroid effecting medication	0.41-3.60mU/l
Reference group(n=1849)	All the above exclusions and any individuals taking any medication.	0.43-3.37mU/l

A fourth retrospective study by Coene et al (2015)[28] showed variation in TSH reference ranges between 5 different testing platforms used in 9 laboratories (details in appendix 8). This resulted in variation in the prevalence of sub clinical thyroid disease. All test results for TSH over the period of a year were collected from the laboratories and their reference ranges plotted. There was inter-laboratory variation in TSH/FT4 both within and between platforms. Beckman Coulter Access and Siemans Immulite users reported a lower prevalence of sub clinical hypothyroidism than Roche Cobas and Abbott Architect users (4.6%, 7.6%, 14.3% and 15.5% respectively).

Differences between laboratories using the same platform were apparent. For example Roche Cobas in one laboratory reported 11.8% sub clinical hypothyroidism whilst another reported 17.8% and sub clinical hyperthyroidism ranged from 3.4% to 11.2%.

Four laboratories changed platforms part way through the year which enabled the study to exclude natural differences in prevalence in the population. Changes in prevalence of the different thyroid conditions were reported after the equipment changed. Three laboratories moved to the Roche Cobas platform and there was a mean increase in prevalence of sub clinical hypothyroidism of 6.2%.

Discussion

QUADAS-2 is a framework designed to assess the quality of primary diagnostic accuracy studies. This evidence update has not applied the QUADAS-2 framework to the studies included for this key question as they are not comparing a particular screening test with the reference standard, but rather carrying out the same test across populations and reporting the upper and lower thresholds by age, gender and other factors. The QUADAS-2 framework with its focus on bias, blinding, applicability, flow and timing cannot be readily applied to the types of studies included here which are attempting to define a reference standard.

This review update has identified one prospective study and four retrospective studies concerned with TSH reference ranges and cut off levels for sub clinical and overt thyroid disease.

None of the studies focussed on the cut off levels of free T4 or T3 although one study[25] did show that age and gender impacted levels of FT4 but not T3.

The prospective study compared baseline TSH at the beginning of the study and the association between baseline values at the upper or lower end of a predetermined reference range and the subsequent likelihood of developing thyroid disease 11 years later. The results indicated that the risk of hypothyroidism is gradual across the reference range with no cut-off point that distinctly separates TSH levels associated with increased risk and TSH levels that are not associated with increased risk. However the authors suggest that as an increasing proportion of people with TSH in the uppermost part of the reference range will go on to develop hypothyroidism and those with TSH between 2.5-4.5mU/l may benefit from follow up testing. At the lower end of the reference range a much more limited effect was seen. Women showed a four fold increase in the likelihood of developing hyperthyroidism at 0.20-0.49mU/l but there were too few men to calculate the increased risk.

Four further studies illustrated the variation in TSH and age, gender, laboratory and platform used. None of the retrospective studies met the PICO criteria but are included as they were all attempting to define a TSH reference range from a thyroid healthy population. These are derived from test results in laboratories or from large epidemiological studies. Each study applied a set of exclusion criteria which have common themes but are not the same and reported slightly different reference ranges when applying 2.5th and 97.5th percentiles. The group comprising individuals with the most rigorous exclusion criteria[26], determined a normal reference range of 0.43-3.37mU/l which is not consistent with either the current recommendations by the British Thyroid Society of 0.4mU/l-4.0mU/l[3] or the US National Academy of Clinical Biochemistry[27] who predicted that the upper TSH reference level was likely to be 2.5mU/l in a population applying the most rigorous exclusion criteria.

One Dutch study illustrated the impact of the variation in the different reference ranges recommended by different platforms within laboratories. Where there was a change in platform in a particular laboratory there was a clear change in prevalence of sub clinical dysfunction for the population it served, leading to an arbitrary change in the number of people prescribed thyroid hormones. At the level of 10mU/l (a typical level for pharmaceutical intervention), Roche cobas users reported an average of 2.5% sub clinical hypothyroidism while for Beckman Coulter and Siemans Immulite users this was 1.3 %. This equates to an extra 300 patients per Roche laboratory who are likely to be receiving FT4 substitution medication.

Efforts to standardise TSH and FT4 are being undertaken by the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). The committee was established in light of the prevalence of thyroid disease, the frequency of laboratory testing, and multiple reports on discrepant measurement results. The IFCC Committee for Standardisation of Thyroid Function Tests aims at equivalence of laboratory test results for FT4 and TSH. Their latest publication reports that standardization and harmonization of FT4 and TSH measurements is feasible from a technical point of view but implementation would need careful preparation with stakeholders.[29]

Summary: Criterion 4: not met

Overall the literature search of publications since the previous review did not identify any new studies that lead to a consensus of the reference ranges and cut-off thresholds of TSH, FT4 or FT3 as a whole population screening test for sub clinical or overt thyroid disease.

A TSH result of 2.5 to 4.5mU/l has been suggested as an upper range which might define sub clinical hypothyroidism. A threshold for determining risk of sub clinical hyperthyroidism disease was not proposed.

No studies focused on the cut off thresholds of FT4 or FT3.

The TSH reference range across testing platforms have been shown to vary significantly thus impacting on prevalence and the number of people identified as eligible for treatment There are significant efforts to attempt to standardize references ranges in the future by the IFCC.

This criterion is not met.

Criterion 9: There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered.

Key question 3: what is the effectiveness of treatment of overt and sub clinical thyroid disease?

Overt thyroid disease

The 2013 UK NSC review referenced the standard treatment for people diagnosed with overt thyroid disease. Treatment of overt hypothyroidism requires the use of levothyroxine (British National Formulary 2017)[30]. The goal of treatment is to restore and maintain serum TSH within the reference range[4]. Prolonged periods of over treatment with thyroxine can lead to symptoms of hyperthyroidism with associated features of atrial fibrillation and osteoporosis.

Options for the treatment for overt hyperthyroidism include the anti-thyroid drug carbimazole, radio-iodine, and surgical removal of part of the thyroid gland[30]. One of the main risks of treatment for overt hyperthyroidism is the development of hypothyroidism[4].

It is not clear if there would be any difference in the treatment outcomes of people with screen detected overt thyroid disease versus clinically detected overt disease.

Sub clinical thyroid disease

The 2013 UK NSC review identified 2 systematic reviews of RCTs on treatment of clinically detected sub clinical thyroid disease and concluded that there was insufficient evidence of benefit and that the potential adverse effects from treatment had not been adequately studied. The apparent short-term (four month) benefits of treating screen-detected subclinical hypothyroidism reported in 15 subjects by Abu-Helalah et al (2010)[31] was limited and the results needed to be confirmed in a larger study with longer follow-up.

Description of the evidence

Key question 3 focuses on evidence about treatment for screen detected sub clinical and overt thyroid disease. The PICO prioritises RCTs and systematic reviews especially where long term outcomes and harms are reported.

Of the 9 studies that were reviewed as full texts relevant to this question 2 papers were subsequently included in this review (details in appendices 12 and 13). During review of the 9 full papers there was reference to an RCT about treatment of sub clinical hypothyroidism in people aged 65 years and over expected to be published in June 2016. The trial results were published in April 2017 after the search date for the literature search for this review. However, on checking that this study met the PICO criteria it was included as part of this review (details in Appendix 14).

Reasons for excluding 7 of the 9 studies identified in the literature search following full text review were either that they were not RCTs, systematic reviews, studies reporting long term outcomes, harms of treatment in screen detected populations or populations with undiagnosed thyroid disease.

The 3 papers included are:

- a retrospective cohort study of 4735 people from the UK, either treated or untreated with levothyroxine for sub clinical hypothyroidism to investigate the association between treatment and non-fatal and fatal cardiovascular events (Razvi et al 2012)[32].
- an evidence review by the US Agency for Healthcare Research and Quality (Rugge et al 2015)[33] evaluating the effectiveness and treatment of overt and sub clinical thyroid dysfunction.
- An RCT of 737 adults ≥65 years with persistent sub clinical hypothyroidism treated with levothyroxine or placebo (Stott et al 2017)[34].

Retrospective cohort study by Razvi et al (2012)[32]

This study retrospectively identified patients with sub clinical hypothyroidism from case records from the UK GP Research Database and analysed outcomes according to subsequent treatment with levothyroxine (Appendix 12). People with TSH outside the normal reference range and FT4 within the normal reference range who had been tested during 2001 with outcomes analysed up to March 2009 were included. Individuals were classed as having sub clinical hypothyroidism if they had recorded TSH serum values of 5.01 to 10 mU/l and normal FT4 values (0.7-1.9ng/dl). These ranges were applied uniformly to overcome differences between different biochemical assays with different reference ranges.

Analyses were performed separately for the lower age group (40 to 70 years) and older age group (>70 years). Median follow up was 7.6 years (0 to 8) in the 40 to 70 age group and 5.2 years (0 to 8) in those 70 and over.

Table 5 below shows that treatment with levothyroxine was associated with reduced all cause mortality, fewer Ischemic Heart Disease (IHD) events, fewer deaths due to circulatory diseases and fewer deaths due to cancer in individuals aged 40 to 70 but not in those who were older.

Table 5: Outcomes of treatment of sub clinical hypothyroidism by age group

	40 to 70 years			**>70 years**		
	Untreated	**Treated**	**Hazard ratio (95% CI)**	**Untreated**	**Treated**	**Hazard ratio (95% CI)**
Patients	1459	1634		823	819	
Fatal and non fatal IHD events	97(6.6%)	68(4.2%)	0.61(0.39-0.95)*	88(10.7%)	104(12.7%)	0.99(0.59-1.33)
All cause mortality	94(6.4%)	55(3.4%)	0.36(0.19-0.66)*	333(40.5%)	288(35.2%)	0.71(0.56-1.08)
Death due to	38(2.4%)	23(1.4%)	0.54(0.37-0.92)*	116(18.3%)	134(17.1%)	0.91(0.56-1.46)

circ. diseases						
Death due to IHD events	27(1.7%)	17(1%)	0.43(0.19-2.05)	70(6.3%)	56(5.5%)	1.04(0.56-1.93)
Death due to malignant neoplasms	35(2.2%)	21(1.2%)	0.59(0.21-0.88)*	73(6.5%)	49(4.6%)	0.51(0.24-1.04)
Fatal and nonfatal CVA	44(3%)	55(3.4%)	1.03(0.51-2.13)	147(17.9%)	145(17.7%)	0.81(0.31-2.12)
Atrial fibrillation	36(2.3%)	35(2%)	0.87(0.59-1.44)	87(7.7%)	86(8.1%)	0.98(0.54-1.76)

Source: Razvi 2012[32] *indicates significantly positive association with treatment

Evidence review by Rugge et al 2015[33]

This evidence review is focused on identifying evidence about the screening and treatment of thyroid dysfunction. It builds on and incorporates the evidence review of screening for sub clinical thyroid dysfunction by Rugge et al in 2011[11]. All but one of the additional references were reported in the UK NSC update in 2013[13]. This was the study by Razvi (2012)[21] which is summarised above.

Rugge et al's[33] literature search was carried out between 2002 to July 2014 and their review concluded that there were:

- no studies comparing clinical benefits or harms in people screened versus not screened for thyroid dysfunction.
- no studies addressing treatment versus no treatment of screen detected undiagnosed overt hypothyroidism.

Treatment of older adults with sub clinical hypothyroidism (Stott at al 2017)[34]

The trial reported by Stott et al (2017)[34] aimed to determine whether levothyroxine provides clinical benefits in older people with sub clinical hypothyroidism. This was a double blind, randomised, placebo controlled, parallel group, multi centre trial involving 737 adults who were at least 65 years of age and who had persisting sub clinical hypothyroidism with TSH between 4.60-19.99mU/l measured at least twice, three months apart. The active intervention was levothyroxine dose of 50µg daily (or 25µg in patients with a body weight<50 kg). Dose adjustment aimed at maintaining a TSH reference range of 0.40-4.59mU/l. Mock adjustment of placebo dosage was also carried out.

Primary outcomes were differences in Hypothyroid Symptoms and Tiredness scores from the ThyPRO scale carried out at baseline and one year in the placebo and treatment groups. Overall the trial reported:

- at 12 months there was no significant difference in the mean hypothyroid symptom score of 16.7±17.5 in the placebo group and 16.6±16.9 in the intervention group (p=0.99).
- there was no significant difference in the mean tiredness score of 28.6±19.5 in the placebo group and 28.7±20.2 in the levothyroxine group (p=0.77).

- there were no differences in the mean change in the hypothyroid symptoms score and the mean tiredness score at 1 year from baseline between the placebo and levothyroxine group.
- there was a non-significant between group difference (p=0.05) in the Tiredness score with a lower value in the levothyroxine group at extended follow up (mean extended follow up 24 months).
- the number of patients with at least one serious adverse event was significantly higher in the placebo group than in the levothyroxine group (p=0.049) but there was no pattern of event type that contributed to this difference.

Discussion

Three papers were identified that were considered relevant to treatment for screen detected cases of sub clinical and overt undiagnosed thyroid dysfunction. One randomized controlled trial and one retrospective study were wholly or partly carried out in the UK focused on treatment of sub clinical hypothyroidism.

Neither the evidence review by Rugge (2015)[33Error! Bookmark not defined.] nor our own search strategy identified trials focused on treatment of screen detected sub clinical hyperthyroidism or undiagnosed overt thyroid dysfunction.

The retrospective study showed that people with sub clinical hypothyroidism between 40 to70 years old had some cardiovascular benefit from being treated with levothyroxine. The same differences were not seen in the older age group (>70 years). The study was observational and may have included people with a prior history of heart disease, despite trying to control for cardiovascular risk factors. This could lead to an unknown level of residual confounding which may impact the results. However in support of some of the cohort study's findings similar results were reported by the double blind, randomized, placebo controlled trial that found no significant excess of serious adverse events in either treated or untreated groups for people aged 65 and over. The RCT also reported that there was no difference in the change in ThyPRO measures for Hypothyroid Symptoms and Tiredness scores between baseline and 12 months for people receiving either levothyroxine or a placebo aged 65 and over.

Summary: Criterion 9- not met

Overall, there is a lack of evidence that people with screen detected sub clinical and overt thyroid disease will benefit from treatment and no new high quality evidence identified considering treatment for undiagnosed overt thyroid dysfunction or sub clinical hyperthyroidism.

Conclusions and implications for policy

This report assesses systematic population screening for thyroid dysfunction against select UK NSC criteria for appraising the viability, effectiveness and appropriateness of a screening programme. This review assessed key questions to determine if new evidence published since January 2011 supports a recommendation for screening for thyroid dysfunction in adults in the UK.

The volume, quality and direction of new evidence published since January 2011 does not indicate that systematic population screening for thyroid dysfunction in adults should be recommended in the UK.

No studies were identified reporting outcomes from people with screen detected thyroid disease. This is important because studies reporting data from populations who have presented to the health service or volunteered for a study may have different characteristics to those taking up the offer of population screening. In addition to a lack of outcomes from screened populations there are uncertainties around the evidence available for key criteria.

Key question 1: what proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?

This question is important because in the context of screening the aim of a programme would be to detect abnormality either due to overt or sub clinical disease. At the point of screening it is important to know whether an abnormal result is likely to be a temporary condition or progress to overt disease. This fluctuation in sub clinical status to normal and back again was flagged in the 2013 UK NSC review and this uncertainty has led to the inclusion of Key question 1 in this review.

The evidence identified as relevant for this question was limited to people with sub clinical hyperthyroidism and one study of an elderly cohort for sub clinical hypothyroidism. No evidence for the proportion of people whose thyroid function normalises following overt disease or the wider adult population for sub clinical hypothyroidism was identified.

The studies suggest there are a number of underlying reasons that can have an effect on the thyroid hormones such as age, sex, type of thyroid disease and the presence of antibodies that mistakenly attack the thyroid gland. However it is not possible to determine an obvious group of people whose thyroid status is more or less likely to revert to normal.

The evidence indicated that thyroid status can resolve to normal in some people whilst others may experience repeated fluctuations from normal to sub clinical status overtime.

The 5 studies tracking people with sub clinical hyperthyroidism and change in thyroid status reported between 31.6% and 53% patients with TSH reverting to within normal reference range values over a period of between 2 months to 9 years. The evidence suggests there are multiple underlying factors that can impact thyroid status change such as age, disease aetiology and baseline TSH values.

The one sub clinical hypothyroidism study focused on an older population (>65 years) and reported that at 2 years 35% of people with sub clinical hypothyroidism at baseline had reverted to normal however at 4 years this had dropped to 17%. This indicates that some people had fluctuated from sub clinical status at baseline to normal at 2 years and back again at 4 years at least once.

Key question 2: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?

Sub-question: is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?

In order for a screening programme to detect as far as possible the correct group of people who will go on to be diagnosed with a condition the screening test needs to be robust with clear cut off levels that will determine who does and who does not have the condition. The 2013 UK NSC

review identified that no optimum cut off levels for TSH, FT4 and FT3 had been agreed for sub clinical and overt disease.

The current review found 5 publications that described methods to determine the TSH reference range in groups of people in different ways.

- One study looked at baseline TSH levels and reported progression to overt disease over 11 years. They determined a sub clinical hypothyroid cut off level (2.5-4.5mU/l) where the risk of progression steeply rose whilst the overt cut off level was predetermined by the testing platform manufacturers. The sub clinical hyperthyroid cut off levels were not reported.
- One study carried out TSH tests on a population of people and then using increasingly strict exclusion criteria determined the TSH range for the sub group with no possible factors that would impact their TSH result. The standard method of applying 2.5th percentile cut offs at either end of the mean range of the results of this group was used to determine the TSH reference range (0.43-3.37mU/l).
- Two other retrospective studies looked at the impact of different factors (eg age and gender) on TSH reference ranges using the standard method of applying 2.5th percentiles cut off thresholds.
- The 5th study showed the variability in reference ranges across different TSH testing platforms and laboratories. Where testing platforms had been changed this resulted in a change in the prevalence of thyroid dysfunction. There was a change of 6.5% in the prevalence of sub clinical hypothyroidism across three laboratories who changed platforms.

Overall the literature search of publications since the previous review did not identify any studies that led to a consensus of the reference ranges and thresholds of TSH, T4 or T3 as a whole population screening test for sub clinical or overt thyroid disease.

Key question3: what is the effectiveness of treatment of overt and sub clinical thyroid disease?

There is standard effective treatment for overt thyroid disease in patients with a clinically diagnosed condition. This was reported in the 2013 UK NSC review.

There is a dearth of evidence for the effectiveness of treatment of people with either clinically diagnosed or screen detected sub clinical thyroid disease. This review found one retrospective study providing some evidence that levothyroxine may confer cardiovascular benefits on younger patients (40-70 years) diagnosed with sub clinical hypothyroidism. However the limitations of this type of study lead to the possibility of residual confounding by participants with unknown cardiovascular disease status. One RCT in older people did not find that levothyroxine improved hypothyroid symptoms or tiredness.

Overall, there is a lack of evidence that people with screen detected sub clinical and overt thyroid disease will benefit from treatment and no new high quality evidence identified considering treatment for undiagnosed overt thyroid dysfunction or sub clinical hyperthyroidism.

Recommendation

This update of the evidence found that no new evidence has been published to change the conclusions of the previous UK NSC review therefore population screening for thyroid disease is not recommended.

Limitations

This rapid review process was conducted over a condensed period of time (~12 weeks). Searching was limited to 3 bibliographic databases and did not include grey literature sources.

The rapid review was guided by a protocol developed *a priori*. Literature search and first pass appraisal were predominantly undertaken by one UK NSC information scientist, and second pass appraisal and study selection by one reviewer, though any queries at both stages were resolved through discussion with a second reviewer, or with the UK NSC evidence team.

Studies available only in non-English language, abstracts, conference reports or poster presentations were not included. Study authors were not contacted and studies that were not published in peer-reviewed journals were not reviewed.

Search strategy

Literature search on screening for thyroid dysfunction in non-pregnant adults
Paula Coles, Information Scientist
January 2017

SCOPE OF THE SEARCH:
Question 1: What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical interventions?
Question 2: Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?
Question 3: What is the effectiveness of treatment of overt and sub clinical thyroid disease?
SOURCES SEARCHED: Medline, Embase, and the Cochrane Library.
DATES OF SEARCH: January 2011 – January 2017
SEARCH STRATEGIES:
Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Natural history

1 Hypothyroidism/ (27791)

2 Hyperthyroidism/ (27507)

3 ((overactive or underactive) adj thyroid).tw. (33)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw. (13328)

5 1 or 2 or 3 or 4 (58922)

6 disease progression/ (149071)

7 Remission, Spontaneous/ (17213)

8 Health Status/ (78777)

9 (natural adj (history or course)).tw. (53857)

10 thyroid status.tw. (2240)

11 (spontaneous adj (course or resolution or remission)).tw. (6531)

12 6 or 7 or 8 or 9 or 10 or 11 (299198)

13 5 and 12 (1790)

14 limit 13 to yr="2011 -Current" (356)

****

Test

1 Hypothyroidism/ (27791)

2 Hyperthyroidism/ (27507)

3 ((overactive or underactive) adj thyroid).tw. (33)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw. (13328)

5 1 or 2 or 3 or 4 (58922)

6 Thyroid Function Tests/ (15180)

7 thyroid function test$.tw. (3332)

8 Immunoassay/ (28393)

9 immunometric assay.tw. (543)

10 thyrotrophin assay.tw. (8)

11 6 or 7 or 8 or 9 or 10 (45674)

12 "Sensitivity and Specificity"/ (346379)

13 (Sensitive$ or specific$).tw. (3459042)

14 Predictive value of tests/ (192671)

15 ((Positive or negative) adj predictive value$).tw. (56870)

16 (PPV or NPV).tw. (15664)

17 ((False or true) adj (negative$ or positive$)).tw. (76598)

18 12 or 13 or 14 or 15 or 16 or 17 (3820626)

19 thyroid status.tw. (2240)

20 (Thyroid stimulating hormone or TSH).tw. (34476)

21 (Thyroxine or T4 or free T4 or FT4).tw. (64507)

22 (tri-iodothyronine or T3 or triiodothyronine or free T3 or FT3).tw. (53231)

23 serum thyrotropin.tw. (1060)

24 antithyroid antibod$.tw. (986)

25 19 or 20 or 21 or 22 or 23 or 24 (112009)

26 Mass Screening/ (101056)

27 (screen$3 or detect$3 or test or tests or testing).tw. (4163801)

28 26 or 27 (4184787)

29 5 and 11 and 18 and 25 (520)

30 5 and 11 and 25 and 28 (1743)

31 screen$3.ti. (159758)

32 (hypothyroidism or hyperthyroidism or thyroid dysfunction or thyroid disease or thyroid deficiency or thyroid failure).ti. (26469)

33 31 and 32 (919)

34 Reference Values/ (166053)

35 ((reference or normal) adj rang$).tw. (50621)

36 cut-off value$.tw. (15188)

37 34 or 35 or 36 (224528)

38 5 and 25 and 37 (1832)

39 29 or 30 or 33 or 38 (4216)

40 limit 39 to yr="2011 -Current" (909)

****

Treatment

1 Hypothyroidism/ (27791)

2 Hyperthyroidism/ (27507)

3 ((overactive or underactive) adj thyroid).tw. (33)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw. (13328)

5 1 or 2 or 3 or 4 (58922)

6 Hormone Replacement Therapy/ (9820)

7 thyroid hormone replacement.tw. (871)

8 (levothyroxine or L-thyroxine or thyroxin$).tw. (32823)

9 Carbimazole/ (1138)

10 carbimazole.tw. (883)

11 (radio-iodine or radioactive iodine or I-131 or RAI).tw. (11793)

12 (surgery or surgical intervention$).tw. (1038840)

13 6 or 7 or 8 or 9 or 10 or 11 or 12 (1091061)

14 Treatment outcome/ (864604)

15 Mortality/ (42364)

16 Morbidity/ (29635)

17 Health Status Indicators/ (24590)

18 (health related quality of life or quality of life).tw. (226200)

19 (QoL or HRQoL).tw. (42371)

20 (questionnaire$ or survey$ or index$).tw. (1523779)

21 ((physical or psychological) adj status).tw. (7863)

22 all cause mortality.tw. (25858)

23 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (2507687)

24 ((treat$ or manage$ or therap$) and (hypothyroidism or hyperthyroidism)).ti. (4471)

25 5 and 13 and 23 (2058)

26 24 or 25 (6209) 27 limit

26 to yr="2011 -Current" (1168)

SIGN filters for RCTs and systematic reviews (http://www.sign.ac.uk/methodology/filters.html) applied (290)

****

Natural history or test or treatment (1424)

Embase 1996 to 2017 Week 03 19 January 2017

Natural history

1 hyperthyroidism/ (18213)

2 hypothyroidism/ (35627)

3 ((overactive or underactive) adj thyroid).tw. (42)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw. (12748)

5 1 or 2 or 3 or 4 (28278)

6 disease course/ (390564)

7 remission/ (130957)

8 health status/ (103210)

9 (natural adj (history or course)).tw. (49274)

10 thyroid status.tw. (1423)

11 (spontaneous adj (course or resolution or remission)).tw. (5672)

12 6 or 7 or 8 or 9 or 10 or 11 (650901)

13 5 and 12 (2545)

14 limit 13 to yr="2011 -Current" (1186)

****

Test cut-offs and reference range

1 hyperthyroidism/ (18213)

2 hypothyroidism/(35627)

3 ((overactive or underactive) adj thyroid).tw.(42)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw.(12748)

5 1 or 2 or 3 or 4 (28278)

6 thyroid function test/ (8190)

7 thyroid function test$.tw. (3791)

8 immunoassay/ (53169)

9 immunometric assay.tw.548

10 thyrotrophin assay.tw.2

11 6 or 7 or 8 or 9 or 10 (62003)

12 "sensitivity and specificity"/267362

13 (Sensitive$ or specific$).tw.2881966

14 predictive value/112635

15 ((Positive or negative) adj predictive value$).tw.67078

16 (PPV or NPV).tw.24476

17 ((False or true) adj (negative$ or positive$)).tw.68397

18 12 or 13 or 14 or 15 or 16 or 173083263

19 thyroid status.tw.142

20 (Thyroid stimulating hormone or TSH).tw.28778

21 (Thyroxine or T4 or free T4 or FT4).tw.46834

22 (tri-iodothyronine or T3 or triiodothyronine or free T3 or FT3).tw.44994

23 serum thyrotropin.tw.626

24 antithyroid antibod$.tw.871

25 19 or 20 or 21 or 22 or 23 or 24 (84034)

26 mass screening/ (34981)

27 (screen$3 or detect$3 or test or tests or testing).tw.3871946

28 26 or 27 (3879226)

29 5 and 11 and 18 and 25 (320)

30 5 and 11 and 25 and 28 (1281)

31 screen$3.ti.146369

32 (hypothyroidism or hyperthyroidism or thyroid dysfunction or thyroid disease or thyroid deficiency or thyroid failure).ti. (15386)

33 31 and 32 (631)

34 Reference Values/44067

35 ((reference or normal) adj rang$).tw.48500

36 cut-off value$.tw.24102

37 34 or 35 or 36 (112202)

38 5 and 25 and 37 (1109)

39 29 or 30 or 33 or 38 (2799)

40 limit 39 to yr="2011 -Current" (1556)

****

Treatment

1 hyperthyroidism/ (18213)

2 hyperthyroidism/ (35627)

3 ((overactive or underactive) adj thyroid).tw. (42)

4 (thyroid adj (dysfunction or disease or deficiency or failure)).tw. (12748)

5 1 or 2 or 3 or 4 (55289)

6 hormone substitution/ (33051)

7 thyroid hormone replacement.tw. (841)

8 (levothyroxine or L-thyroxine or thyroxin$).tw. (19052)

9 carbimazole/ (2180)

10 carbimazole.tw. (595)

11 (radio-iodine or radioactive iodine or I-131 or RAI).tw. (9240)

12 (surgery or surgical intervention$).tw. 1040082

13 6 or 7 or 8 or 9 or 10 or 11 or 12 (1097008)

14 treatment outcome/ (728685)

15 mortality/ (685770)

16 morbidity/ (349632)

17 health status indicator/ (1902)

18 (health related quality of life or quality of life).tw. (292862)

19 (QoL or HRQoL).tw. (64587)

20 (questionnaire$ or survey$ or index$).tw. (1558707)

21 ((physical or psychological) adj status).tw. (8715)

22 all cause mortality.tw. (33403)

23 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 (3061442)

24 ((treat$ or manage$ or therap$) and (hypothyroidism or hyperthyroidism)).ti. (2434)

25 5 and 13 and 23 (3411)

26 24 or 25 (5490)

27 limit 26 to yr="2011 -Current" (2538)

SIGN systematic review and RCT filters (http://www.sign.ac.uk/methodology/filters.html) applied (446)

****

Natural history or test or treatment (2983)

#1 MeSH descriptor: [Hypothyroidism] this term only (335)

#2 MeSH descriptor: [Hyperthyroidism] this term only (282)

#3 ("overactive thyroid" or "over active thyroid" or "underactive thyroid" or "under active thyroid"):ti,ab,kw (2)

#4 ("thyroid dysfunction" or "thyroid disease" or "thyroid deficiency" or "thyroid failure"):ti,ab,kw (392)

#5 #1 or #2 or #3 or #4 Publication Year from 2011 to 2017 (296)

Table 4: All searches carried out on 19 January 2017

Database	No. of citations retrieved
Medline	**1424**
Embase	**2983**
Cochrane Library	**296**
Total	**4703**

After automatic and manual de-duplication, 3463 unique references were sifted for relevance to the review.

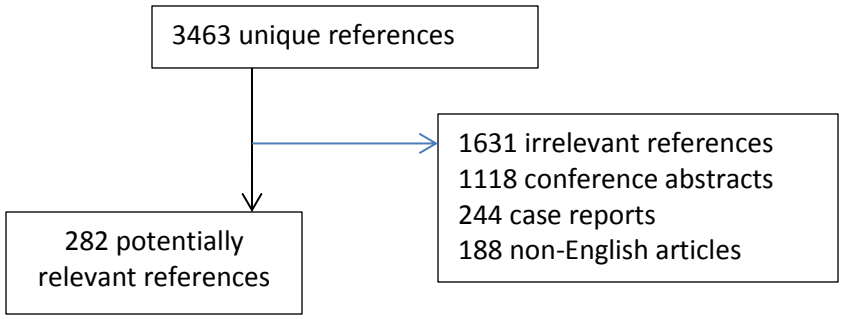
Inclusions

• Non-pregnant, general populations
• Studies based on UK populations, or comparable populations in Europe, USA, Canada, Australia and New Zealand.
• Any type of study deemed to be relevant to any of the three questions. (At this stage of sifting, only study types mentioned in the exclusions below have been removed.)

Exclusions

• Animal studies, case reports, non-English publications, animal studies, conference abstracts and comments
• Pregnant, neonatal or child populations
• High risk populations

Figure 4: Flow diagram summarising the results of the reference sifting process



These 282 references were classified as presented in Table 5

282 references were deemed to be potentially relevant
These references are categorised below. There may be some overlap between categories.
No judgement on quality or bias has been made at this stage.

After 940 duplicates were removed, 3463 unique references were sifted by title and abstract, and where necessary and available the full text, for potential relevance to the review. 282 papers remained and were passed to the SPH reviewer for further consideration.

These 282 references were classified as presented in Table 6

Table 6: Summary of the relevant references by category

Category	No. of citations
Screening and/or treatment systematic reviews	20
Guidelines and recommendations	11
Reversion to normal	8
Test *Reference ranges (32)* *Reference ranges – prediction of morbidity (26)* *Test (30)* Incidence/prevalence of condition	**88**
Treatment Treatment for hypothyroidism (44) Levothyroxine (24) Combination of treatments (10) Other treatments (4) Non-systematic reviews (6) Treatment for sub clinical hypothyroidism (27) Treatment for hyperthyroidism (47) Radioactive iodine (22) Surgery (7) Methimazole (3) Combination of treatments (6) Other treatments (3) Non-systematic reviews (6) Treatment for sub clinical hypothyroidism (8) Treatment for thyroid dysfunction (13)	**139**
Overviews	**16**
Total	**282**

Key question PICOS

Question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Comments	Includes both hyper and hypothyroidism Studies of screen detected populations should be prioritised if available Results should be reported by duration of follow up
Population	Non pregnant adults
Intervention	N/a
Comparator	N/a
Outcomes	Continued presence or resolution of thyroid dysfunction

Question	Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction?
Sub-questions	Is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
Comments	Includes hyper and hypothyroidism Report outcomes by age The previous review found an insufficient volume of prospective studies to be confident that reported sensitivity and specificity values would be replicated in a screening programme setting. None were identified for overt and sub clinical hyperthyroidism or overt hypothyroidism. One small prospective study was identified for sub clinical hypothyroidism but required larger studies to explore the hypothesis further and the methods by which the test values were derived was considered controversial Prospective studies of consecutively enrolled subjects should be prioritised
Population	Non-pregnant adults
Intervention	Thyroid function test (TSH, T4 and T3)
Comparator	Open
Outcomes	Studies reporting clinical performance measures and SRs of these: • Sensitivity • Specificity • False positive rate • False negative rate • PPV/ NPV

Question	What is the effectiveness of treatment of overt and sub clinical thyroid disease?
Comments	RCTs, systematic reviews and other studies of screen detected sub clinical disease should be prioritised if available

	Studies of long term outcomes should be prioritised Harms of treatment should be reported if described
Population	Non-pregnant adults
Intervention	Standard care for overt and sub clinical dysfunction, for example: • For hypothyroidism: levothyroxine • For hyperthyroidism: carbimazole; radio-iodine; surgical intervention
Comparator	Placebo, no treatment N/a if observational
Outcomes	TSH levels Cognitive function Measures of cardiovascular morbidity (eg blood pressure, BMI, cholesterol) Patient reported physical and psychological health status (eg via commonly used thyroid disease indexes, general health questionnaires etc)

Appendix Tables

Appendix number	1
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Zhyzhneuskaya S, Addison C, Tsatlidis V, Weaver JU, Razvi S. The natural history of sub clinical hyperthyroidism in graves' disease: The rule of thirds. Thyroid. 2016;26(6):765-9**
Study details	Prospective cohort study
Study objectives	To assess the natural progression of a group of patients with sub clinical hyperthyroidism due to Graves Disease.
Inclusions	Consecutive patients with sustained sub clinical hyperthyroidism (minimum of 2 occasions at least 3 months apart) referred to secondary care endocrinology clinic between August 2002 and December 2013.
Exclusions	Individuals on medication that could affect thyroid function such as levothyroxine, anti thyroid drugs amiodarone and lithium were excluded.
Population	44 patients with sub clinical hyperthyroidism due to Graves Disease

Intervention/ test	Patients tests: TSH, free thyroxine (T4), free triiodothyronine(T3) thyroid peroxidase antibody(TPO), TSH receptor antibody (TRAb) and thyroid uptake scan. In the 1st year patients assessed at 3, 6 and 12 months then every 6 months. Median follow-up was 32 months (2-111).
Comparator	N/A
Results	Of the 44 patients 13(30%) continued to have sub clinical hyperthyroidism, 15 (34%) progressed to the overt condition and 15(34%) patients reverted to normal thyroid function. One patient (2%) developed overt hypothyroidism. Multivariate analysis confirmed that positive TPO status and increased age were the only significant predictors of progression to overt hyperthyroidism. Positive TPO Hazard ratio (HR) was 10.15 (95% CI 1.83-56.23, p<0.01) and for each increase in age by a year the HR was 1.06 (95% CI 1.02-1.10, p<0.01).

Appendix number	2
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Das G, Ojewuyi TA, Baglioni P, Geen J, Premawardhana LD, Okosieme OE. Serum thyrotrophin at baseline predicts the natural course of sub clinical hyperthyroidism. Clinical Endocrinology. 2012;77(1):146-51**
Study details	Retrospective cohort study
Study objectives	Clarify relationship between TSH and progression to overt hyperthyroidism in patients with sub clinical hyperthyroidism followed up in a hospital outpatients department.
Inclusions	Patients with a sub clinical hyperthyroidism based on at least 2 hormone tests 3 months apart (TSH level below the reference range with normal T3 and T4 levels)
Exclusions	Patients with low TSH from non-thyroid causes, patients receiving medication known to affect thyroid function such as levothyroxine and anti-thyroid drugs.
Population	323 patients with sub clinical hyperthyroidism between 2003-2010
Intervention/ test	Serum TSH and free thyroxine (T4) were documented at baseline and during follow up (6-93 months)
Comparator	N/A
Results	Of the 323 patients followed up 183(56.7%) continued to have persistent sub clinical hyperthyroidism, 102(31.6%) reverted to normal thyroid function and 38(11.8%) progressed to overt hyperthyroidism. Only fully suppressed TSH at

	baseline <0.10mU/l was significantly associated with progression to overt hypothyroidism (hazard ratio 3.4 95% CI 1.6-7.0 p<0.001). Partially suppressed TSH, age, disease aetiology (Graves disease, toxic nodular goitre or indeterminate diagnosis) and baseline T4 levels did not influence progression to overt hyperthyroidism.

Appendix number	3
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Somwaru LL, Rariy CM, Arnold AM, Cappola AR. The natural history of sub clinical hypothyroidism in the elderly: the cardiovascular health study. Journal of Clinical Endocrinology & Metabolism. 2012;97(6):1962-9**
Study details	Longitudinal cohort study
Study objectives	Estimate persistence, resolution and progression of sub clinical hypothyroidism over 4 yrs
Inclusions	Individuals of at least 65 years of age who were part of the US Cardiovascular Health Study
Exclusions	Use of medication that would affect thyroid function (402 people)
Population	3996 people with blood samples stored from the US Cardiovascular Health Study.
Intervention/ test	Blood testing of serum TSH, free T4 total T3 and TPO antibodies
Comparator	N/A
Results	Of the total 3996 individuals on whom thyroid tests were performed at baseline: • 3057(85%) were eurthyroid • 459(12.8%) sub clinically hypothyroid • 22(0.61%) overtly hypothyroid • 12 (0.22%) overtly hyperthyroid • 44(1.2%) sub clinically hyperthyroid After 2yrs a further 84(2.7%) individuals who had been euthyroid had developed sub clinical hypothyroidism Of the 459 with sub clinical hypothyroidism at baseline 90 were excluded at the 2 year follow up point due to inadequate sera remaining for testing, death or no clinic visit made.

	Of the remaining 369 at 2 years: • 208(56%) remained sub clinically hypothyroid • 128(35%) reverted to euthyroidism • 8(2%) became overtly hypothyroid • 25(7%) initiated treatment with hormone replacement therapy At 4 years: • Of the 128 who were sub clinically hypothyroid at baseline who were euthyroid at year 2, 41(32%) were again sub clinically hypothyroid, 62 remained euthyroid, 5 had started medication and 9 had died. The remaining 11 did not have TSH measured. • Of the 208 who were sub clinically hypothyroid at baseline and yr 2, 122(59%) remained sub clinically hypothyroid, 16(8%) resolved to euthyroid 4 become overtly hypothyroid, 23 started taking thyroid medications, 11 died and the remainder had no TSH measured.

Appendix number	4
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Poola R, Mathiason MA, Caplan RH. A retrospective study of the natural history of endogenous sub clinical hyperthyroidism. Wisconsin Medical Journal. 2011;110(6):277-80**
Study details	Retrospective study of medical records
Study objectives	To examine the natural history of endogenous sub clinical hyperthyroidism
Inclusions	TSH test results from patients seen between 2002 and 2006
Exclusions	Exclusions include; patients with elevated free T4 , free T3, T3 levels; those with non-thyroid illness or who were acutely ill when TSH was measured; those taking thyroid medication that affects thyroid function, pregnant women and those without at least one follow up TSH measurement and patients were follow up was less than 6 months.
Population	116 patients with low TSH levels; normal free T4, T3 and free T3 measurements and at least 6 months follow up.
Intervention/ test	From retrospective analysis TSH, free T4, T3 and free T3 were determined from the notes of patients at baseline and at least 6 months follow up. Results of scans to assess whether thyroid gland was nodular, non-nodular or diffusely enlarged

	where documented.
Comparator	N/A
Results	TSH levels associated with nodular thyroid glands (18 patients) were less likely to revert to normal (17%) compared 59% in patients without nodules (98 patients). There were no differences in clinical outcome related to initial TSH levels or age (p=0.435). Aetiology of SH in 49% of patients could not be determined from patient notes.
Comments	

Appendix number	5
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Schouten BJ, Brownlie BEW, Frampton CM, Turner JG. Sub clinical thyrotoxicosis in an outpatient population-predictors of outcome. Clinical Endocrinology. 2011;74(2):257-61**
Study details	Retrospective study of patients attending a thyroid outpatient clinic.
Study objectives	Determine rate and risk factors for development of thyrotoxicosis
Inclusions	Consecutive patients diagnosed with sub clinical hyperthyroidism between Jan 1997 and Dec 2002
Exclusions	Patients were excluded if they; had non-thyroidal causes of subnormal TSH; used medication that affects thyroid function; were pregnant; had a severe life threatening illness; had a past history of treated thyrotoxicosis; had painless thyroiditis.
Population	96 patients met the criteria
Intervention/ test	Test results retrieved from patient notes include: TSH, total T4, T3, free T4, thyroid scinitiscan
Comparator	N/A
Results	Progression to overt thyrotoxicosis was: Yr 1 8%, Yr 2 16% Yr 3 21% yr5 26% Cumulative % requiring treatment for overt disease at 5 years was 9% Graves disease, 21% multi-nodal goitre, and 61% Autonomous nodule. Analysis of the influence of age, gender, family history, symptoms at presentation, presence of thyroid nodules, baseline TSH level, scintiscan category on development of overt disease was undertaken. Underlying thyroid pathology determined by

	scinitigraphy was the only independent predictor of outcome (p=0.003).

Appendix number	6
Relevant criteria	**Criterion 1: The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including the development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious treatable disease.**
Relevant Key question	What proportion of overt and sub clinical thyroid dysfunction reverts to normal function without clinical intervention?
Publication details	**Vadiveloo T, Donnan PT, Cochrane L, Leese GP. The thyroid epidemiology, audit, and research study (TEARS): The natural history of endogenous sub clinical hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2011;96(1):E1-E8**
Study details	Record-linkage technology was used retrospectively to identify patients with SH in the general population of Tayside, Scotland, from January 1, 1993, to December 31, 2009
Study objectives	Objective of the study was to define the rates of progression to frank hyperthyroidism and normal thyroid function.
Inclusions	All Tayside residents with at least 2 measurements of TSH below the reference range for at least 4 months from baseline and normal free T(4)/total T(4) and total T(3) concentrations at baseline were included as potential cases. Using a unique patient identifier, data linkage enabled a cohort of SH cases to be identified from prescription, admission, and radioactive iodine treatment records.
Exclusions	Cases younger than 18 yr of age were excluded from the study
Population	SH in the general population of Tayside.
Intervention/ test	The status of patients was investigated at 2, 5, and 7 yr after diagnosis.
Comparator	N/A
Results	2024 cases with SH were identified, a prevalence of 0.63% and an incidence of 29 per 100,000. Most SH cases without thyroid treatment remained as SH at 2 (81.8%), 5 (67.5%), and 7 yr (63.0%) after diagnosis. Few patients (0.5-0.7%) developed hyperthyroidism at 2, 5, and 7 yr. The percentage of SH cases reverting to normal increased with time: 17.2% (2 yr), 31.5% (5 yr), and 35.6% (7 yr), and this was more common in SH patients with baseline TSH between 0.1 and 0.4 mU/liter

Appendix number	7

Relevant criteria	**5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed**
Relevant key question	Key Question: Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction? Sub-questions : Is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
Publication details	**Asvold BO, Vatten LJ, Midthjell K, Bjoro T. Serum TSH within the reference range as a predictor of future hypothyroidism and hyperthyroidism: 11-Year follow-up of the HUNT study in Norway. Journal of Clinical Endocrinology and Metabolism. 2012;97(1):93-9**
Study details	Prospective population based study linked to the Norwegian Prescription Database
Study objectives	Assess TSH within the reference range and subsequent risk of hypo- or hyperthyroidism.
Inclusions	10,083 women and 5023 men older than 20 years without previous thyroid disease who had a baseline TSH of 0.20-4.5mU/litre and who participated at a follow up examination 11 years later.
Exclusions	• People under 20 years and over 70 years • People with current or previous thyroid disease • TSH outside the lab reference range • free T4 below the lab lower reference range • people with missing information on self reported thyroid disease.
Population	Inhabitants in Nord-Trøndelag County Norway
Test	TSH serum concentration Free T4 TPO antibodies
Comparator / reference standard	TSH reference range = 0.2-4.5 mU/litre T4 =8-20 pmol/litre TPO =<200U/ml
Results	Follow up median period was 11.1 yr (range 9.4-12.8 yr). At follow up: • 355(3.5%)of 10,083 women and 63(1.3%) of 5023 men developed hypothyroidism • 96 (1.1) women and 26 (0.6%) men developed hyperthyroidism • Baseline TSH within the reference range in women was positively associated with subsequent hypothyroidism($p<0.001$) compared to baseline TSH of 0.50-1.4mU/litre • The association in men was similar ($p<0.001$) but at any given level the risk of hypothyroidism was lower than in women. • At baseline of 0.5-1.4mU/l probability of becoming hypothyroid in women was 1.1% (CI 0.8-1.4)and men 0.3% (CI 0.1-0.6)

	• At baseline of 4.0-4.5mU/l probability of becoming hypothyroid in women was 31.5% (CI 24.6-39.3) and in men was 14.7 (CI 7.7-26.2) • In people with baseline 4.1-4.5 mU/l TPO antibodies were measured . People with TPO antibodies were at higher risk of hypothyroidism (women with TPO 43.3% CI 29.5-58.1% and men 21.5% CI 7.1-49.6%) compared to without(women 20.3% CI 12.1-31.9% and men12.9% CI 5.4-27.8). • The risk of hyperthyroidism was higher in women with baseline TSH of 0.20-0.49mU/l (3.9% CI1.8-8.4%) than baseline of 0.5-0.99 mU/l (1.4% CI 0.9-2.1%). • Low numbers of hyperthyroid men precluded precise estimates of association with baseline TSH.

Appendix number	8
Relevant criteria	**5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed**
Relevant key question	Key Question: Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction? Sub-questions : Is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
Publication details	**Coene KL, Demir AY, Broeren MA, Verschuure P, Lentjes EG, Boer AK. Sub clinical hypothyroidism: a 'laboratory-induced' condition? European Journal of Endocrinology. 2015;173(4):499-505**
Study details	Retrospective analysis of anonymous TSH results from Dutch external quality assessment program and examination of TSH/freeT4 reference ranges and prevalence of thyroid disease among 9 Dutch laboratories 4 of which switched platforms.
Study objectives	Examine prevalence of sub clinical hypothyroidism based on results from 9 different Dutch laboratories using 5 different platforms.
Inclusions	Nine laboratories submitted results of all TSH/fT4 carried out over a year.
Exclusions	
Population	All samples tested for TSH and fT4 over a one year period
Test	TSH and fT4 using 5 different platforms in 9 different laboratories. 4 of the laboratories switched platforms mid-year.
Comparator / reference standard	Varied - provided by individual manufacturers and laboratories.

Results	• There was high inter-laboratory variation in TSH/fT4 both within and between platforms. • Beckman Coulter Access and Siemans Immulite users reported a lower prevalence of sub clinical hypothyroidism than Roche Cobas and Abbott Architect users (4.6%, 7.6%, 14.3% and 15.5% respectively). • Differences between laboratories using the same platform were apparent eg: Roche Cobas in one laboratory reported 11.8% sub clinical hypothyroidism whilst another reported 17.8% and sub clinical hyperthyroidism ranged from 3.4% 50 11.2%. • Four labs switched platforms part way through the year which enabled the study to exclude natural differences in prevalence in the population. Changes in prevalence of the different thyroid conditions were reported after the platforms changed. Three labs moved to the Roche Cobas platform and there was an average increase in prevalence of sub clinical hypothyroidism of 6.2%. With an average of 26,000 tests being carried out by laboratory the shift of these three laboratories could result in an additional 5000 patients per year receiving a sub clinical hypothyroidism diagnosis. For fT4 no relevant differences were observed. • At the level of 10mU/l (a typical level for pharmaceutical intervention), Roche cobas users reported an average of 2.5% sub clinical hypothyroidism while for Beckman Coulter and Siemans Immulite users this was 1.3 %. This equates to an extra 300patients per Roche laboratory receiving fT4 substitution medication.

Appendix number	9
Relevant criteria	**5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed**
Relevant key question	Key Question: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction? Sub-questions : is there agreement on what is a healthy level of each of the T3, T4 and TSH hormones in non-pregnant adults?
Publication details	**Langen VL, Niiranen TJ, Maki J, Sundvall J, Jula AM. Thyroid-stimulating hormone reference range and factors affecting it in a nationwide random sample. Clinical Chemistry and Laboratory Medicine. 2014;52(12):1807-13**
Study details	Finnish nationwide epidemiological survey using a stratified 2 stage cluster sample.
Study objectives	To determine TSH reference range on the Abbott Architect ci8200 integrated system
Inclusions	Adults over 30 years of age
Exclusions	Of 6247 people were excluded who: • Were under 30 years of age • Declined to participate

	• Were unable to provide a blood sample or had it taken before 8am and after 6pm • Were non-ambulatory • Used thyroid hormones or anti-thyroid agents • Had a personal history of thyroid problems or goiter • Were pregnant or breastfeeding The remaining 5709 were the thyroid health sub-population • Of these 5709 those with a TSH range <0.4mU/l or >2.5mU/l with TPOAB >5.6IU/l based on equipment manufacturers upper limit of reference range were excluded The remaining 4586 were the risk factor free population • Of these 4586 those taking medications that have a potential effect on thyroid function tests were excluded The remaining 3453 were the sub-group with no thyroid effecting medications. • Of these 3453 participants all those taking any medications were excluded The remaining 1849 was the reference sub population.
Population	8028 people 30 years and over was randomly drawn from the population register of which 6247 (77.8) agreed to participate in a health interview, health examination and provided a sample of blood.
Test	TSH – all samples tested on one machine.
Comparator / reference standard	TSH reference ranges were established directly from the 2.5th and 97.5th percentile s of the TSH measurements in the thyroid health population and all its subsets. The validity of this approach was controlled by defining the TSH reference range from the reference sub-population and also from data on the 95% confidence limits of its TSH values which were transformed with the best suitable function to obtain a Gaussian distribution.
Results	• Exclusion of TPOAb-positive participants from the thyroid health population resulted in decrease in TSH reference range upper limit from 4.43mU/l to 3.71mU/l • The exclusion of participants using thyroid affecting medication lowered the TSH URL by 0.1mU/l • The exclusion of participants with any medications lowered the TSH URL by 0.3 mU/l

Appendix number	10
Relevant criteria	**5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed**
Relevant key question	Key Question: Has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction? Sub-questions : Is there agreement on what is a healthy level of each of the T3,

	T4 and TSH hormones in non-pregnant adults?
Publication details	**Milinkovic N, Ignjatovic S, Zarkovic M, Jovicic S, Radosavljevic B, Singh S, et al. Indirect estimation of age-related reference limits of thyroid parameters: a cross-sectional study of outpatients' results. Scandinavian Journal of Clinical & Laboratory Investigation. 2014;74(5):378-84**
Study details	Retrospective cross sectional analysis of laboratory information of 22,860 blood samples recording TSH, fT4 and T3
Study objectives	Defining adequate reference limits for thyroid hormones (TSH, fT4 and T3) for sub clinical thyroid disease by age and gender.
Inclusions	Outpatient results of first blood test reported per patient
Exclusions	Those results for people: • <18 • For whom this is a second or subsequent test • With anti-TPO and Tg results • Positive anti-TPO and Tg results • Thyroid hormones above or below the reference ranges • Pregnant women
Population	Results from blood tests of outpatients in adults 18 and over.
Test	TSH, fT4 and T3 using Abbott Architect ci8200
Comparator / reference standard	TSH, fT4 and T3 reference ranges using 2.5 and 97.5 percentiles thresholds
Results	• Change in mean TSH for males was 0.42mIU/land 0.69 ($p<0.0001$) between <30 years and >70 years. • Changes in mean fT4 concentration of 0.39 pmol/l was found between <30 year olds and 41-50 year olds. • There was no significant difference in changes of the mean of T3 in the populations • A significant difference in TSH was observed between men and women over age 70 years ($p<0.05$) • A significant difference between male and female fT4 values in 3rd and 4th decade was observed($p<0.05$)

Appendix number	11
Relevant criteria	**5. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed**
Relevant key question	Key Question: has a test cut-off been identified which is suitable for population screening for overt and sub clinical thyroid dysfunction? Sub-questions : is there agreement on what is a healthy level of each of the T3,

	T4 and TSH hormones in non-pregnant adults?
Publication details	**Vadiveloo T, Donnan PT, Murphy MJ, Leese GP. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: The thyroid epidemiology, audit, and research study (TEARS). Journal of Clinical Endocrinology and Metabolism. 2013;98(3):1147**
Study details	Retrospective analysis of a TSH test results of residents in Tayside whose TSH was tested between July 2003 and December 2009.
Study objectives	To examine the association of TSH with age, gender and diabetes in a large population based cohort without evidence of thyroid disease
Inclusions	Tayside residents
Exclusions	Test results were excluded from: • People with identifiable thyroid disease (determined from record linkage of data) • People under 18 years of age • People ever treated with radioactive iodine • People who have taken T4, carbimazole, propylthiouracil, amidarone or liothyronine at any time • People who had ever tested positive for thyroid antibodies • People who had undergone thyroid surgery • People diagnosed with thyroid cancer or a pituitary disorder • People who were in hospital when the test was taken • Women who were pregnant • People with a result over 20mU/l
Population	All Tayside residents who had thyroid function tests between 2003 and 2009. A cohort of residents with no thyroid disease was identified (from record linkage).
Test	TSH test
Comparator / reference standard	Reference range of TSH with cut-off thresholds at 2.54 and 97.5 percentiles.
Results	• The 153,127 TSH measurements identified were grouped according to age of the person and level of TSH (group 1 >0.4Um/l, group 2 =0.4-4.0Um/l, group 3 <4.0Um/l) • The proportion of TSH measurements in groups 1 and 3 increased progressively with age whilst group 2 decreased with age • At 31-40 years - group 1 = 1.1%, group 2=96.4% group 3=2.4% whilst those older than 90 years - group 1=3.5%, group 2=86.2% and group 3=10.2% • There was an increase in median with increasing age ($p<0.001$) • There was a decrease in the cut-off point for the 2.5^{th} centile with increasing age and an increase in the median and 97.5^{th} centile with age • TSH moved towards higher concentrations with increasing age regardless of the presence or absence of thyroid antibodies • The median TSH was significantly different in men and women ($p<0.001$) but the pattern of increase with age was similar

	• There was a significant difference in the TSH median in those with and without diabetes($p<0.001$)

Appendix number	**12**
Relevant criteria	**Criterion 9: There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered**
Relevant Key question	What is the effectiveness of treatment of overt and sub clinical thyroid disease?
Publication details	**Razvi S, Weaver JU, Butler TJ, Pearce SHS. Levothyroxine treatment of sub clinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. Archives of Internal Medicine. 2012;172(10):811**
Study details	Retrospective cohort study of UK individuals with TSH between 5.01-10mU/l and normal FT4 first tested in 2001 and followed for up to 8 years.
Study objectives	To investigate the association between levothyroxine treatment of sub clinical hypothyroidism and ischaemic heart disease, morbidity and mortality.
Inclusions	Individuals aged 40 or older, with first ever TSH level between 5.01-10.00mU/l and normal FT4.
Exclusions	Individuals with : • History of ischaemic heart disease • History of cerebrovascular disease • Poor quality records with lack of continuous follow up or incomplete and inaccurate data • Previous treatment with amirodarone hydrochloride, lithium carbonate, or in the previous 12 months with corticosteriods
Population	Case records of GP research database
Intervention/ test	TSH at baseline for age groups 40-70 years and >70 years
Comparator	N/A

Results	Treatment with levothyroxine in individuals aged 40 to 70 was associated with: • reduced all cause mortality, 0.36(0.19-0.66) • fewer Ischemic Heart Disease (IHD) events, 0.61(0.39-0.95) • fewer deaths due to circulatory diseases 0.54(0.37-0.92) • fewer deaths due to cancer 0.59(0.21-0.88) The same effects were not found in the >70 years age group

Appendix number	**13**
Relevant criteria	**Criterion 9: There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered**
Relevant Key question	What is the effectiveness of treatment of overt and sub clinical thyroid disease?
Publication details	**Rugge J, Bougatsis C & Chou R. *Screening for and treatment of thyroid dysfunction; an evidence review for the US Preventative Services Task Force*. Agency for Healthcare Research and Quality (2016) Rockville.**
Study details	Evidence review on screening and treatment of sub clinical and overt thyroid disease
Study objectives	To update the US Preventative Services Task Force (USPSTF) 2004 review of screening to inform clinical practice.
Inclusions	Randomised controlled trials and controlled observational studies on the effects of screening for or treatment of sub clinical or overt thyroid disease on clinical and intermediate outcomes.
Exclusions	Non-English language articles, studies only published as abstracts
Population	Community living, non-pregnant adults without a clear history of thyroid disease or clear symptoms of overt thyroid disease who were screened and treated for thyroid disease.
Intervention/ test	Screening, treatment of overt and sub clinical thyroid disease using hormone replacement therapy, anti thyroid medications and ablation surgery.
Comparator	No screening, no treatment or observation.
Results	**Sub clinical hypothyroidism** <u>Cardiovascular events and mortality</u>- one retrospective cohort study suggested levothyroxine use versus no treatment was associated with lower risk for fatal or non-fatal ischaemic heart disease events in patients in the age group 40-70 years (Hazard ratio 0.61[95% CI; 0.39-0.95]). Patients older than 70 did not have the

	same association (HR 0.99[95% CI 0.59-1.33]). Similar positive associations were reported for the younger age group for all cause mortality, death due to circulatory diseases and cancer mortality but no association was apparent for those over 70 years of age. *Cognitive function Two trials were identified that appeared to evaluate screen detected population. Neither trial found treatment effects on various measures of cognitive function (cognitive skills and performance (p=0.57), cognitive status p=0.18, speed of cognitive processing (p=0.59).* *Blood pressure – 3 trials assessed effects of treatment on blood pressure, none found a significant difference between treatment and placebo groups.* *Cholesterol - In eight good and fair quality trials mean differences between treatment and no treatment varied with three trials reporting statistical differences in mean cholesterol levels (p<0.0001).* *LDL cholesterol- Three of eight trials reported significantly significant differences in mean LDL values with treatment (p<0.0001, p-0.03, p<0.0001)* *HDL cholesterol – None of eight fair and good quality trials found significant differences in HDL values between the treatment and control groups.* *Triglycerides – None of eight fair and good quality trials found significant differences between treatment and control groups for mean level of triglycerides.* *Body mass index or weight – Six good or fair quality trials reported BMI or weight in treatment and placebo groups. Five trials reported BMI and two trials reported weight. None of the trials found a significant difference in mean BMI or weight between the treatment and control groups.* *Quality of life Five trials were identified using various doses of levothyroxine and measuring quality of life. No differences were found between treatment and placebo in any study.* *Harms –One trial directly compared harms between treated and non treated adults and found no difference in withdrawals due to side effects after 12 months (p=0.49).* ***Sub clinical hyperthyroidism*** *Two poor quality trials reported effects of treatment on blood pressure and one trial reported effects on body mass index and bone mineral density and lipids. No differences between means of treatment and placebo groups for any of the variables were reported.* **Overt thyroid disease** No studies examined the effect of treatment of overt thyroid disease on outcomes.
Comments	The findings from the retrospective cohort study reporting cardiovascular events associated with treatment with levothyroxine are reported separately (Razvi,2012)[32]

	Findings in italics were previously reported in the UK NSC update (2011)[15]

Appendix number	**14**
Relevant criteria	**Criterion 9: There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered**
Relevant Key question	What is the effectiveness of treatment of overt and sub clinical thyroid disease?
Publication details	**Stott DJ, Rodondi N, Kearney PM, Ford I, et al Thyroid hormone therapy for older adults with sub clinical hypothyroidism (2017) *New England Journal of Medicine*. DOI: 10.1056/NEJMoa1603825**
Study details	Multi centre (UK, Ireland, Netherlands & Switzerland) double blind, randomised, placebo controlled parallel group trial of 737 adults ≥65 years of age with persistent sub clinical hypothyroidism treated with placebo or levothyroxine.
Study objectives	To determine the clinical benefits of levothyroxine in older people with sub clinical thyroidism.
Inclusions	People≥65 years
Exclusions	• People with prescriptions for levothyroxine, anti -thyroid drugs, amiodarone, or lithium • People who have had thyroid surgery • People in receipt of radioactive iodine within the previous 12 months • People with dementia • Hospitilisation, major illness or elective surgery within 4 weeks • Acute coronary syndrome within previous 4 weeks • Terminal illness
Population	Participants identified from clinical laboratory and general practice databases who were ≥65 years of age with TSH of 4.6-19.99mU/l and fT4 within the reference range, measured on 2 occasions at least 3 months to 3 years apart.
Intervention/ test	Active intervention – levothyroxine dose of 50µg daily (or 25µg in patients with a body weight<50 kg). Dose adjustment was aimed to result in TSH reference range of 0.40-4.59mU/l.
Comparator	Placebo with mock dose adjustment
Results	Primary outcomes were differences in hypothyroid symptoms and tiredness scores from the ThyPro scale carried out at baseline and one year in the placebo and treatment groups. • At 12 months the mean hypothyroid symptom score was 16.7±17.5 in the

	placebo group and 16.6±16.9 in the intervention group (p=0.99) • The mean tiredness score was 28.6±19.5 in the placebo group and 28.7±20.2 in the levothyroxine group (p=0.77) • There were no differences in the mean change in the hypothyroid symptoms score and the mean tiredness score at 1 year from baseline between the placebo and levothyroxine group • There was a between group difference (p=0.05) in the tiredness score with a lower value in the levothyroxine group at extended follow up (mean extended follow up 24 months) • The number of patients with at least one serious adverse event was higher in the placebo group than in the levothyroxine group (p=0.049) but there was no pattern of event type that contributed to difference.

References

[1] Madariaga, APalacios S, Guillen-Grima F, Galofre J The incidence and prevalence of thyroid dysfunction in Europe: A meta analysis Journal of Clinical Endocrine Metabolism, 2014, 99(3):923-931.
[2] Madariaga, APalacios S, Guillen-Grima F, Galofre J The incidence and prevalence of thyroid dysfunction in Europe: A meta analysis Journal of Clinical Endocrine Metabolism, 2014, 99(3):923-931.
[3] Okosieme O, Gilbert J, Abraham P, Boelaert K, Dayan C, Gurnell M, et al. Management of primary hypothyroidism: Statement by the British Thyroid Association Executive Committee. Clinical Endocrinology. 2016;84(6):799-80
[4] Association for Clinical Biochemistry , British Thyroid Association and the British Thyroid Foundation. UK guidelines for the use of thyroid function tests. 2006. Available from http://www.btf-thyroid.org/images/stories/pdf/tft_guideline_final_version_july_2006.pdf (Accessed May 2017)
[5] McQuade C, Skugor M, Brennan DM, Hoar B, Stevenson C, Hoogwerf BJ. Hypothyroidism and moderate subclinical hypothyroidism are associated with increased all-cause mortality independent of coronary heart disease risk factors: A précis database study. Thyroid 2011;21 (8):837-43
[6] Hogervorst E, Huppert F, Matthews FE, Brayne C. Thyroid function and cognitive decline in the MRC Cognitive Function and Ageing Study. Psychoneuroendocrinology 2008;33(7):1013-22
[7] Biondi B, Bartalena L, Cooper DS, Hegedus L, Laurberg P, Kahaly GJ. The 2015 European Thyroid Association guidelines on diagnosis and treatment of endogenous sub clinical hyperthyroidism. European Thyroid Journal. 2015;4(3):149-63
[8] Gammage MD, Parle JV, Holder RL, Roberts LM, Hobbs FDR, Wilson S et al. Association between serum free thyroxine concentration and atrial fibrillation. Archives of Internal Medicine 2007; 167 (9):928-934

[10] National Institute of Health and Care Excellence Hypothyroidism Clinical Knowledge Summary April 2016 https://cks.nice.org.uk/hypothyroidism#!scenario:1(accessed June 1st 2017)
[11]Rugge B, Balshem H, Sehgal R, Relevo R, Gorman P, Helfand M. Screening and treatment of sub clinical hypothyroidism or hyperthyroidism. Agency for Healthcare Research & Quality (AHRQ). 2011. http://www.effectivehealthcare.ahrq.gov/ehc/products/129/750/Hypo-Hyper-Thyroid_CER24_20111114.pdf
[12] National Institute of health and Care Excellence, Hyperthyroidism Clinical Knowledge Summary June 2016
https://cks.nice.org.uk/hyperthyroidism (Accessed June 1st 2017)
[13] Pittam G. Allaby M. Coles S. Screening for thyroid disease. A report for the UK National Screening Committee. February 2013
[14] Abu-Helalah M, Law MR, Bestwick JP, Monson JP, Wald NJ. A randomized double-blind crossover trial to investigate the efficacy of screening for adult hypothyroidism. Journal of Medical Screening 2010; 17(4):164-69
[15] UK NSC. Adult screening for thyroid dysfunction: briefing note. October 2016
[16] Zhyzhneuskaya S, Addison C, Tsatlidis V, Weaver JU, Razvi S. The natural history of sub clinical hyperthyroidism in graves' disease: The rule of thirds. Thyroid. 2016;26(6):765-9

[17] Das G, Ojewuyi TA, Baglioni P, Geen J, Premawardhana LD, Okosieme OE. Serum thyrotrophin at baseline predicts the natural course of sub clinical hyperthyroidism. Clinical Endocrinology. 2012;77(1):146-51

[18] Poola R, Mathiason MA, Caplan RH. A retrospective study of the natural history of endogenous sub clinical hyperthyroidism. Wisconsin Medical Journal. 2011;110(6):277-80

[19] Schouten BJ, Brownlie BEW, Frampton CM, Turner JG. Sub clinical thyrotoxicosis in an outpatient population-predictors of outcome. Clinical Endocrinology. 2011;74(2):257-61

[20] Vadiveloo T, Donnan PT, Cochrane L, Leese GP. The thyroid epidemiology, audit, and research study (TEARS): The natural history of endogenous sub clinical hyperthyroidism. Journal of Clinical Endocrinology and Metabolism. 2011;96(1):E1-E8

[21] Somwaru LL, Rariy CM, Arnold AM, Cappola AR. The natural history of sub clinical hypothyroidism in the elderly: the cardiovascular health study.Journal of Clinical Endocrinology & Metabolism.2012;97(6):1962-9

[22] Haentjens P, Van Meerhaeghe A, Poppe K, Velkeniers B. Subclinical thyroid dysfunction and mortality: An estimate of relative and absolute excess all-cause mortality based on time-to-event data from cohort studies. European Journal of Endocrinology 2008; 159 (3):329-341

[23] Asvold BO, Vatten LJ, Midthjell K, Bjoro T. Serum TSH within the reference range as a predictor of future hypothyroidism and hyperthyroidism: 11-Year follow-up of the HUNT study in Norway. Journal of Clinical Endocrinology and Metabolism. 2012;97(1):93-9

[24] Coene KL, Demir AY, Broeren MA, Verschuure P, Lentjes EG, Boer AK. Sub clinical hypothyroidism: a 'laboratory-induced' condition? European Journal of Endocrinology. 2015;173(4):499-505

[25] Milinkovic N, Ignjatovic S, Zarkovic M, Jovicic S, Radosavljevic B, Singh S, et al. Indirect estimation of age-related reference limits of thyroid parameters: a cross-sectional study of outpatients' results. Scandinavian Journal of Clinical & Laboratory Investigation. 2014;74(5):378-84

[26] Langen VL, Niiranen TJ, Maki J, Sundvall J, Jula AM. Thyroid-stimulating hormone reference range and factors affecting it in a nationwide random sample. Clinical Chemistry and Laboratory Medicine. 2014;52(12):1807-13

[27] Demers LM, & Spencer CA, editors. Laboratory medicine practice guidelines; laboratory support for the diagnosis and monitoring of thyroid disease. Washington, DC: National Academy of Clinical Biochemistry 2002.

[28]Vadiveloo T, Donnan PT, Murphy MJ, Leese GP. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: The thyroid epidemiology, audit, and research study (TEARS). Journal of Clinical Endocrinology and Metabolism. 2013;98(3):1147

[29] Thienpont LM, Van Uytfanghe K, Van Houcke S, Das B, Faix JD, MacKenzie F, et al. A Progress Report of the IFCC Committee for Standardization of Thyroid Function Tests. European Thyroid Journal. 2014;3(2):109-16

[30] Joint Formulary Committee, *British National Formulary 73*, 2017, London. BMJ Group and Pharmaceutical Press

[31] Abu-Helalah M, Law MR, Bestwick JP, Monson JP, Wald NJ. A randomized double-blind crossover trial to investigate the efficacy of screening for adult hypothyroidism. Journal of Medical Screening 2010; 17(4):164-69

[32] Razvi S, Weaver JU, Butler TJ, Pearce SHS. Levothyroxine treatment of sub clinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. Archives of Internal Medicine. 2012;172(10):811

[33] Rugge JB, Bougatsos C, Chou R. Screening and treatment of thyroid dysfunction: An evidence review for the U.S. preventive services task force. Annals of Internal Medicine. 2015;162(1):35-45

[34] Stott DJ, Rodondi N, Kearney PM, Ford I, et al Thyroid hormone therapy for older adults with sub-clinical hypothyroidism (2017) New England Journal of Medicine. DOI: 10.1056/NEJMoa1603825