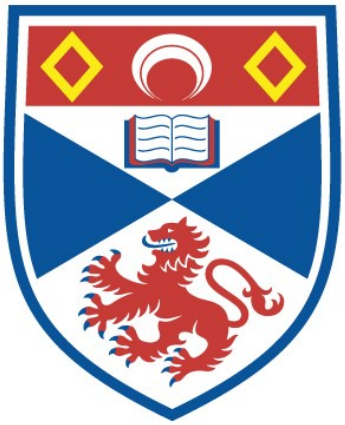
Investigating the impact of the management of subclinical hypothyroidism on long-term clinical outcomes

Brenda Sarange Bauer

A thesis submitted for the degree of PhD
at the
University of St Andrews



2025

Declaration

Candidate's Declaration

I, Brenda Sarange Bauer, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 40,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in October 2020.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date: 17 December 2024

Signature of candidate:

Supervisor's Declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date: 13th December 2024

Signature of supervisor:

Permission for Publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Brenda Sarange Bauer, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy
No embargo on print copy.

Electronic copy
Embargo on part (Chapter 4 Target trial emulation, Chapter 5 Discussion, Appendix III) of electronic copy for a period of 3 years on the following ground(s):
• Publication would preclude future publication

Supporting statement for electronic embargo request
The trial emulation manuscript is yet to be published but I plan to submit it to a peer-reviewed journal at a later date.

Title and Abstract
• I require an embargo on the abstract only.

Date: 17 December 2024

Signature of candidate:

Date: 13th December 2024

Signature of supervisor:

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Brenda Sarange Bauer, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date: 17 December 2024

Signature of candidate:

Abstract

Contents

List of Figures

List of Tables

List of Abbreviations

ACE-SA Association of Clinical Endocrinologists of South Africa
ADDE Annual District Death Extract
ALF Anonymous Linking Field
ALF-E Anonymous Linking Field - Encrypted
AMSTAR Assessing the Methodological Quality of Systematic Reviews
ATA American Thyroid Association
AWISS All Wales Injury Surveillance Systems
BMI Body Mass Index
CCA Corrected Covered Area
CCI Charlson Comorbidity Index
CCW Clone-censor-weight
CI Confidence interval
CPRD Clinical Practice Research Datalink
CVD Cardiovascular disease
DHCW Digital Health & Care Wales
EDDS Emergency Department Dataset
EHR Electronic Health Records
EMIS Egton Medical Information Systems
ETA European Thyroid Association
FT3 Free Triiodothyronine; Free T3
FT4 Free Thyrotropin; Free T4
GP General Practitioner
GRADE Grading of Recommendations, Assessment, Development, and Evaluations
HDRUK Health Data Research UK
HIRU Health Information Research Unit
HR Hazard ratio
IEMO Institute for Evidence-based Medicine in Old age
IGRP Information Governance Review Panel
IPCW Inverse probability of censoring weighting
IPTW Inverse probability of treatment weighting
IPW Inverse Probability Weights
IU International Units
KNHANES Korean National Health and Nutrition Examination Survey

KTA Korean Thyroid Association
LATS Latin American Thyroid Society
LT4 Levothyroxine
MA Meta-analysis
MACE Major Adverse Cardiovascular Events
NHANES National Health and Nutrition Examination Survey
NHS National Health Service
NICE National Institute for Health and Care Excellence
NT Nested trial
ONS Office for National Statistics
OPDW Outpatient Database for Wales
OPRD Outpatient Referral Dataset
PEDW Patient Episode Dataset for Wales
PLoS Public Library of Science
PP Per-protocol
PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QoL Quality of life
RCT Randomised Controlled Trial
RR Relative risk
SAIL Secure Anonymised Information Linkage
SCH Subclinical hypothyroidism
SD Standard Deviation
SEMDSA Society for Endocrinology, Metabolism and Diabetes of South Africa
SR Systematic review
T3 Triiodothyronine
T4 Thyroxine
TFT Thyroid Function Test
TPOAb Thyroid peroxidase antibody
TRH Thyrotropin Releasing Hormone
TRUST Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism
TSH Thyroid Stimulating Hormone
TTE Target Trial Emulation
ULN Upper Limit of Normal
UTREC University of St Andrews Teaching and Research Ethics Committee
WDSD Welsh Demographic Service Dataset
WIMD Welsh Index of Multiple Deprivation
WLGP Welsh Longitudinal General Practice
WRRS Wales Results Reporting Service

Acknowledgements

To my supervisors, Prof Colin McCowan, Dr Amaya Azcoaga-Lorenzo and Dr Utkarsh Agrawal, I will forever be grateful for this opportunity and the support you gave me through this journey. Thank you for the many questions, which were always on point!

Thank you also to the staff at the School of Medicine and within the wider University of St Andrews who directly or indirectly helped me along the way.

To my parents, David and Margaret, thank you for getting me here and leading by example.

To Fabian and our precious lovelies Lars and Liana, thank you most of all, for the love that has kept me going thus far and continues to sustain me.

Funding

My PhD was funded by the University of St Andrews School of Medicine through the Maitland Ramsay Postgraduate Research Scholarship and Handsel Tuition Fee Scholarship -Medicine.

Statement of contribution

I, Brenda Bauer, confirm that I have composed this portfolio-based thesis myself. I led in all the aspects of the research contained in this thesis, as reflected in being the first author of the respective articles. This work was done with input from my supervisors, who offered advice on the research design and methodology and reviewed all the drafts before publication. They are listed as my co-authors.

Peer-reviewed publications in this thesis

1. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, McCowan C. Management strategies for patients with subclinical hypothyroidism: a protocol for an umbrella review. Systematic reviews. 2021 Dec; 10:1-6. https://doi.org/10.1186/s13643-021-01842-y

2. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C (2022) The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review. PLOS ONE 17(5): e0268070. https://doi.org/10.1371/journal.pone.0268070

3. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C. Subclinical hypothyroidism in Wales from 2000 to 2021: A descriptive cohort study based on electronic health records. PLOS ONE. 2024 May 21;19(5): e0298871. https://doi.org/10.1371/journal.pone.0298871

1

Introduction

This chapter begins with a comprehensive description of subclinical hypothyroidism (SCH), including its definition, diagnosis, management, and related prevailing controversies. The research questions, aims and objectives that guided this project are then presented, followed by the structure of the thesis. This chapter, therefore, provides the background information and justification for the rest of the thesis.

1.1 Thyroid

1.1.1 Thyroid gland

The thyroid gland is an endocrine organ whose anatomical position is the base of the neck. It is frequently described as 'butterfly-shaped' because it comprises two larger lobes, one on either side of the trachea, connected in the front by an isthmus (Figure 1.1). The role of the thyroid gland is hormone production; thyroxine (T4) and triiodothyronine (T3) are the two main thyroid hormones, as well as calcitonin [1].

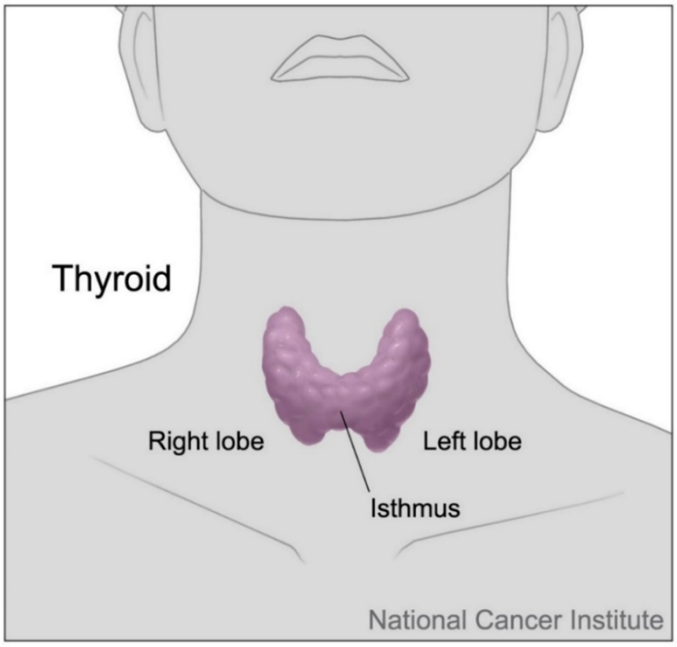


Figure 1.1. *Anatomical features of the thyroid gland (reproduced from the National Cancer Institute courtesy of Don Bliss [2]).*

1.1.1.1 Thyroid hormones

T3 and T4 have diverse roles in the body, influencing all the major organ systems. Their functions include regulating body temperature, metabolism, growth and development of the brain [3]. Calcitonin regulates blood calcium levels and differs markedly from T3 and T4 in its production, mechanism of action and regulation [4]. For the remainder of this text, 'thyroid hormone(s)' will refer to T3 and T4 only.

1.1.1.2 Thyroid homeostasis

Homeostasis refers to a state of balance - thyroid hormones are regulated through the hypothalamic-pituitary-thyroid axis (Figure 1.2). In response to decreasing amounts of thyroid hormones in the blood, the hypothalamus in the brain produces Thyrotropin Releasing Hormone (TRH). TRH stimulates the pituitary gland in the brain to release TSH, which in turn stimulates the production of thyroid hormones. It should be noted, however, that most of the T3 circulating in the body is a product of enzymatic action on T4 (deiodination) in peripheral organs such as the liver [5].

In contrast, high thyroid hormone levels activate the negative feedback loop, inhibiting TRH production. External factors such as environmental pollutants are also thought to influence this axis at the hypothalamic level [7, 8].

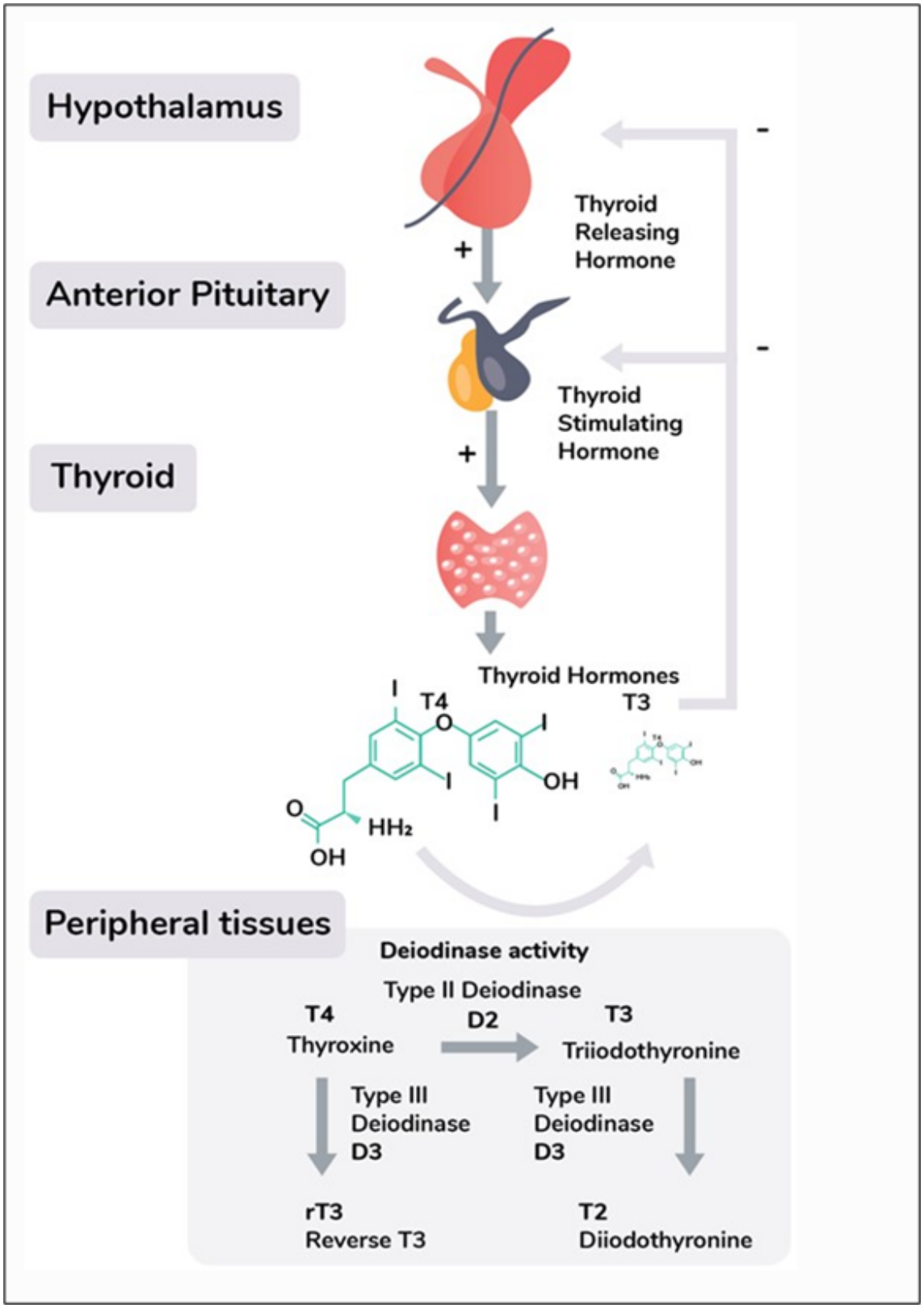


Figure 1.2. *The hypothalamic-pituitary-thyroid axis (reproduced from Keestra et al. [6] which is licensed under the Creative Commons CC BY license).*

1.2 Subclinical Hypothyroidism

1.2.1 Definition

Subclinical hypothyroidism (SCH) is an endocrine disorder characterised by levels of circulating thyroid hormones (free T3 and free T4) within normal range and Thyroid Stimulating Hormone (TSH) above the upper reference limit [9–11]. SCH is biochemically distinguishable from overt hypothyroidism – where TSH is also raised, but free thyroid hormone levels are lower than normal [12].

The upper limit of the TSH normal range varies, but based on the commonly used threshold of 4 mIU/L [13, 14], SCH may be considered to be mild if TSH concentrations are between 4 and 9.9 mIU/L or severe if TSH concentrations are higher than or equal to 10 mIU/L [15]. It has been estimated that up to 75% of SCH patients have mild disease [16].

1.2.2 Prevalence

The prevalence of SCH is estimated to be 4-20% of the adult population, depending on age, race, sex and iodine intake [15, 17, 18]. Other studies have reported between 1% and 15% prevalence in elderly populations [19, 20]. Prevalence likely varies between studies because of the inconsistent application of diagnostic reference ranges, selection criteria and sampling strategies [19, 21].

TSH reference ranges tend to vary across populations, and even laboratories in the same geographical area could apply different values depending on the equipment used for the assays [22]. In general, SCH is more common in women partly due to their underlying predisposition to autoimmune thyroid disease, a known cause of SCH [15, 23] (Table 1.1).

Table 1.1. *SCH prevalence in the literature (adapted from Yoo and Chung [24]).*

Year	Country	Age(years)	TSH threshold (mIU/L)	Prevalence, % (women)	Prevalence, % (men)
1977	UK (the Whickham survey) [25]	>18	>6.0	7.5	2.8
1990	USA (nursing home) [26]	>60	>4.5	14.6	9.7
1993	Japan (health examination) [27]	Mean 46	>5.0	2.1	0.4
2000	USA (the Colorado study) [28]	≥18	>5.1	9.1 (men and women)	
2002	USA (NHANES III) [29]	≥12	>4.5	4.3 (men and women)	
2006	The Netherlands [30]	>18 (46% >69)	>4.0	4.9	3.0
2017	South Korea (KNHANES VI) [31]	≥10	>6.86	4	2.3
2019	Europe (meta-analysis) [20]		>4.5	4.8	2.7

Note:
Abbreviations: TSH, thyroid-stimulating hormone; NHANES, National Health and Nutrition Examination Survey; KNHANES, Korean National Health and Nutrition Examination Survey.

1.2.3 Causes

The most common cause of SCH is autoimmune diseases that affect the thyroid gland, particularly Hashimoto's disease, which is found in up to 80% of cases [15]. Other causes of TSH elevation include iodine deficiency, radioiodine treatment for hyperthyroidism, medication including amiodarone and lithium, recovery from non-thyroidal illness as well as the surgical removal of sections of the thyroid [5, 15, 32, 33].

1.2.4 Natural history

Progression of SCH to overt hypothyroidism occurs in approximately 2-6% of patients per year – it is more likely in women and those with both elevated TSH and antithyroid autoantibodies [9, 15, 16]. Subsequently, overt hypothyroidism may cause non-specific but progressively debilitating symptoms based on its severity, duration and the patient's treatment status [12, 34].

However, around 60% of cases, particularly those with mild SCH, have been found to revert to normal thyroid hormone levels (euthyroidism) over time without ever receiving treatment [16, 17, 35, 36]. The mechanism of such reversal is poorly understood [36], but previous findings suggest that the hormonal changes could be seasonal [35].

1.2.5 Clinical presentation

Patients with SCH are typically asymptomatic [9, 37]. However, they can sometimes have mild, non-specific symptoms such as fatigue, dry skin and constipation [14, 15, 38] that may be inadvertently overlooked or attributed to other illnesses.

1.2.6 Diagnosis

Since it is a biochemical imbalance, SCH is only diagnosed through laboratory blood tests measuring thyroid hormones and TSH, also known as thyroid function tests (TFTs). Therefore, SCH is often detected by chance when TFTs are ordered in addition to other investigations or in a routine test panel [11, 39].

For TFTs, optimal accuracy in diagnosing thyroid dysfunction is attained via assays for TSH, free T3 (fT3) and free T4 (fT4) [40]. The distinction between 'free' and 'bound' thyroid hormone is that the latter is bound to plasma proteins – in the blood – and is therefore not biologically active [41]. Total hormone assays, which combine free and bound hormone levels, tend not to be diagnostically useful due to biological variations in the amounts of thyroid-binding proteins.

The magnitude of changes in TSH levels in response to thyroid hormone levels is markedly higher than the inverse, so blood TSH levels are considered the most reliable primary indicator for SCH [42]. TSH production varies throughout the day and across individuals [15, 16, 43], so a single abnormal TSH assay in isolation may not be sufficient for diagnosis [44]. As a result, it is recommended that positive tests should be repeated for confirmation [42]. The NICE guidelines, for instance, recommend a second test 3 to 6 months after the initial TFT [45].

The usual practice is to limit FT3 to the diagnosis of hyper-, rather than hypothyroidism, whereas TSH and FT4 measurements are the gold standards for diagnosing SCH [42, 45]. Therefore, it should be noted that FT4 assays may be affected by illness, medications such as anticonvulsants or lithium carbonate and factors that influence the circulating amounts of thyroid hormone-binding proteins [40].

Based on an abnormal test result, laboratory protocols may also warrant the measurement of thyroid peroxidase antibodies (TPOAb), previously linked to the progression from subclinical to overt hypothyroidism [15, 15, 36]. Table 1.2 shows the types of thyroid disorders detected using TFTs.

Table 1.2. *Patterns of thyroid dysfunction obtained from TFTs [40].*

TSH level	FT4 or FT3 level	Thyroid Disorder
High	Normal	Subclinical hypothyroidism
High	Low	Primary hypothyroidism
Low	High	Overt hyperthyroidism
Low	Normal	Subclinical hyperthyroidism
Low	Low	Central hypothyroidism
High	High	Central hyperthyroidism

1.2.6.1 Reference Range

Since it is a biochemical disorder, the diagnosis of SCH directly depends on what TSH levels fall in the reference range [46]. The ideal comparator for a thyroid hormone or TSH assay would be a prior measurement taken from the same individual, given that their concentrations are relatively stable for long periods of time. In the absence of an individual's previous assays, the best alternative is to apply population-based reference ranges, which are typically obtained from healthy people with no thyroid disease, irrespective of their demographic characteristics (e.g. age, ethnicity) [47]. However, it has been found that TSH distribution curves are not normally distributed but instead exhibit a left skew towards serum TSH levels higher than 4.5 mIU/L [47], as was demonstrated in the National Health and Nutrition Examination Survey (NHANES) III [48]. Possible explanations for this observation are the inclusion of individuals with early, undetected thyroid disease and the progressive physiological increase of TSH as people grow older [49–51].

Population-based reference ranges for TSH fail to account for the genetic differences that influence physiologically normal TSH levels across individuals. A recent study comparing genetically derived TSH ranges to population-based references found that almost 30% of study participants previously classified as having thyroid disorders were biochemically euthyroid based on their genetic profiles [52].

Furthermore, the sensitivity and specificity of lab assay equipment can bias the results of TSH measurements [53, 54] since they tend to vary depending on the manufacturers' chosen calibration and standardisation methods. Some assays contain antibodies which are capable of detecting TSH isoforms that are biologically inactive, therefore yielding misleading results [14, 55]. Laboratories also tend to apply manufacturer-provided reference ranges rather than adapting them for the local population [56, 57]. Therefore, it has been recommended that clinicians should view TSH ranges flexibly in light of assay bias [55, 58].

A study comparing TSH assays from two manufacturers, Roche Diagnostics and Abbot Laboratories, found that the lack of standardisation in the reference ranges provided by the companies resulted in diagnostic discordance [58]. Less than half of the samples (44%) had SCH according to both assays, though the Roche version was reported to have a narrower range [58]. Overdetection and overdefinition are the two underlying causes of overdiagnosis [59]; therefore, TFTs must be correctly interpreted.

1.2.6.2 Differential Diagnoses

Causes of transiently elevated TSH levels include obesity, pituitary gland tumours that secrete TSH, and kidney failure [15, 16]. It is also possible for the increase to be outside the reference range but non-pathological [15, 60], given that TSH reference ranges may differ due to physiological factors like age [48, 61, 62] and exogenous factors such as testing laboratory standards, even within the a single geographical region [63].

1.2.7 Treatment

Levothyroxine (LT4), a synthetic form of T4, is the drug of choice for treating SCH. Its T3 analogue, liothyronine (LT3), is not typically used but may be combined with levothyroxine for specific resistant cases [5, 64]. Even so, evidence supporting long-term combination therapy is limited and notably relates to mixed forms of hypothyroidism rather than directly to SCH [64, 65].

Furthermore, LT3 and LT4 may affect SCH patient physiology differently, such as the degree of weight loss observed [66]. After the medication is initiated, patients usually have periodic blood tests to check their TSH levels [67]. Figure 1.3 shows the potential pathways that patients could follow after SCH is diagnosed.

Given the characteristically elevated TSH levels, the goal of treating SCH is usually to normalise TSH to within the reference range [45], but achieving biochemical and clinical euthyroidism requires periodic monitoring and dose adjustment [68]. Even so, patients on long-term, high-dose levothyroxine treatment may have suppressed TSH levels (<0.03 mIU/L), which are associated with a higher risk of cardiovascular disease and fractures [69]. On the other hand, persistently high TSH levels while on treatment may be caused by poor treatment compliance, impaired absorption of levothyroxine or higher levothyroxine demand, for example, as a result of pregnancy or weight gain [70, 71].

The alternative to medication – typically for patients with mild SCH with few or no symptoms – is patient monitoring with repeated TFTs at regular intervals [72]. In case of progressively increasing TSH or worsening health, medication with levothyroxine may then be initiated when needed (Figure 1.3).

Clinical guidelines for the management of SCH vary in their recommendations (Table 1.3) but are united by the common theme of inadequate or weak evidence [9, 32, 73–75]. It has also been noted that some patient subgroups, young adults, for instance, are relatively underrepresented in SCH trials, so although more research is generally required, this need is especially pertinent for specific sub-populations [76].

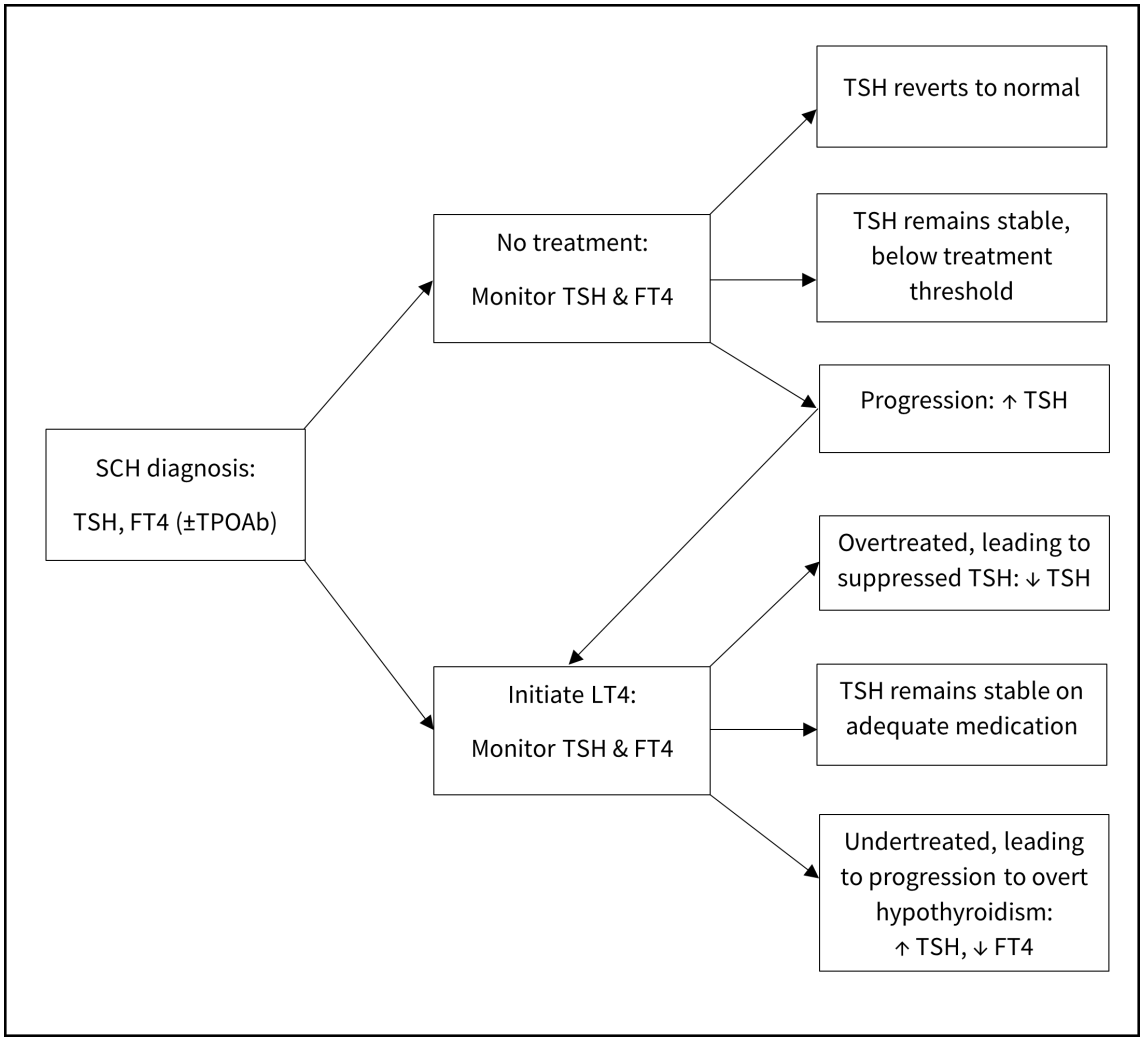


Figure 1.3. *Potential pathways following the diagnosis of SCH.*

Table 1.3. *Clinical guidelines for managing SCH (adapted from Ku et al. [77]).*

Guidelines	Consideration of LT4 treatment	Observation without LT4 treatment
ATA (2012) [78]	TSH >10 mIU/L, age <70 years	TSH <10 mIU/L, age >70 years
	TSH 4–10 mIU/L, age <65 years, symptoms (+)	TSH 4–10 mIU/L, age >65 years
ETA (2013) [32]	TSH >10 mIU/L, age <70 years	TSH <10 mIU/L symptoms (–), age <70 years
	TSH <10 mIU/L, age <70 years, symptoms (+)	TSH <10 mIU/L, age >70 years
	TSH <10 mIU/L, age >70 years, symptoms (+) or high cardiovascular (CV) risk	
LATS (2013) [79]	TSH >10 mIU/L	TSH 4.5-10 mIU/L, age > 65 years symptoms (–)
	TSH 4.5-10 mIU/L, age < 65 years, CV risk	
	TSH 4.5-10 mIU/L, age > 65 years symptoms (+)	
	TSH 4.5-10 mIU/L, TPOAb, autoimmune thyroiditis	
SEMDSA/ACE-SA (2015) [75]	TSH > 10 mIU/L	TSH 4-10 mIU/L, age > 65
	TSH 4-10 mIU/L, age < 65 years, CV risk factors, TPOAb, pregnancy, psychiatric illness, T2DM or symptoms (+)	
	TSH 4-10 mIU/L, TPOAb	
Clinical practice guideline (2017) [16]	TSH >10 mIU/L, age <70 years	TSH >10 mIU/L, age >70 years
	Especially, symptoms (+) or CV risk factors	
	6 months of LT4 treatment in the cases of TSH >4.5 and <7 mIU/L with symptoms	
	6 months of LT4 treatment in the cases with TSH >7 and <10 mIU/L, age <70 years	
	6 months of LT4 treatment in the cases with symptoms (+) regardless of age, CV risk factors, TPOAb	
NICE guideline (2018) [45]	TSH >10 mIU/L, age <70 years	TSH >10 mIU/L, age >70 years
	TSH 4–10 mIU/L, age <65 years, symptoms (+)	TSH 4–10 mIU/L, age >65 years
Clinical practice guideline (2019) [73]	Only women who are trying to become pregnant or patients with TSH >20 mIU/L	Almost all adults
KTA (2023) [80]	TSH >10 mIU/L, age <70 years	TSH 6.8–10 mIU/L, age <70 years
		All elderly patients

Note:
Abbreviations: LT4, levothyroxine; ATA, American Thyroid Association; TSH, thyroid-stimulating hormone; ACE-SA, Association of Clinical Endocrinologists of South Africa, ETA, European Thyroid Association; TPOAb, thyroid peroxidase antibody; NICE, National Institute for Health and Care Excellence; KTA, Korean Thyroid Association; LATS, Latin American Thyroid Society; SEMDSA, Society for Endocrinology, Metabolism and Diabetes of South Africa.

1.2.8 Screening

Screening for thyroid disease is not recommended because of insufficient evidence of its benefits [15] but also due to the high likelihood of false positives, for instance, in cases of transiently elevated TSH [76, 81]. A community study on the risk of abnormal TSH levels in subsequent tests found that among subjects with SCH, more than 60% of TSH measurements reverted to normal during the 5-year follow-up period [17], which implies that screening may result in the transitory inflation of SCH incidence. Furthermore, there is probably limited clinical value in identifying asymptomatic SCH cases [76] because of the inconclusive nature of the existing evidence on the effects of treatment.

1.3 Controversies associated with SCH

1.3.1 Overtesting

There has been a steady increase in the ordering of TFTs in primary care [82–84], with widely varying patterns reported between practices [38]. However, overtesting is not limited to GP practices and routinely occurs within hospitals too [39, 85, 86]. For SCH, 'overtesting' refers to ordering TFTs without sufficient cause [82, 84]. Overtesting may be considered akin to screening because both typically involve patients with no discernible symptoms.

There is evidence of the overuse of TFTs in primary care in Canada and the UK [38, 87], and its strongest predictor is, unsurprisingly, the frequency of GP-patient encounters. It stands to reason that a high number of GP visits would result in more investigations, including TFTs (which could result in an incidental SCH diagnosis).

Following an analysis of electronic medical records, over 30% of all adults attending primary care practices in Ontario, Canada, were found to have received at least one TFT over a 2-year follow-up period without clear clinical indications [84]. Crucially, less than 5% of the tests in the two years returned abnormal results, and the authors estimated that the number of unnecessary tests translated to approximately $25 million [84]. The extent of the issue is that nationwide campaigns have been designed to discourage the use of poorly evidenced tests, treatments and procedures, including TFTs. The Choosing Wisely Canada initiative [88] is an example that reduced the number of monthly TFTs in participating primary care practices by 11% over two years compared to non-participating practices [84].

The probable consequence of inappropriate test ordering is a corresponding increase in the number of diagnoses of SCH (overdiagnosis). Overusing diagnostic tests increases the probability of false positives [89], leading to 'a cascade of further investigations and unnecessary treatments' [38]. A UK study found that the frequency of monitoring TFTs for patients on LT4 did not align with clinical guidelines – those with disordered TSH or fT4 had longer intervals, whereas intervals for people with baseline TFTs within the reference range were too short, thus contributing to overtesting [90]. On balance, it should also be noted that overdiagnosis misrepresents true prevalence estimates [61] and can, therefore, have far-reaching implications.

1.3.2 TSH normal range

TSH measurement is the gold standard for detecting thyroid disorders because of the log-linear relationship between TSH and T4 [91], such that the magnitude of change in TSH levels is amplified and, therefore, easier to detect than more minor changes in T4 [22]. However, several factors influence TSH levels, including certain medications, diurnal variation, sex, age and ethnicity [15, 22, 92]. The definition of 'normal' also varies, given that the distribution of serum TSH skews to the right [13]. It has also been established that an individual's TSH levels may be fixed in the upper or lower part of the reference range [47].

Thus arises the call for age and sex-specific reference ranges for TSH [93], which would prevent the misdiagnosis of thyroid disorders and the burdens associated with managing lifelong conditions, including the initiation of treatment. In a Japanese study on the impact of changing the normal TSH ranges, Yamada et al. [49] reported that in the 60-69 age group, 62% of men and 78% of women with SCH were re-classified as euthyroid. However, the cohort was mainly composed of people having their annual check-ups, and the authors acknowledge that this limits the generalisability of their findings because they were likely of higher socioeconomic standing [49].

Most laboratories do not use age-specific reference ranges; this has been reported in the US [5, 48] and the UK [61]. The failure to apply different ranges depending on patient age may be explained by the fact that most current testing guidelines also recommend single cut-offs for TSH [18]. It is crucial to consider the age-specific differences because normal ranges influence the diagnosis, treatment and monitoring of SCH. Elderly patients tend to have high TSH levels, so it has been suggested that (i) their threshold for diagnosis should be raised to prevent overdiagnosis and (ii) their levothyroxine dosages should be lowered to prevent overtreatment [14, 94].

1.3.3 Clinical management

The main controversy in SCH relates to patient management – there are decades-long arguments for and against treatment [95–98] but no clear evidence of its harms and benefits. In the UK, the two most recent recommendations differ in their thresholds; one advises against the treatment of SCH with levothyroxine except in three groups (young adult patients aged 30 and under, those with severe symptoms or very high TSH levels of >20 mIU/L, and women attempting to conceive) [73]. The other guideline, by the National Institute for Health and Care Excellence (NICE), acknowledges the lack of evidence on the long-term effects of levothyroxine prescribed for SCH but asks physicians to consider treatment for adults with TSH of ≥ 10 mIU/L [74]. The guidance by Bekkering et al. [73] was based on a systematic review [99] that included insufficient data on the under-65 age group and patients with TSH over 10 mIU/L and was otherwise underpowered when the results of the TRUST trial [100] were excluded [101]. Therefore, the recommendations relating to these two subgroups were likely based on extrapolated data [98].

It is plausible that the lack of clarity on what justifies the initiation of therapy, be it the diagnosis of SCH or materially, a specific diagnostic attribute like the TSH threshold, increases the frequency of inconsistent clinical care. From a different viewpoint, the patient on levothyroxine must make lifestyle changes like maintaining adherence and

compliance with the medication. They will likely be on lifelong treatment [95] and require repeat appointments to monitor their TSH levels. The latter may translate to related expenses – prescription costs, lost time and wages – and potential secondary health complications [76]. Other possible concerns following the initiation of levothyroxine therapy include drug malabsorption and increased thyroid hormone metabolism or binding capacity, which negatively influence subsequent TFTs [102].
Also, given that some SCH patients' TSH levels spontaneously revert to normal without medication [36, 100], it may be that treatment is not always necessary. This possibility must be emphasised because it remains unclear if symptomatic and asymptomatic patients should be managed differently. Additionally, past trials on this topic have not evaluated the subsequent reversal of symptoms after initiating levothyroxine, nor any potential long-term effects of receiving treatment [103].

1.3.4 Overtreatment

The 2020/2021 Prescription Cost Analysis for England reports that levothyroxine was the third most dispensed medicine in the country, with the number having risen slightly compared to 2019 – 33.1 million and 33 million prescriptions, respectively [104]. Levothyroxine is prescribed for both overt hypothyroidism and SCH; however, it is more probable that the larger proportion of these figures is due to SCH because of its higher estimated prevalence in the general population or, possibly, overtesting and the use of lower TSH testing thresholds [14, 64, 105].
Overuse of levothyroxine may lead to adverse effects such as heart failure and bone density changes [63, 69, 106]. It is also understood that not all patients on treatment have TSH levels within the normal range [28], although this may be due to anomalous TFT results [102] or incorrect levothyroxine dosages [94]. A study of thyroid hormone users aged 65 years and over showed that many SCH patients were overtreated with levothyroxine to the point of hyperthyroidism [94]. As such, over-replacement is another potential concern fuelled by the insufficiency of definitive evidence of the clinical importance of treating SCH.

1.3.5 Consequences of SCH

Evidence on the long-term adverse effects of SCH is mostly unclear – it has previously been associated with cardiovascular complications [10, 16, 107, 108], frailty fractures [69], cognitive dysfunction, anxiety and depression [109, 110]. However, these studies were limited in their generalisability because of small sample sizes and varying diagnostic or therapeutic reference ranges. More significantly, other studies with similar limitations had negative findings [111–114]. Some studies have found no link between SCH and poor health-related quality of life [11, 23], but others reported more total symptoms among SCH patients than those with normal thyroid function [28].

1.4 Rationale for this Thesis

Uncertainty regarding the management of SCH, particularly the pleiotropic effects of levothyroxine for SCH outcomes, poses a challenge to clinicians and patients. The most

remarkable indicator of this issue is the disparities between clinical guidelines. More research on this topic is thus required to facilitate informed decision-making by healthcare practitioners.

An additional but critical concern is the potential for steadily increasing SCH diagnoses due to overtesting, compounded by the relative insufficiency of evidence on long-term treatment effects. This interplay of factors could result in more SCH patients receiving suboptimal clinical care over lengthy durations. From the perspective of healthcare management, overtesting, overdiagnosis, and overtreatment may all contribute to a massive waste of limited resources. Furthermore, considering the paucity of conclusive evidence [15, 16], the chances of new, adequately-powered randomised controlled trials (RCTs) on SCH to assess long-term outcomes are probably low.

1.5 Aims & Objectives

1.5.1 Thesis aim

The management options for SCH are that patients are either: (i) started on treatment (levothyroxine) or (ii) followed up with no pharmacological intervention. The overarching aim of this PhD was to investigate the impact of these strategies on long-term clinical outcomes. By comparing the impact of treatment and non-treatment on long-term health, this thesis will add to the evidence base of whether patients benefit from treatment.

1.5.2 Thesis objectives

1. To review the existing literature on the impact of SCH and its treatment on long-term clinical outcomes.
2. To characterise the epidemiology and clinical management of SCH using electronic health record (EHR) data.
3. To investigate the impact of treating SCH with levothyroxine on long-term clinical endpoints using routinely collected health data.

1.6 Thesis structure

This thesis is presented as a portfolio of research articles in the format required by the University of St Andrews. There are four papers – three have been published in peer-reviewed journals and are incorporated as such. The last manuscript is in publishable format – submission is planned to follow the completion of this thesis. Because of the portfolio-based format, there is an element of duplication throughout because all the articles relate to the same topic.

The relatively large number of published evidence syntheses on SCH precluded conducting another conventional systematic review. Hence, to avoid 'duplication of effort' [115], umbrella review methodology was selected instead for examining the evidence base. This is presented in Chapter 2 in its published format.

Chapter 3 contains the published article of a descriptive study based on electronic health records in the SAIL Databank. This chapter describes the cases of SCH in Wales between January 2000 and December 2021 to provide additional real-world context on this topic.

RCTs are considered the gold standard for assessing treatment effects [116]. However, over and above project resource limitations, emulation of a target trial was deemed an appropriate alternative to a randomised trial because the framework has been validated for causal inference [116, 117]. The trial emulation forms Chapter 4.

In Chapter 5, I summarise the findings from the umbrella review, descriptive study and trial emulation. I also discuss the relevance of these results, the thesis strengths and limitations, and potential areas for future research on this topic.

2

Umbrella review

2.1 Introduction

Research is best informed by identifying what is already known on a topic. This chapter describes an umbrella review that was performed to explore the published literature, specifically systematic reviews and meta-analyses, on how the management of SCH is associated with various long-term outcomes. It therefore addresses the first objective of this thesis.

The review protocol was registered on PROSPERO [118] and published in BMC Systematic Reviews [119]. Searches were conducted on electronic databases, and the identified reviews were screened, their quality assessed, and their findings collated. The review manuscript was then published in PLoS ONE [120]. Details of the two articles are as follows:

1. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, McCowan C. Management strategies for patients with subclinical hypothyroidism: a protocol for an umbrella review. Systematic reviews. 2021 Dec;10:1-6. https://doi.org/10.1186/s13643-021-01842-y

2. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C (2022) The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review. PLOS ONE 17(5): e0268070. https://doi.org/10.1371/journal.pone.0268070

 Supplementary materials are included in Appendix I.

2.2 Methods

2.2.1 Published protocol

Bauer *et al. Syst Rev* (2021) 10:290
https://doi.org/10.1186/s13643-021-01842-y

Systematic Reviews

PROTOCOL Open Access

Management strategies for patients with subclinical hypothyroidism: a protocol for an umbrella review

Brenda S. Bauer*, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal and Colin McCowan

Abstract

Background: Subclinical hypothyroidism is a thyroid disorder diagnosed from the laboratory blood test results of otherwise asymptomatic patients. It has been associated with poor cardiovascular outcomes, mortality and progression to overt thyroid hormone deficiency. Current guidelines on the management of subclinical hypothyroidism differ because of conflicting evidence on long-term treatment benefits. Even though there are several existing systematic reviews on its clinical outcomes, no definitive conclusion has been reached yet. As such, a new synthesis could help provide more insight and consensus on this topic. To this purpose, this umbrella review will evaluate and synthesise current evidence on the long-term clinical outcomes of the different management strategies for subclinical hypothyroidism.

Methods: This is a protocol for an umbrella review on the management strategies for subclinical hypothyroidism. We will conduct literature searches in multiple electronic databases (from inception onwards), namely MEDLINE, EMBASE, Scopus, Web of Science, Cochrane Database of Systematic Reviews, JBI Evidence Synthesis, Epistemonikos database, PDQ Evidence and the PROSPERO register. There will be no restriction on the date or language of publication. Additional material will be identified through grey literature searches and citation chaining. Review inclusion criteria will be patients with subclinical hypothyroidism, receiving treatment or monitoring, no restrictions on the comparators used and with cardiovascular events, frailty fractures, quality of life and all-cause mortality as primary outcomes of interest. Two reviewers will independently screen all citations, full-text articles and abstract data on a pre-piloted form in duplicate. Methodological quality (or bias) of included studies will be appraised using AMSTAR-2. Any conflicts that arise will be resolved through discussion or involving a third reviewer. A narrative synthesis will be provided with information presented in the main text and tables to summarise and explain the characteristics and findings of the included reviews. Even so, it is not expected that a meta-analysis will be performed due to review variability. Study limitations and methodological quality assessments will also be reported to provide context for the overall summary of evidence.

Discussion: This review will provide a comprehensive summary of the effects of the pharmacological and non-pharmacological management of subclinical hypothyroidism on specific long-term clinical outcomes. It is anticipated that the findings of this umbrella review will aid in the development of consensus-based clinical recommendations for subclinical hypothyroidism, as well as highlight areas for future research. Review findings will be disseminated primarily through peer-reviewed publications.

*Correspondence: bsb1@st-andrews.ac.uk
School of Medicine, University of St Andrews, North Haugh, St Andrews KY16 9TF, UK

Systematic review registration: PROSPERO CRD42021235172

Keywords: Subclinical hypothyroidism, Levothyroxine, Systematic review, Review of systematic reviews, Umbrella review

Background

Subclinical hypothyroidism (SCH) is a disorder of the thyroid gland in which blood levels of free circulating hormones are normal, but those of thyroid-stimulating hormone (TSH)—which stimulates thyroid hormone production—are elevated. It can only be diagnosed through laboratory tests, and diagnosed patients are typically asymptomatic [1, 2]. As such, the detection of SCH is often incidental [3, 4], and in approximately 2 to 5% of patients, SCH has been found to progress to overt hypothyroidism [2]. SCH has also been linked to an increased risk of cardiovascular disease and mortality [5, 6], frailty fractures [7], cognitive dysfunction, anxiety and depression [8]. Crucially, however, these associations are based on differing conclusions from observational studies and small randomised trials with relatively brief follow-up periods [9].

SCH is treated through the replacement of thyroid hormone using the drug levothyroxine [10]. Even so, the decision to begin replacement therapy has long been controversial because of conflicting findings on whether treatment is beneficial for long-term outcomes [1, 2]. Recently published guidelines on the management of SCH differ in their recommendations, as a result. One evidence-based guideline recommends applying a TSH threshold of 10 mU/L for prescribing levothyroxine because of potential long-term benefits such as cardiovascular outcomes and symptom improvement [11, 12]. On the other hand, Bekkering et al. [13] considered a systematic review of 21 trials that found minimal to no evidence of clinical benefit from replacing thyroid hormones in SCH [14]. In response, a strong recommendation was issued against treatment for most adult patients, except patients with TSH levels greater than 20 mIU/L and pregnant women [13].

It is widely acknowledged that inadequate research has been conducted on the long-term clinical outcomes of managing SCH, especially as inconsistencies remain in the findings of the studies that have been performed to date [1, 9, 15–17]. Since thyroid hormone replacement is a lifelong treatment, it is vital to investigate how levothyroxine affects health in subclinical disease. Equally important are the clinical effects of follow-up with no treatment—for patients who do not meet treatment thresholds, for example—in determining the optimal timing of treatment, as well as the suitability of certain patient groups to receive treatment.

The umbrella review approach is well-suited to the synthesis of a body of contentious evidence, as it allows for a rigorous and systematic assessment of the literature [18]. We will employ this methodology to summarise and compare systematic reviews of various clinical outcomes of the management strategies of subclinical hypothyroidism, which may be either to prescribe treatment or to monitor the patient with no pharmacological intervention. Specifically, the review questions are:

Q1: What is the impact of levothyroxine treatment on long-term clinical outcomes for patients with subclinical hypothyroidism?

Q2: What is the impact of follow-up without treatment on long-term clinical outcomes for patients with subclinical hypothyroidism?

Methods

Protocol development

This protocol was registered in the PROSPERO register [19] as CRD42021235172. The methods described below are based mainly on the 'Umbrella Reviews' chapter of the JBI Manual for Evidence Synthesis [20], though some elements—protocol length, referencing style, critical appraisal and data collection tools, in particular—have been adapted for our purposes. The Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) guidelines [21, 22] have been followed in reporting this protocol, for which a completed checklist is provided as an additional file [see Additional file 1].

Inclusion criteria

It is anticipated that all the systematic reviews obtained for this review will have clearly defined inclusion and exclusion criteria, in keeping with systematic review norms and guidelines. Therefore, it will be possible to apply the following criteria while selecting the relevant literature. A summary of the screening criteria is presented in Table 1.

Participants

The population of interest is restricted to patients diagnosed with subclinical hypothyroidism, regardless of age, setting and the country in which the studies took place. Reviews relating solely to pregnant women, children and adolescents will be excluded because these are special patient groups with additional clinical considerations.

Table 1 Review inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Participants	Patients diagnosed with subclinical hypothyroidism, no demographic or etiological restrictions	Pregnant women Children and adolescents Diagnosis of overt hypothyroidism
Intervention	Management of SCH, i.e. treatment using levothyroxine or follow-up without treatment	No reported treatment status
Comparator	No restriction on comparison or control groups	
Outcomes	Cardiovascular outcomes Cerebrovascular outcomes Quality of life measures Frailty fractures All-cause mortality Other reported outcomes (secondary)	No reporting of any of the primary outcomes of interest
Study design	Quantitative systematic reviews and meta-analyses of empirical research	Any other study types (e.g. narrative reviews, scoping reviews, qualitative syntheses)
Review characteristics	Articles in any language Any period of study or date of publication	

Intervention

Inclusion is restricted to systematic reviews of studies involving the management of SCH, whether (i) using levothyroxine for treatment or (ii) follow-up with no treatment. Studies that do not report the treatment status of participants will not be included.

Comparator

Any comparison groups will be eligible for inclusion, depending on whether one was used in the synthesis. Therefore, reviews that compare the effects of treatment against no treatment will be included, as well as those that report findings from only one of the two strategies.

Outcomes

The primary clinical outcomes of interest are cardiovascular (e.g. heart disease, heart failure, peripheral vascular disease), cerebrovascular (i.e. stroke), quality of life measures (e.g. Underactive Thyroid-Dependent Quality of Life score, Short-Form 36, Thyroid-Related Quality-of-Life Patient-Reported Outcome Measure), frailty fractures and all-cause mortality. Secondary outcomes (e.g. improvements in clinical symptoms, cognitive dysfunction) will also be included if reported in addition to the above.

Study design

Only quantitative systematic reviews and meta-analyses of empirical studies will be eligible for inclusion, regardless of whether the studies were randomised clinical trials or observational. Narrative and scoping reviews, as well as purely qualitative reviews, will be excluded during study selection. Any systematic reviews that include theoretical or opinion articles will also be considered ineligible.

Primary studies will not be considered, even when gaps are identified in the evidence within included systematic reviews.

Review characteristics

There will be no limitations on the year of publication or study period to allow for temporal comparisons in study findings. Publications in languages other than English will be included in the first instance; if translation is not possible, they will be excluded, but their details reported.

Information sources and search strategy

Comprehensive searches will be carried out on multiple electronic databases: MEDLINE, EMBASE, Scopus, Web of Science, Cochrane Database of Systematic Reviews, JBI Evidence Synthesis, Epistemonikos database, PDQ Evidence and the PROSPERO register, from inception onwards. There will be no additional filters based on the date or language of publication.

We will use controlled vocabularies and search terms directly related to the review questions such as 'treatment', 'levothyroxine' and 'subclinical hypothyroidism' which will be modified, as needed, to account for database-specific differences. Search filters will be applied to retrieve only systematic reviews. The MEDLINE search strategy, developed with the assistance of an academic librarian, is shown in Additional file 2.

The reference lists of selected reviews will also be checked for eligible syntheses (backward citation chaining) and Google Scholar used for forward citation chaining. A search will also be performed for grey literature,

on the WorldCat and Open Grey databases and Internet search engines.

These searches will be updated in the later stages of the review (i.e. during data synthesis) to identify any relevant systematic reviews that will have been published in the interim.

Study selection

All the references retrieved from the searches will be imported to EndNote X9 [23] to remove duplicate records. The remaining citations will then be imported to Covidence [24] and screened independently by a set of two reviewers in duplicate—first by titles and abstracts—against the inclusion and exclusion criteria described above. In situations where it is impossible to identify inclusion from the title and abstract alone, these articles will progress to full-text review.

Afterwards, articles that pass through the initial screening will be obtained and read in full to determine their eligibility for inclusion. Any disagreements in study selection will be resolved through discussion or the involvement of a third reviewer to reach a consensus. Updated systematic reviews will be included but treated as a single study to prevent duplication during data extraction. All decisions at this stage will be recorded and presented in a PRISMA flow diagram in subsequent reports.

Data extraction

A pair of reviewers working independently will use a standardised, pre-piloted form to extract data in duplicate. Specifically, data will be collected on first author, year of publication, reported a protocol, objective(s), reported strategies to search literature, number of databases searched and date of last search, any restrictions (e.g. language, geographic or date), inclusion/exclusion criteria, intervention(s) of interest and comparators, patient population, main outcomes of interest, type of study designs included (e.g. randomised controlled trials, observational studies or both), number of included studies, number of studies reporting data for meta-analyses, effect metric(s) reported (e.g. risk ratio), methods to assess study risk of bias, statistical methods to combine studies, summary meta-analytic estimates including heterogeneity measures, additional analyses (e.g. subgroup analysis or sensitivity analysis), metabias assessment (e.g. publication bias across studies), funding source and conflicts of interest. Where presented, data on the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) rating for individual systematic reviews will also be collected. A complete list of fields to be extracted from included reviews is included in Additional File 3.

Disagreements arising from data extracts will be resolved by discussion with a third reviewer to reach a consensus. Where necessary, review authors will be contacted for further information on incomplete or missing data.

Quality assessment

The critical appraisal of all selected systematic reviews will be conducted in tandem with data extraction, using the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) tool. This checklist was designed to assess the methodological quality of systematic reviews of randomised trials [25] and is currently in its second version, AMSTAR-2 [see Additional File 4]. In recognition of the increasing number of systematic reviews incorporating data from non-randomised and observational studies, the original checklist was updated, published and subsequently validated [26, 27]. Syntheses are judged on 16 domains, including the suitability of the research question and inclusion criteria, the search strategy, the characteristics and critical appraisal of included studies and publication bias. Most domains are rated either 'Yes' or 'No' though some have the additional option of 'Partial Yes'.

Discrepancies in the independent assessments made by each pair of reviewers will be resolved by discussion with a third reviewer to reach a consensus. The results of the quality assessments will be applied in the overall synthesis and presentation of findings so that it will also be possible to compare the included reviews by methodological quality. However, the primary studies from included systematic reviews will not be evaluated individually.

Data synthesis

Review findings will be synthesised narratively, as it is anticipated that there will be several differences in inclusion criteria, methods of synthesis and outcome measures. Overall outcome measures will be presented in tabular form, accompanied by detailed descriptions of review characteristics and quality assessments.

If there are sufficient data from the included systematic reviews, patient characteristics (e.g. age, sex) and methodological differences (e.g. search strategies, definitions of clinical outcomes) will be used to stratify the findings, to allow for further comparisons in the management options for SCH based on these criteria.

There is a considerable burden involved in performing a meta-analysis of existing systematic reviews, given the likelihood of primary studies being counted more than once [28]. This is because of the complexity of taking each review apart and then combining the results of several individual studies, many of which are likely to have different review questions and inclusion criteria. As such,

it is anticipated that a meta-synthesis of included meta-analyses will not be performed; key statistical data will only be summarised.

Confidence in cumulative evidence

The GRADE ratings described within the included systematic reviews will be reported in this umbrella review. However, it is anticipated that not all studies will report these measures, especially older syntheses published prior to the first GRADE guidelines [29]. For such reviews, no new GRADE assessments will be conducted because they involve an assessment of primary studies. As such, this is beyond the scope of this umbrella review.

Discussion

This is a protocol outlining the processes through which an umbrella review will be performed. It is anticipated that this review of systematic reviews will be useful in summarising and comparing the syntheses of evidence on the management of SCH. As such, its findings may either aid in the development of, or reinforce future evidence-based clinical guidelines. Furthermore, the review will be useful for the identification of any potential biases or gaps that could explain the contradictions in the literature on this topic. Knowledge gaps identified in the literature can also inform future studies and systematic reviews.

The key strength of this overview will be to provide a comprehensive summary of current evidence on the management of SCH through the application of robust and established methods to source, select, appraise and synthesise existing systematic reviews. This information will be of interest to researchers, clinicians and patients with SCH seeking a high-level overview of the evidence; this will be the first umbrella review on this topic, to the authors' knowledge.

This type of evidence synthesis—the umbrella review—though useful, is also subject to several limitations. First, inclusion in this review is restricted to systematic reviews, but additional empirical studies on the same topic are likely to have since been published. These new findings would, therefore, not be captured in the scope of this secondary synthesis. For this reason, all searches will be updated at least once, towards completion of the review.

Another potential challenge when applying meta-review methodology is overlap in primary research. Study results included in more than one systematic review can cause misleading findings through a multiplier effect because a specified set of findings would be counted more than once. Therefore, a crucial element of data extraction and the subsequent synthesis will be to identify all primary studies and report all instances of overlap.

A third limitation is the differences in inclusion criteria between included studies that impede more quantitative forms of synthesis when conducting an overview. However, given the aim of this review of systematic reviews to collate and summarise all the synthesised literature on the clinical management of SCH, a descriptive and tabular presentation of findings should suffice.

Protocol amendments

Any amendments to this protocol in the carrying out of this umbrella review will be documented and reported in both the PROSPERO register and any subsequent publications.

Dissemination plans

The findings of this umbrella review will be disseminated through publication in peer-reviewed journals, via social media networks and relevant conferences.

Abbreviations

AMSTAR: A MeaSurement Tool to Assess systematic Reviews; GRADE: Grading of Recommendations, Assessment, Development and Evaluation; IU: International units; PICO: Population, Intervention, Comparator and Outcomes; PRISMA: Preferred Reporting Items for Systematic review and Meta-Analyses; PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols; SCH: Subclinical hypothyroidism; TFT: Thyroid Function Test; TSH: Thyroid-stimulating hormone.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13643-021-01842-y.

Additional file 1. PRISMA-P checklist.

Additional file 2. MEDLINE search strategy.

Additional file 3. Data extraction template.

Additional file 4. AMSTAR-2 checklist.

Acknowledgements

We thank Vicki Cormie, a senior librarian at the University of St Andrews, for her assistance in developing the search strategy for this review.

Authors' contributions

BSB led the design of the protocol, search strategy, drafting and revising of the manuscript. CM, AA and UA provided oversight and critically reviewed and revised the manuscript. All the authors read and approved the final manuscript. CM is the guarantor of this review.

Funding

This umbrella review will be conducted as part of a PhD studentship funded by the School of Medicine, University of St Andrews. The University of St Andrews played no role in the design and development of this protocol.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 12 February 2021 Accepted: 21 October 2021
Published online: 01 November 2021

References

1. Javed Z, Sathyapalan T. Levothyroxine treatment of mild subclinical hypothyroidism: a review of potential risks and benefits. Therapeutic Adv Endocrinol Metab. 2016;7:12–23.
2. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. Jama. 2004;291:228–38.
3. Chubb S, Davis W, Inman Z, Davis T. Prevalence and progression of subclinical hypothyroidism in women with type 2 diabetes: the Fremantle Diabetes Study. Clin Endocrinol. 2005;62:480–6.
4. Schübel J, Feldkamp J, Bergmann A, Drossard W, Voigt K. Latent hypothyroidism in adults. Deutsches Ärzteblatt International. 2017;114:430.
5. Rodondi N, Den Elzen WP, Bauer DC, Cappola AR, Razvi S, Walsh JP, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. Jama. 2010;304:1365–74.
6. Razvi S, Weaver JU, Butler TJ, Pearce SH. Levothyroxine treatment of subclinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. Arch Internal Med. 2012;172:811–7.
7. Flynn RW, Bonellie SR, Jung RT, MacDonald TM, Morris AD, Leese GP. Serum thyroid-stimulating hormone concentration and morbidity from cardiovascular disease and fractures in patients on long-term thyroxine therapy. J Clin Endocrinol Metab. 2010;95:186–93.
8. Kong WM, Sheikh MH, Lumb PJ, Freedman DB, Crook M, Doré CJ, et al. A 6-month randomized trial of thyroxine treatment in women with mild subclinical hypothyroidism. Am J Med. 2002;112:348–54.
9. Cooper DS, Biondi B. Subclinical thyroid disease. Lancet. 2012;379:1142–54.
10. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocrine Reviews. 2008;29:76–131.
11. Vasileiou M, Gilbert J, Fishburn S, Boelaert K. Thyroid disease assessment and management: summary of NICE guidance. BMJ. 2020;368:m41.
12. UK National Guideline Centre. Thyroid disease: assessment and management. 2019.
13. Bekkering G, Agoritsas T, Lytvyn L, Heen A, Feller M, Moutzouri E, et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. BMJ. 2019;365:l2006.
14. Feller M, Snel M, Moutzouri E, Bauer DC, de Montmollin M, Aujesky D, et al. Association of thyroid hormone therapy with quality of life and thyroid-related symptoms in patients with subclinical hypothyroidism: a systematic review and meta-analysis. Jama. 2018;320:1349–59.
15. Baumgartner C, Blum MR, Rodondi N. Subclinical hypothyroidism: summary of evidence in 2014. Swiss Med Weekly. 2014;144:w14058.
16. Biondi B. Natural history, diagnosis and management of subclinical thyroid dysfunction. Best Pract Res Clin Endocrinol Metab. 2012;26:431–46.
17. Chrysant SG. The current debate over treatment of subclinical hypothyroidism to prevent cardiovascular complications. Int J Clin Practice. 2020;74:e13499.
18. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evidence-based Mental Health. 2018;21:95–100.
19. Centre for Reviews and Dissemination. PROSPERO International prospective register of systematic reviews: University of York; 2021 [cited 2021 15/01]. Available from: https://www.crd.york.ac.uk/prospero/.
20. Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Chapter 10: Umbrella reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020.
21. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015: elaboration and explanation. Brit Med J. 2015;349:g7647. https://doi.org/10.1136/bmj.g7647.
22. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Systematic Reviews. 2015;4:1. https://doi.org/10.1186/2046-4053-4-1.
23. The EndNote Team. EndNote. EndNote X9 ed. Philadelphia: Clarivate; 2013.
24. Veritas Health Innovation. Covidence systematic review software Melbourne, Australia2013. Available from: www.covidence.org.
25. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.
26. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
27. Lorenz RC, Matthias K, Pieper D, Wegewitz U, Morche J, Nocon M, et al. A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. J Clin Epidemiol. 2019;114:133–40.
28. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Med Res Methodol. 2011;11:1–6.
29. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64:383–94.

Publisher's Note

2.3 Published article

PLOS ONE

RESEARCH ARTICLE

The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review

Brenda S. Bauer*, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan

Population and Behavioural Science Division, School of Medicine, University of St Andrews, St Andrews, Scotland, United Kingdom

* bsb1@st-andrews.ac.uk

OPEN ACCESS

Citation: Bauer BS, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C (2022) The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review. PLoS ONE 17(5): e0268070. https://doi.org/10.1371/journal.pone.0268070

Editor: Mahmoud Nassar, Icahn School of Medicine at Mount Sinai / NYC Health + Hospitals Queens, New York, UNITED STATES

Received: February 25, 2022

Accepted: April 21, 2022

Published: May 19, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0268070

Abstract

Aim

This umbrella review summarises and compares synthesised evidence on the impact of subclinical hypothyroidism and its management on long-term clinical outcomes.

Methods

We conducted comprehensive searches on MEDLINE, EMBASE, Scopus, Web of Science, Cochrane Database of Systematic Reviews, JBI Evidence Synthesis, the PROSPERO register, Epistemonikos Database and PDQ Evidence from inception to February and July 2021 using keywords on subclinical hypothyroidism, treatment with levothyroxine, monitoring and primary outcomes (all-cause mortality, cardiovascular events, stroke, frailty fractures and quality of life). Only systematic reviews and meta-analyses on adult patient populations were considered. Study selection, data extraction and quality appraisal using AMSTAR-2 were done independently by two reviewers and discrepancies were resolved through discussion. Overlap across the selected reviews was also assessed, followed by a narrative synthesis of findings.

Results

A total of 763 studies were identified from literature searches; 20 reviews met inclusion criteria. Methodological quality ratings were high (n = 8), moderate (n = 7), and low (n = 5), but no reviews were excluded on this basis. Though there was slight overlap across all reviews, some pairwise comparisons had high corrected covered area scores. Compared to euthyroidism, untreated subclinical hypothyroidism was associated with a higher risk of cardiovascular events or death if Thyroid Stimulating Hormone was above 10mIU/L at baseline. Treatment was associated with a lower risk of death from all causes for patients younger than 70 years and possibly better cognitive and quality of life scores than untreated individuals. Evidence on the risk of strokes and fractures was inconclusive.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This umbrella review was conducted as part of a PhD project (BSB) funded by the School of Medicine, University of St Andrews, Scotland, United Kingdom. Publication fees were funded under an Open Access membership agreement between the University of St Andrews Library and PLOS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

In the long term, treatment of subclinical hypothyroidism may be beneficial for some patient groups. However, the findings of this review are negatively impacted by the relative sparseness and poor quality of available evidence. Additional large and adequately powered studies are needed to investigate this topic further.

Systematic review registration

PROSPERO (CRD42021235172)

Introduction

Subclinical hypothyroidism (SCH) is characterised by elevated Thyroid Stimulating Hormone (TSH) levels in contrast to free thyroid hormone–usually thyroxine/T4 –within the reference range [1–3]. The leading cause of SCH is Hashimoto's thyroiditis, a chronic autoimmune disorder that affects more women than men [2]. Prevalence varies worldwide but has been estimated to be between 4% to 9%, increasing with age to more than 20% for women over 60 years old [2, 4]. This broad range can also be explained by differences in race, dietary intake of iodine and diagnostic cut-offs for SCH [2]. Patients usually exhibit few, if any, symptoms, so it is common for SCH to be detected incidentally from a routine blood test panel [5, 6].

The reference ranges used for measurements of thyroid hormones and TSH vary between laboratories because they are highly dependent on the reference population [7]. Nonetheless, a distinction is sometimes made between mild and severe SCH with a TSH measurement of 10mIU/L as the cut-off [8, 9]. Approximately 60% of cases with mild SCH revert to normal TSH levels over time [3, 9]. Furthermore, depending on the initial severity of their condition, female SCH patients and those that are antithyroid peroxidase antibody-positive are more likely to develop overt hypothyroidism [4]. This progression occurs in around 2–4% of cases per year [2, 3].

Measurement of TSH and thyroid hormone levels is achieved through thyroid function tests (TFTs) which are frequently ordered unnecessarily without medical indications [10–12]. One of the main pitfalls that could likely result from inappropriate TFTs is that more asymptomatic patients are diagnosed as having SCH. Following diagnosis, there are two options for the management of SCH, thyroid replacement therapy with levothyroxine or follow-up without prescribing medication [2, 3]. Even so, a reasonable expectation for the latter is that treatment would be initiated in the event of a patient's worsening state, provided that the progression can be attributed to SCH. Regardless of the strategy, patients require periodic blood tests to monitor TSH levels for increasing severity or improvement of SCH [1, 3, 13].

The management of SCH is controversial–there has been no definitive evidence on the benefits of replacing thyroid hormones, especially the long-term clinical consequences. This is partly because few adequately powered randomised trials have investigated this topic [4, 14], the most notable being the Thyroid Hormone Replacement for Untreated Older Adults with Subclinical Hypothyroidism (TRUST) trial [15]. The TRUST Study Group reported that in their trial with 737 adults over 65 years old, levothyroxine treatment did not improve patient symptoms nor lower the risk of cardiovascular events and fractures [15]. On the other hand, other smaller trials and observational studies have linked treatment of SCH to improved patient outcomes [2, 4, 8]. Also, current UK clinical guidelines for the management of SCH

differ in their recommendations for treatment thresholds and exclusions [14, 16]; these differences can be directly ascribed to inconsistencies in the existing evidence base.

Based on a systematic review by Feller et al. [17], a clinical guideline panel found no evidence to recommend thyroid hormone replacement for SCH patients, except for those with TSH levels above 20mIU/L and women that are pregnant or trying to conceive [14]. For the outcomes they considered, for example, quality of life, cognitive function and cardiovascular events, the panel found no vital difference between treated and untreated groups, irrespective of patient age. Moreover, they noted the issue of practicality regarding medication–patients require long-term treatment and follow-up and even risk developing hyperthyroidism in case of overuse [14]. In contrast, the latest NICE guidance recommendation is for physicians to consider treating adults with TSH of $\geq$ 10 mIU/L to improve SCH patient outcomes [16]. The reviewing committee found little evidence on SCH treatment but emphasised that additional factors–such as the presence of symptoms–should be considered, over and above TSH levels [13].

Therefore, an umbrella review was performed to collate and compare existing literature on the long-term effects of SCH treatment and follow-up with no medication. Umbrella reviews–also called overviews, meta-reviews or reviews of reviews [18]–are ideally suited to exploring discrepancies in the literature by allowing for a broader scope of inquiry than a typical systematic review [18, 19]. The review questions of interest were: (i) what is the impact of levothyroxine treatment on patient outcomes in subclinical hypothyroidism? and (ii) what is the impact of monitoring without treatment on clinical outcomes for patients with subclinical hypothyroidism? Rather than restrict the focus of this overview to a direct comparison of these management strategies, we sought also to identify what is known for either option.

Methods

To a large extent, umbrella reviews are conducted similarly to typical systematic reviews. However, the critical difference between these methods is that the former use existing systematic reviews and meta-analyses as the units of synthesis [18]. These will subsequently be referred to as 'primary reviews' in this paper, in contrast to 'primary studies', the empirical studies included in the systematic reviews. The protocol for this overview was registered on PROSPERO (CRD42021235172) and the methods followed have previously been described in detail [20]; there were no deviations from the registered protocol. The reporting of this overview follows a checklist developed for overviews of systematic reviews based on recommendations from existing guidelines [21].

Search strategy

Comprehensive searches were performed on multiple databases from inception to February 2021, namely MEDLINE, EMBASE, Scopus, Web of Science, Cochrane Database of Systematic Reviews, JBI Evidence Synthesis, the PROSPERO register, Epistemonikos Database and PDQ Evidence. The key search terms used included 'subclinical hypothyroidism', 'monitoring', 'treatment' and 'levothyroxine', in both free text and subject headings; the search syntax was modified to match the different databases. The searches were updated in July 2021, during the latter stages of data extraction, to identify any systematic reviews and meta-analyses that had been published as the review was in progress. The MEDLINE search strategy is provided in S1 Appendix. No additional date or language filters were applied. Grey literature searches were also performed, and the reference lists of eligible studies were scanned to identify other potentially relevant primary reviews.

Eligibility criteria

The following eligibility criteria were considered to determine inclusion in this umbrella review. Only systematic reviews and meta-analyses on SCH, either in part or as a whole, were considered, irrespective of whether the primary review included randomised trials or observational studies. Generally, all primary reviews had to report on the clinical outcomes of adult patients (>18 years old) with SCH, regardless of the diagnostic thresholds that were initially applied, for example, reference ranges for TSH and thyroid hormones. Patients with overt hypothyroidism were excluded, as were children and pregnant women, whose thyroid hormone requirements differ from the rest of the population.

The two eligible interventions were: (i) treatment with levothyroxine, and (ii) patient follow-up without medication. Another essential requirement was that the treatment status of patients was reported in the systematic reviews, such that it would be possible to distinguish between treated and untreated groups. There were no additional restrictions on study comparators and settings.

The primary outcomes of interest were all-cause mortality, defined as the death of patients with SCH, irrespective of the cause, at least 12 months from baseline or the start of follow-up; cardio- and cerebrovascular outcomes such as heart failure, arrhythmias, stroke, peripheral vascular disease, coronary heart disease; quality of life as measured using suitable instruments (or otherwise described as 'symptoms' particularly in older publications); and, frailty fractures, defined as fractures resulting from low-impact trauma, usually due to pre-existing disease. Other long-term clinical outcomes reported in the included systematic reviews, for example, cognitive function, were considered secondary outcomes.

Study selection

References retrieved from the searches were imported into Covidence (www.covidence.org/) and initially screened in duplicate for eligibility by title and abstract. After that, the full texts of selected primary reviews were obtained and read independently by pairs of reviewers who assessed each paper against the selection criteria. When needed, primary review authors were contacted to provide additional information.

Data extraction

A data extraction form was developed on Covidence and piloted by two reviewers. This form was used to extract information on citation details, primary study selection criteria, search parameters, selection and quality assessment methods and primary review findings relating to the outcomes of interest for this umbrella review. Where provided, effect estimates were extracted alongside their 95% confidence intervals and the treatment status of the assessed group(s). This process was done by two reviewers working independently, and discrepancies in the extracted data were resolved through discussion to reach a consensus.

Table 1. AMSTAR-2 overall confidence ratings (from Shea et al. [22]).

Rating	Interpretation
High	≤1 non-critical weakness
Moderate	>1 non-critical weakness
Low	1 critical flaw +/- non-critical weaknesses
Critically low	>1 critical flaw +/- non-critical weaknesses

https://doi.org/10.1371/journal.pone.0268070.t001

Quality appraisal

The quality of the selected systematic reviews was independently assessed by two reviewers using the Assessment of Multiple Systematic Reviews (AMSTAR-2) tool [22], an instrument with 16 questions on the methodological quality of systematic reviews. These questions include whether a comprehensive literature search was conducted, justification for excluding studies, and risk of bias assessments for included studies (S2 Appendix). Overall ratings and their meanings are shown in Table 1. AMSTAR-2 was chosen over the risk of bias in systematic reviews (ROBIS) tool [23] because while both assess strongly related aspects (methodological quality vs risk of bias), the former is advantageous for inter-rater reliability and usability [24–26]. Disagreements between the reviewers were similarly resolved through discussion.

It has been suggested that GRADE criteria can be applied to systematic reviews [27]. However, this approach was initially designed for empirical studies, hence the paucity of relevant guidance on how best to achieve this [18, 28]. Therefore, we did not perform any secondary GRADE assessments on the included primary reviews but extracted any reported quality ratings.

Assessing overlap

One of the unique challenges in conducting an umbrella review is overlap–the inclusion of the same primary study or trial in more than one selected systematic review or meta-analysis [29, 30]. Any subsequent synthesis of more than one of these primary reviews would result in 'double-counting' and biased findings because the contribution of a subset of the data would have been multiplied by some factor [30, 31]. The proposed methods for dealing with overlap are: (i) selecting only the most recent systematic review or the one with the largest number of studies, (ii) selecting only the primary review of the highest quality, or (iii) including all primary reviews but evaluating the amount of overlap [18, 30, 31].

To assess overlap between the included primary reviews, we calculated the corrected covered area (CCA) using the formula described by Pieper et al. [29]:

$$CCA\ (Corrected\ CA) = \frac{N - r}{rc - r}$$

where N–number of included primary studies in selected reviews

r–number of index publications

c–number of included primary reviews

A matrix of the included systematic reviews and their primary studies was created to identify the numerators and denominators shown above. CCA is interpreted in banded thresholds: 5% or less indicates slight overlap, 6% to 10% shows moderate overlap, 11% to 15% for high overlap and values greater than 15% indicate very high overlap [29].

Synthesis of results

A narrative synthesis of results was performed due to high levels of study overlap and considerable heterogeneity in primary review inclusion criteria and reported outcomes. Summaries of the included primary reviews are presented below in tabular form alongside corresponding effect estimates such as odds and hazard ratios (where reported). Of note, no further re-analysis of empirical study data was performed, as previously stated in the umbrella review protocol [20]. The extracted data were grouped according to the clinical outcome of interest, regardless of the degree of overlap among sets of studies. Moreover, because there is currently no agreed-upon solution for the issue of low-quality systematic reviews in overviews [18, 26], all selected papers were included in the narrative synthesis.

Results

In total, 763 records were retrieved from the initial and updated searches. After screening by title, abstract and full-text, 20 syntheses were selected for inclusion in this umbrella review. Notably, two otherwise eligible primary reviews were excluded based on all their included studies having been used in later publications by the same authors [32, 33]. Authors were unable to provide further information for two other publications [34, 35]. One item of grey literature, a systematic evidence review commissioned by a government agency for healthcare research, was included [36]. A list of the systematic reviews that were excluded after reading full texts is provided in S3 Appendix. The PRISMA flowchart [37] showing the stages of study selection is presented in Fig 1.

The characteristics of the selected primary reviews are shown in Table 2. Of the 20 included syntheses, four were systematic reviews [36, 38–40], five were published as combined systematic reviews and meta-analyses [17, 41–44], six were labelled meta-analyses [45–50], and five were individual participant data analyses [51–55]. The majority were published earlier than the TRUST trial [15], with only seven primary reviews published during or after 2017 [17, 38, 41–43, 50, 51]. Most of the primary reviews synthesised observational data, although three papers only included RCTs [17, 36, 40]. Generally, SCH was defined using similar TSH thresholds (> 4.5 mIU/L) and normal T4, but some studies subdivided this further into degrees of SCH or TSH elevation e.g., mild vs moderate (Table 2).

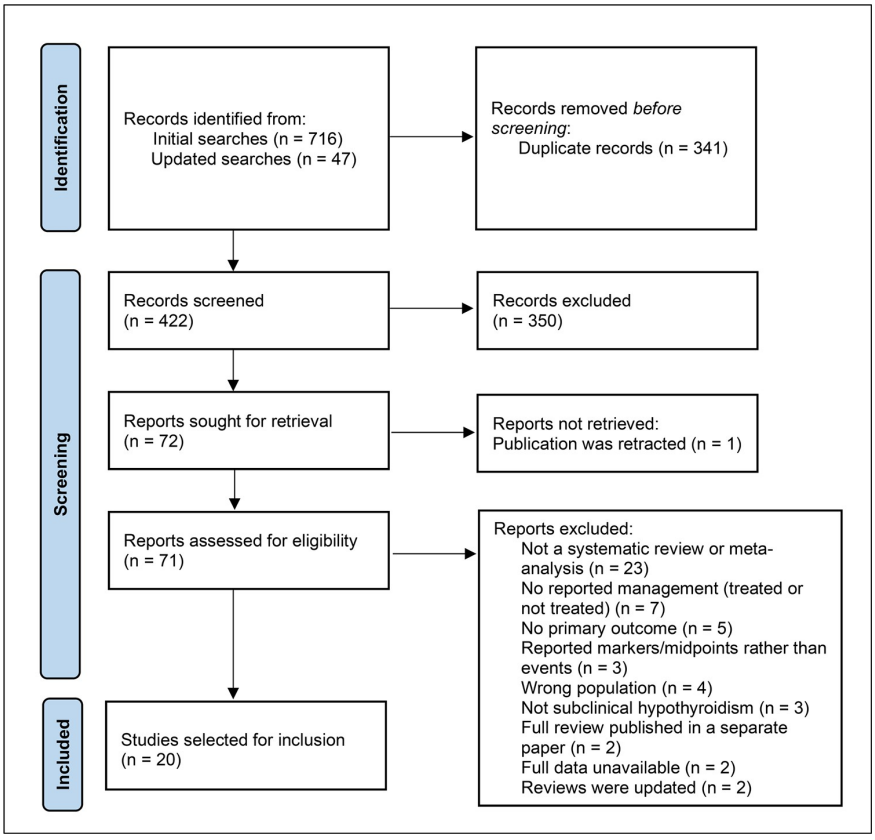


Fig 1. PRISMA flowchart showing study selection (adapted from Page et al. [37]).

https://doi.org/10.1371/journal.pone.0268070.g001

Table 2. Characteristics of the included systematic reviews and meta-analyses.

Study	Design	Study aim	Included studies	Definition of SCH	SCH patients (%)	Summary of findings	AMSTAR-2 Overall confidence
Baumgartner et al. (2017) [51]	IPD	To examine the risk of AF in individuals with thyroid function within the normal range and SCH	11 cohort studies (IPD)	TSH level between 4.5 and 19.9 mIU/L with fT4 levels in the reference range	1958 (6.5)	The reviewers found no link between SCH and the risk of AF; this was the same for individuals with TSH levels within the normal range.	High
Blum et al. (2015) [45]	MA	To assess the association of subclinical thyroid dysfunction with fractures	13 cohort studies	TSH level of 4.50 to 19.99 mIU/L with normal FT4 levels	4092 (5.8)	There was no observed association between SCH and fracture risk.	Moderate
Chaker et al. (2015) [52]	IPD	To evaluate the association between SCH and stroke	17 cohort studies	TSH levels of 4.5 to 19.9 mIU/L with normal T4 levels	3451 (7.3)	There was no overall increase in the risk of stroke events and fatal stroke in patients with SCH than euthyroid patients, except for patients younger than 65 years.	Moderate
Collet et al. (2014) [53]	IPD	To compare the risks of CHD mortality and events associated with SCH by thyroid antibody status	6 cohort studies	TSH 4.5 to 19.9 mIU/L and normal T4 level	1691 (4.4)	Thyroid antibodies were found to have no effect on CHD events and mortality though SCH patients with higher TSH levels were generally at higher risk of developing these outcomes.	Moderate
Dhital et al. (2017) [41]	SR + MA	To look at the association between thyroid function profile and outcomes after acute ischemic stroke	12 cohort studies	Elevated TSH and normal fT4 (study-specific cut-offs)	Unclear	SCH was associated with better functional outcomes after acute ischemic stroke, but this depended on the initial levels of free T3.	Low
Feller et al. (2018) [17]	SR + MA	To examine the association of THT with quality of life and thyroid-related symptoms in adults with SCH	21 RCTS	Thyrotropin and free thyroxine levels above and within centre-specific reference ranges, respectively	2192 (100)	There was no association between treatment of SCH and improving thyroid-related symptoms and quality of life (primary outcomes) or cognitive function, depressive symptoms and the other secondary outcomes.	High
Gencer et al. (2012) [54]	IPD	To clarify the association between subclinical thyroid dysfunction and HF events	6 cohort studies	TSH level of 4.5 to 19.9 mIU/L with normal FT4 levels	2068 (8.1)	Patients with TSH levels higher than 10mIU/L faced a significantly higher risk of HF events.	Moderate
Helfand (2004) [36]	SR	To evaluate the benefits of screening for subclinical thyroid dysfunction	8 RCTs	Elevated TSH and normal T4	Unclear	Evidence of an association between treatment and reduced symptoms was demonstrated only for SCH patients with TSH >10 mIU/L and those with a history of Graves' disease.	Low
Peng et al. (2021) [42]	SR + MA	To investigate whether THT is associated with decreased mortality in adults with SCH	2 RCTs and 5 cohort studies	Grade 1 (TSH level 5.0–10 mIU/L); Grade 2 (TSH level >10 mIU/L) with free thyroxine level within the reference range	21055*3 (100)	Treatment was found to benefit SCH patients younger than 65 years; all-cause mortality decreased by 50%, and cardiovascular mortality decreased by 46%. However, the same did not apply to patients older than 65 years. There was also no overall benefit of treatment on mortality.	High

(*Continued*)

Table 2. (Continued)

Study	Design	Study aim	Included studies	Definition of SCH	SCH patients (%)	Summary of findings	AMSTAR-2 Overall confidence
Razvi et al. (2008) [46]	MA	To examine the influence of age and gender on IHD and mortality in SCH	15 cohort studies	Mild SCH—TSH levels < 10 mIU/L	2,531 (8.7)	The overall incidence of IHD and mortality was not significantly higher for patients with SCH, but IHD prevalence was found to be significantly elevated for patients younger than 65 years.	High
Reyes Domingo et al. (2019) [38]	SR	To synthesize the evidence on the effects of screening and subsequent treatment for thyroid dysfunction	5 RCTs and 3 cohort studies*2	Study-specific	Unclear	Evidence was found linking treatment for SCH with reduced all-cause mortality for patients younger than 65 years, but it was determined to be of low quality.	High
Rodondi et al. (2006) [47]	MA	To determine whether SCH is associated with an increased risk for CHD	5 cohort, 6 cross-sectional and 3 case-control studies	Elevated TSH and a normal T4 (no pre-specified cut-offs)	1409 (10.8)	Compared to euthyroid patients, CHD was 1.6 times more likely in patients with SCH; this association was constant throughout the included studies but less pronounced in the prospective cohorts.	High
Rodondi et al. (2010) [55]	IPD	To assess the risks of CHD and total mortality for adults with SCH	11 cohort studies	Serum TSH level of 4.5 mIU/L or greater to less than 20 mIU/L, with a normal T4 concentration	3450 (6.2)	SCH patients with TSH levels higher than 10mIU/L had a significantly higher risk of CHD events and mortality than euthyroid patients.	High
Rugge et al. (2015) [39]	SR	To assess the benefits and harms of screening and treatment of subclinical and undiagnosed overt hypothyroidism and hyperthyroidism in adults*	13 RCTs and 1 cohort study	4.5–10.0 mIU/L (mildly elevated) or ≥10 mIU/L (markedly elevated) TSH levels with normal thyroxine	Unclear	Reviewers found a potential association between SCH and cardiovascular disease but inconclusive evidence that treatment would be beneficial; SCH treatment was also not associated with improved cognitive function or quality of life.	Moderate
Singh et al. (2008) [48]	MA	To compare the relative risk for incident CHD events, cardiovascular-related and total mortality associated with subclinical thyroid abnormalities	6 cohort studies	Serum TSH above 4.0–5.0 mIU/L with normal free T4 (range 0.7–1.8 ng/dL)	1365 (10.2)	SCH was linked to a significant risk of CHD at baseline and both CHD and cardiovascular mortality during follow-up. On the other hand, all-cause mortality was not found to be increased with SCH.	Low
Sun et al. (2017) [43]	SR + MA	To explore the relationship between subclinical thyroid dysfunction and the risk of cardiovascular outcomes	16 cohort studies	TSH levels >3.6 to 6 mIU/L (study-specific)	5178 (7.2)	There was a significantly higher risk of CHD and cardiovascular mortality for SCH patients younger than 65 years, but the same effect was not observed for patients older than 80 years. A slightly higher risk of AF and HF was also associated with SCH.	Moderate

(*Continued*)

Table 2. (Continued)

Study	Design	Study aim	Included studies	Definition of SCH	SCH patients (%)	Summary of findings	AMSTAR-2 Overall confidence
Villar et al. (2007) [40]	SR	To assess the effects of thyroid hormone replacement for SCH	12 RCTs	TSH level above the upper limit of the reference range with normal values of total T4 or free T4 (FT4), with or without T3 or free T3 (FT3) measurements	350 (100)	It was not possible to assess the benefits of SCH treatment on reducing cardiovascular mortality. However, there was also no significant impact of levothyroxine on health-related quality of life and symptoms.	High
Wirth et al. (2014) [44]	SR + MA	To assess the risk for hip and non-spine fractures associated with subclinical thyroid dysfunction	7 cohort studies	TSH level greater than 4.5 to 20.0 mIU/L and an FT4 level in the reference range	Unclear	No association between SCH and fracture risk was found, but the reviewers could not assess the effects of treatment vs no treatment due to insufficient data.	Moderate
Yan et al. (2016) [49]	MA	To identify the relationship between subclinical thyroid dysfunction and the risk of fracture	5 cohort studies	TSH level greater than 4.0 to 5.5 mIU/L (study-specific)	2580 (0.9)	A link between SCH and higher fracture risk was not found, but the reviewers acknowledge that they had limited data.	Low
Yang et al. (2019) [50]	MA	To assess the association between subclinical thyroid dysfunction and the clinical outcomes of HF patients	14 cohort studies	Elevated TSH values in the presence of normal FT4 values	2308 (10.9)	Both adjusted and unadjusted analyses showed a significantly higher risk of all-cause mortality and cardiovascular death associated with SCH for patients with heart failure.	Low

THT—Thyroid Hormone Therapy; SR–Systematic Review; SR + MA–Systematic Review and Meta-Analysis; MA–Meta-analysis; IPD–Individual Participant Data analysis; SCH–Subclinical Hypothyroidism; CHD–Coronary Heart Disease; AF–Atrial Fibrillation; RCT–Randomised Controlled Trial; HF–Heart Failure; IHD–Ischaemic Heart Disease; Thyroxine–T4, fT4, thyroid hormone; Thyrotropin–Thyroid Stimulating Hormone (TSH)

*this was an update to Helfand et al. [36], but because the searches did not overlap, this was considered a separate review.

*[2]only for the relevant research question on clinical outcomes for SCH.

*[3]the authors report potential overlap between the studies; hence the estimate may be incorrect.

https://doi.org/10.1371/journal.pone.0268070.t002

Coverage of the primary outcomes was good, given that all the outcomes of interest were reported in at least two publications. However, it is crucial to note that most relevant results were obtained via subgroup or sensitivity analyses in the primary reviews. As such, they were not necessarily representative of the overall findings shown in Table 2.

All-cause mortality

Seven publications reported findings on all-cause mortality; of these, three primary reviews compared rates between treated and untreated patients [38, 39, 42], three compared untreated and euthyroid individuals [43, 46, 48], and one compared both treated and untreated SCH groups with euthyroid participants [50]. There was no statistically significant difference in the overall numbers of deaths from all causes for patients with SCH between those who were and were not on treatment (Table 3).

Taking age into account, lower estimates of all-cause mortality were reported for patients younger than 70 years on treatment (RR 0.50, 95% CI 0.29 to 0.85 [42]; HR 0.36, 95% CI 0.19 to 0.66 [38, 39]). However, these estimates were based on one study in Rugge et al. [39], and another paper rated the same evidence as being of very low certainty [38]. On the other hand, older patient groups demonstrated no significant association between levothyroxine treatment and all-cause mortality (Table 3).

Table 3. Review findings on all-cause mortality.

Study	Outcome	Treatment status	Comparator	Effect estimate (95% CI)
Peng et al. (2021) [42]	All-cause mortality	Treated	Untreated	RR 0.95 (0.75–1.22)
	All-cause mortality; age <65–70 years			RR 0.50 (0.29–0.85)
	All-cause mortality; age > = 65–70 years			RR 1.08 (0.91–1.28)
Reyes Domingo et al. (2019) [38]	All-cause mortality; adults (>18 years)	Treated	Untreated	HR 1.91 (0.65–5.60)
	All-cause mortality; adults (<65 or <70 years)			IRR 0.63 (0.40–0.99
				HR 0.36 (0.19–0.66)
	All-cause mortality; adults (>65 years)			HR 1.91 (0.65–5.60)
	All-cause mortality; females			IRR 0.99 (0.85–1.16)
				1.08 (0.80–1.48)
	All-cause mortality; males			IRR 1.24 (0.89–1.16)
				1.43 (0.87–2.34)
Rugge et al. (2015) [39]	All-cause mortality; 40–70 years	Treated	Untreated	HR 0.36 (0.19–0.66)
	All-cause mortality; >70 years			HR 0.71 (0.56–1.08)
Yang et al. (2019) [50]	All-cause mortality	Untreated	Euthyroid	HR 1.48 (1.29–1.70)
		Treated	Euthyroid	HR 1.48 (1.14–1.94)
Razvi et al. (2008) [46]	IHD/all-cause mortality; <65 years	Untreated	Euthyroid	OR 1.32 (0.95–1.83)
	IHD/all-cause mortality; > 65 years			OR 0.87 (0.51–1.45)
Sun et al. (2017) [43]	Total mortality	Untreated	Euthyroid	RR 1.01 (0.90–1.15)
Singh et al. (2008) [48]	All-cause mortality	Untreated	Euthyroid	RR 1.115 (0.990–1.255)

HR–Hazard Ratio; RR–Relative Risk; IRR–Incident rate Ratio; IHD–Ischemic Heart Disease; OR–Odds Ratio.

https://doi.org/10.1371/journal.pone.0268070.t003

Only one of four comparisons of all-cause mortality between untreated and euthyroid study participants was statistically significant (HR 1.48, 95% CI 1.29 to 1.70) [50]. The same review found that death was more likely among SCH patients on treatment than in euthyroid persons (HR 1.48, 95% CI 1.14 to 1.94) [50]. However, the population of interest for this review all had heart failure, thereby limiting the generalisability of these findings to other SCH patients.

Cardiovascular outcomes

Cardiovascular outcomes were the most extensively reported outcomes of interest across the included reviews (n = 13) as shown in Table 4. No difference was found in the number of incident atrial fibrillation events between untreated persons and euthyroid controls, irrespective of age and TSH level [51]. Similarly, the difference between treated and untreated SCH patients was not statistically significant, though notably, the evidence was rated as being of very low to moderate certainty [38].

Compared to euthyroidism, untreated SCH was significantly associated with a higher likelihood of CHD and heart failure if patients had TSH levels above 10mIU/L (HR 2.17, 95% CI 1.19 to 3.93) [55], (HR 2.37, 95% CI 1.59 to 3.54) [54] or were thyroid peroxidase antibody-negative (HR 1.25, 95% CI 1.06 to 1.47); HR (3.76, 95% CI 1.77 to 8.01) [53]. It was also reported that untreated SCH was associated with higher odds of ischaemic heart disease (OR 1.58, 95% CI 1.07 to 2.35) [46] and a higher risk of developing coronary heart disease during follow-up (RR 1.188, 95% CI 1.024 to 1.379) [48] than euthyroid participants. However, one primary review found that incident CHD was not associated with untreated SCH [43]–this difference may have resulted from the reviewers' decision to restrict the inclusion of primary studies based on quality appraisal scores.

Table 4. Reported primary review findings on cardiovascular outcomes.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Baumgartner et al. (2017) [51]	Atrial fibrillation	Untreated	Euthyroid (TSH 3.50–4.49 mIU/L)	For TSH 4.5–6.9 mIU/L: HR 0.87 (0.66–1.16)
				For TSH 7.0–9.9 mIU/L: HR 1.22 (0.78–1.92)
				For TSH 10.0–19.9 mIU/L: HR 1.56 (0.84–2.90)
Reyes Domingo et al. (2019) [38]	Atrial fibrillation; adults (>18y)	Treated	Untreated	HR 0.80 (0.35–1.80)
	Atrial fibrillation; adults (<65 or <70)			HR 0.76 (0.26–1.73)
	Atrial fibrillation; adults (>65y)			HR 0.80 (0.35–1.80)
Collet et al. (2014) [53]	CHD events	Untreated	Euthyroid	SH With -ve TPOAb HR 1.25 (1.06–1.47)
				SH With +ve TPOAb HR 1.12 (0.88–1.41)
				SH with TSH ≥10.0 mIU/L and neg. TPOAb HR 3.76 (1.77–8.01)
				SH with TSH ≥10.0 mIU/L and pos. TPOAb HR 1.19 (0.61–2.32)
Rodondi et al. (2006) [47]	CHD	Untreated	Euthyroid	OR 2.06 (1.36–3.14)
Rodondi et al. (2010) [55]	CHD	Untreated	Euthyroid	For TSH 4.5–19.99 mIU/L: HR 1.17 (0.91–1.50)
				For TSH 10–19.99 mIU/L: HR 2.17 (1.19–3.93)
Sun et al. (2017) [43]	CHD	Untreated	Euthyroid	RR 1.02 (0.92–1.14)
Singh et al. (2008) [48]	CHD (during follow-up)	Untreated	Euthyroid	RR 1.188 (1.024–1.379)
Gencer et al. (2012) [54]	Heart failure events; TSH 4.5–19.9 mIU/L	Untreated	Euthyroid	HR 1.26 (0.93–1.69)
	Heart failure events; TSH 10.0–19.9 mIU/L			HR 2.37 (1.59–3.54)
Razvi et al. (2008) [46]	IHD incidence; < 65 yrs	Untreated	Euthyroid	OR 1.58 (1.07–2.35)
	IHD incidence; > 65 yrs			N/P
Rugge et al. (2015) [39]	IHD; 40–70 yrs	Treated	Untreated	HR 0.61 (0.39–0.95)
	IHD; >70 yrs			HR 0.99 (0.59–1.33)
Reyes Domingo et al. (2019) [38]	Fatal and non-fatal cardiovascular events (not AF); adults (>18y)	Treated	Untreated	HR 0.89 (0.47–1.69)
	Fatal and non-fatal cardiovascular events (not AF); adults (<65 or <70)			HR 0.61 (0.39–0.95)
				HR 1.03 (0.51–2.13)
				IRR 1.11 (0.61–2.02)
	Fatal and non-fatal cardiovascular events (not AF); adults (>65y)			HR 0.89 (0.47–1.69)
	Fatal and non-fatal cardiovascular events (not AF); females			IRR 0.99 (0.70–1.38)
				0.99 (0.70–1.40)
	Fatal and non-fatal cardiovascular events (not AF); males			IRR 1.41 (0.83–2.40)
				1.36 (0.79–2.35)

CHD–Coronary heart disease; HR–Hazard Ratio; RR–Relative Risk; IRR–Incident rate Ratio; IHD–Ischemic Heart Disease; OR–Odds Ratio.

https://doi.org/10.1371/journal.pone.0268070.t004

SCH patients receiving treatment and younger than 70 years were significantly less likely to develop IHD than untreated individuals (HR 0.61, 95% CI 0.39 to 0.95) [38]. However, these findings were based on a single empirical study in both primary reviews, in which it was reported that the GRADE rating for this evidence was very low. A similar association was not found for patients older than 70 years nor subgroups based on sex [38].

Table 5. Reported primary review findings on cardiovascular mortality.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Peng et al. (2021) [42]	Cardiovascular mortality	Treated	Untreated	RR 0.99 (0.82–1.20)
	Cardiovascular mortality; age <65–70 years			RR 0.54 (0.37–0.80)
	Cardiovascular mortality; age > = 65–70 years			RR 1.05 (0.87–1.27)
Reyes Domingo et al. (2019) [38]	Cardiovascular deaths; adults (>18y)	Treated	Untreated	OR 2.01 (0.18–22.27)
	Cardiovascular deaths; adults (<65 or <70)			HR 0.54 (0.37–0.92) IRR 0.55 (0.25–1.20)
	Cardiovascular deaths; adults (>65y)			OR 2.01 (0.18–22.27)
	Cardiovascular deaths; females			IRR 0.96 (0.77–1.21)
	Cardiovascular deaths; males			IRR 1.32 (0.83–2.08)
Rugge et al. (2015) [39]	Cardiovascular deaths (40–70 years)	Treated	Untreated	HR 0.54 (0.37–0.92)
Collet et al. (2014) [53]	CHD mortality	Untreated	Euthyroid	SH With -ve TPOAb HR 1.34 (1.07–1.69)
				SH With +ve TPOAb HR 1.28 (0.94–1.72)
				SH with TSH ≥10.0 mIU/L and negative TPOAb HR 1.95 (0.76–4.98)
				SH with ≥10.0 mIU/L and positive TPOAb HR 1.92 (1.09–3.36)
Rodondi et al. (2010) [55]	CHD mortality	Untreated	Euthyroid	For TSH 4.5–19.99 mIU/L: HR 1.25 (1.04–1.51)
				For TSH 10–19.99 mIU/L: HR 1.85 (1.13–3.05)
Sun et al. (2017) [43]	Cardiovascular mortality	Untreated	Euthyroid	RR 0.86 (0.56–1.32)
Singh et al. (2008) [48]	Cardiovascular mortality	Untreated	Euthyroid	RR 1.278 (1.023–1.597)
Yang et al. (2019) [50]	Cardiac death and/or hospitalization	Untreated	Euthyroid	HR 1.32 (1.08–1.60)
		Treated		HR 1.36 (1.12–1.66)

CHD–Coronary Heart Disease; HR–Hazard Ratio; RR–Relative Risk; IRR–Incident rate Ratio; OR–Odds Ratio.

https://doi.org/10.1371/journal.pone.0268070.t005

We distinguished between all-cause mortality and cardiovascular mortality (n = 8), which primary review authors defined as deaths arising from cardiovascular diseases (Table 5). Considering treated vs untreated SCH, an association between treatment and cardiovascular death was found only for adult patients younger than 65–70 years [38, 39, 42]. Stratifying the results by sex did not yield statistically significant findings. In addition, whereas the risk of cardiovascular mortality was found to be higher for untreated SCH patients compared to euthyroid controls in four primary reviews [48, 50, 53, 55], Sun et al. [43] reported a lower nonsignificant estimate (RR 0.86, 95% CI 0.56 to 1.32) [43]. A possible reason for this difference, despite very high overlap between pairs of these primary reviews, could be discrepancies in the determination of treatment status. Furthermore, Sun et al. [43] rated the quality of evidence for cardiovascular mortality in their primary review as low because of high heterogeneity.

Only one primary review considered the relationship between thyroid peroxidase antibody status and cardiovascular mortality. Collet et al. [53] found that untreated thyroid antibody-negative SCH was significantly associated with cardiovascular mortality (HR 1.34 95% CI 1.07 to 1.69), but the same did not apply for antibody-positive SCH (HR 1.28, 95% CI 0.94 to 1.72) [53], except for patients that also had TSH levels above 10mIU/L (HR 1.92, 95% CI 1.09 to 3.36) [53]. Finally, compared to euthyroid controls, death and hospitalisation due to cardiovascular causes were more likely to occur among treated SCH patients with heart failure [50].

Table 6. Primary review findings on stroke.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Chaker et al. (2015) [52]	Stroke events	Untreated	Euthyroid	HR 0.96 (0.70–1.31)
	Fatal stroke			HR 1.27 (0.74–2.16)
Dhital et al. (2017) [41]	Stroke–modified Rankin scale	Untreated	Euthyroid	OR after 1 month 2.58 [1.13–5.91]
				OR after 3 months 2.28 [1.33–3.91]
	Stroke–mortality after 3 months			OR 0.20, (0.04–1.12)

HR–Hazard Ratio; OR–Odds Ratio.

https://doi.org/10.1371/journal.pone.0268070.t006

Stroke

There were no direct comparisons of the risk of stroke between treated and untreated patients with SCH. One primary review compared untreated individuals with SCH with euthyroid controls, and no significant difference was found in either the incidence of strokes or deaths arising from strokes [52]. Dhital et al. [41] found that functional outcomes for untreated SCH (based on the modified Rankin scale) were twice as likely to be better than those for euthyroid controls 1 and 3 months after acute ischemic stroke (Table 6).

Fractures

No primary reviews that compared the risk or likelihood of fractures between euthyroid individuals and SCH patients that did or did not receive treatment found a significant difference (Table 7). This result was similar across various types of fractures, for example, hip fractures (HR 1.02, 95% CI 0.87 to 1.19 [45]; HR 1.10, 95% CI 0.81 to 1.50 [44]) and spine fracture (HR 1.16, 95% CI 0.66 to 2.04) [45].

Quality of life and presence of symptoms

For the 5 studies that explicitly reported quality of life outcomes, no statistically significant differences were found between patients who did and did not receive treatment for SCH (Table 8). Similarly, thyroid-related symptoms, fatigue, mental and general well-being scores were not significantly associated with treatment status. However, Helfand [36] reported that specific subgroups–patients with TSH values greater than 10 mIU/L and those with a history of Graves' disease seemed to benefit from treatment. Graves' disease is an autoimmune thyroid disorder treated with antithyroid medication, radiotherapy or surgery [56]. Nonetheless, it is

Table 7. Primary review findings on fractures.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Blum et al. (2015) [45]	Hip fracture	Untreated	Euthyroid	HR 1.02 (0.87–1.19)
	Any fracture			HR 1.11 (0.94–1.30)
	Non-spine fracture			HR 1.13 (0.93–1.38)
	Spine fracture			HR 1.16 (0.66–2.04)
Reyes Domingo et al. (2019) [38]	Fractures; adults (all >65)	Treated	Untreated	HR 1.06 (0.41–2.76)
Yan et al. (2016) [49]	Fractures (any)	Untreated	Euthyroid	RR 1.25 (0.85–1.84)
		Treated		RR 1.22 (0.61–2.47)
Wirth et al. (2014) [44]	Hip fractures	Untreated	Euthyroid	HR 1.10 (0.81–1.50)
	Non-spine fractures			HR 1.11 (0.60–2.05)

HR–Hazard Ratio; RR–Relative Risk.

https://doi.org/10.1371/journal.pone.0268070.t007

Table 8. Primary review findings on quality of life and symptoms.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Feller et al. (2018) [17]	General QoL	Treated	Untreated	SMD -0.11 (-0.25–0.03)
Reyes Domingo et al. (2019) [38]	Thyroid QoL—less than 12 mo	Treated	Untreated	MD 0.0 (-2.0–2.1)
	Thyroid QoL—more than 12 mo			MD 1.0 (-1.9–3.9)
				-0.5 (-2.2–1.3)
Rugge et al. (2015) [39]	Quality of life	Treated	Untreated	Multiple
Rugge et al. (2015) [39]	Thyroid-related symptoms	Treated	Untreated	SMD 0.01 (-0.12–0.14)
	Fatigue and tiredness			SMD -0.01 (-0.16–0.15)
	Depressive symptoms			SMD -0.10 (-0.34–0.13)
Helfand (2004) [36]	Symptoms	Treated	Untreated	Multiple
Villar et al. (2007) [40]	Symptoms, mood and quality of life	Treated	Untreated	Multiple
Reyes Domingo et al. (2019) [38]	Fatigue/tiredness—less than 12 mo	Treated	Untreated	MD 0.4 (-2.1–2.9)
	Fatigue/tiredness—more than 12 mo			MD -3.5 (-7.0–0.0)
	Mental well-being			Multiple
	Physical well-being			MD -0.1 (-0.3–1.0)
				-0.1 (-0.3–1.0)
	General well-being			Multiple

SMD–Standardised Mean difference; MD–Mean Difference.

https://doi.org/10.1371/journal.pone.0268070.t008

noted that the single study that this finding was based upon was a small trial of 33 participants, all of whom had previously treated Graves' disease [36].

Secondary outcomes

Some of the included papers (n = 4) reported on cognitive function (Table 9), which was assessed using various tools such as the Letter-Digit Coding Test and Mini-Mental State Examination. All of the primary reviews found no significant difference in cognitive function between treated and untreated groups [17, 38, 39] except Villar et al. [40]. However, this result was based on only one included study with an unclear risk of bias assessment.

Overlap

The extent of overlap in this umbrella review is shown in Fig 2, an intersection heatmap of the calculated CCA between pairs of the 20 included primary reviews. As shown, only one of the included evidence syntheses [41] had a unique set of primary publications compared to all the other primary reviews. Overall, excluding the diagonal, the pairwise comparisons showed slight (66.7%), moderate (10%), high (2.6%) and very high (20.8%) overlap. However, it should be noted that these values were obtained with no consideration of the specific outcomes

Table 9. Primary review findings on cognitive function.

Study	Outcome	Treatment status	Comparator	Effect estimates (95% CI)
Feller et al. (2018) [17]	Cognitive function	Treated	Untreated	Difference 1.01 (95% CI −0.56 to 2.46)
Reyes Domingo et al. (2019) [38]	Cognitive function	Treated	Untreated	Multiple (no difference)
Villar et al. (2007) [40]	Cognitive function	Treated	Untreated	MD 2.4 (0.3–4.5)
Rugge et al. (2015) [39]	Cognitive function	Treated	Untreated	Multiple (no difference)

MD–Mean Difference.

https://doi.org/10.1371/journal.pone.0268070.t009

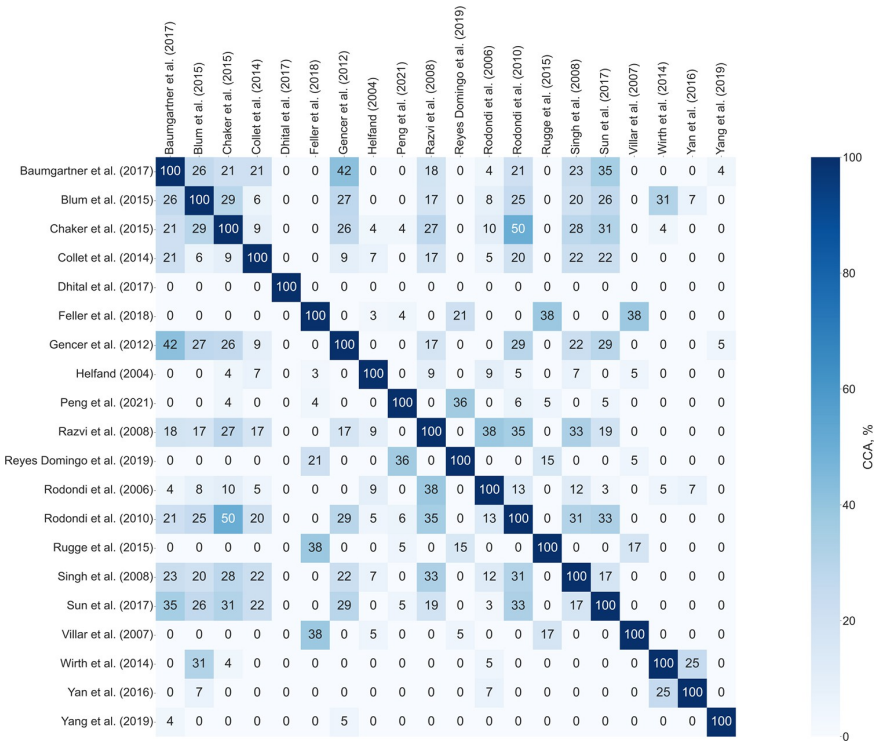


Fig 2. Heatmap showing pairwise calculated CCA.

https://doi.org/10.1371/journal.pone.0268070.g002

reported in each of the primary reviews and therefore require cautious interpretation. That is because, although calculating CCA involved mapping all the primary studies included in each publication, not all provided findings relevant to this umbrella review. The fundamental reason for this complexity is that we were only interested in estimates reported with the participants' corresponding treatment status. On balance, CCA for the entire overview was calculated to be 5.12%, the higher limit for slight overlap [29].

Quality appraisal

Overall confidence in review findings was found to be high for eight primary reviews [17, 38, 40, 42, 46, 47, 51, 55], moderate for seven primary reviews [39, 43–45, 52–54] and low for five primary reviews [36, 41, 48–50]. Three syntheses did not include a meta-analysis and therefore could not be assessed for questions 11, 12 and 15 [36, 38, 39]. The breakdown of checklist questions is shown in 0 10.

Discussion

This umbrella review on the impact of the management of SCH on clinical outcomes covers evidence from 20 selected systematic reviews and meta-analyses of RCTs and observational studies. Across the outcomes of interest, the synthesised literature found can be summarised as follows. We found that the treatment of SCH may be associated with a reduced likelihood of death from all causes for patients under 70 years old. On the other hand, the relationship between SCH treatment status and the risk of death compared to the euthyroid population

Table 10. Results of the AMSTAR-2 assessments.

Review	Question																Overall confidence in results
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Baumgartner et al. (2017) [51]	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	High
Blum et al. (2015) [45]	Y	N	N	Y	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate
Chaker et al. (2015) [52]	Y	N	Y	PY	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Moderate
Collet et al. (2014) [53]	Y	N	N	PY	Y	N	N	PY	N	N	Y	N	N	N	Y	Y	Moderate
Dhital et al. (2017) [41]	Y	N	N	PY	Y	Y	N	Y	PY	N	Y	Y	Y	N	N	Y	Low
Feller et al. (2018) [17]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	High
Gencer et al. (2012) [54]	Y	PY	Y	PY	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate
Helfand (2004) [36]	Y	PY	Y	Y	N	N	N	Y	PY	N	NMA	NMA	N	N	NMA	N	Low
Peng et al. (2021) [42]	Y	Y	N	PY	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Razvi et al. (2008) [46]	Y	N	Y	PY	Y	Y	PY	Y	Y	N	Y	Y	Y	Y	Y	Y	High
Reyes Domingo et al. (2019) [38]	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	NMA	NMA	Y	Y	NMA	Y	High
Rodondi et al. (2006) [47]	Y	N	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	High
Rodondi et al. (2010) [55]	Y	PY	Y	PY	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Rugge et al. (2015) [39]	N	PY	Y	Y	Y	Y	N	N	PY	N	NMA	NMA	Y	N	NMA	Y	Moderate
Singh et al. (2008) [48]	Y	N	Y	PY	N	N	N	Y	N	N	Y	N	N	N	N	N	Low
Sun et al. (2017) [43]	Y	N	Y	PY	Y	Y	N	PY	Y	N	Y	Y	Y	Y	N	Y	Moderate
Villar et al. (2007) [40]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	High
Wirth et al. (2014) [44]	Y	PY	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Moderate
Yan et al. (2016) [49]	Y	N	N	PY	Y	Y	N	Y	Y	N	Y	Y	N	Y	N	Y	Low
Yang et al. (2019) [50]	Y	N	N	PY	Y	Y	N	N	N	N	Y	N	N	Y	Y	Y	Low

Y- Yes; N–No; PY–Partial Yes; NMA–No Meta-Analysis.

https://doi.org/10.1371/journal.pone.0268070.t010

remains unclear. Increased risk of all-cause mortality for untreated SCH was reported in only one primary review [50], for which inclusion was restricted to patients with comorbid heart failure.

We also found that compared to euthyroidism, untreated SCH patients with very high TSH (>10mIU/L) may be at greater risk of cardiovascular events and death from cardiovascular disease [54, 55]. The same effects were observed for thyroid peroxidase antibody-negative patients [53]. Even so, there was discordance in findings between the primary reviews; whereas seven primary reviews reported a higher risk of CHD and cardiovascular mortality for untreated SCH patients than euthyroid persons [46–48, 50, 53–55], one primary review did not [43]. We rated the latter as having more than one non-critical weakness according to the AMSTAR-2 checklist; the others were either 'high' or 'moderate' in overall confidence in their results (Table 10). A high degree of overlap was calculated between the studies reporting cardiovascular outcomes, as high as 38%. Therefore, it was not easy to ascertain the precise source of the difference in results.

It was not possible to investigate the impact of treatment on the risk of stroke because the only available comparisons were of untreated SCH and euthyroidism. The finding that untreated patients had better functional outcomes one month following stroke was reported only in one low-quality study [41] and is, therefore, inconclusive. In a similar vein, there was insufficient evidence of the impact of treatment or no treatment of SCH on fracture risk. Overall confidence in the results of three out of the four primary reviews [38, 44, 45] was rated as 'high' or 'moderate'. However, none of the effect estimates was statistically significant, so it is also not possible to make conclusions on this relationship based on the quantity of evidence.

Reported findings on quality of life and the presence of symptoms between treated and untreated SCH patients were mainly of no statistical significance. As such, we cannot definitively state whether levothyroxine treatment improves or worsens these outcomes. Nonetheless, medication potentially benefits two patient groups–patients with severe SCH from the start and those that previously received treatment for autoimmune hyperthyroidism/Graves' disease [36].

The secondary outcome reported in the included primary reviews, cognitive function, was only compared between treated and untreated SCH patients. Given that the majority of findings were similar, it may be said that among patients with SCH, levothyroxine may have no significant impact on cognitive function, notwithstanding the type of assessment tool used [17, 38–40]. Crucially, however, two points must be emphasised. First, that the amount of evidence in favour of this statement is notably low, considering that few primary reviews that reported on cognitive function. Second, that only the primary outcomes were included in the literature searches, so the findings in this review cannot accurately reflect the body of evidence regarding the relationship between SCH and cognitive function.

With reference to the number of primary reviews, the volume of evidence was discernibly skewed in favour of cardiovascular outcomes (n = 13) rather than all-cause mortality (n = 7), stroke (n = 2), fractures (n = 4), quality of life (n = 5) and cognitive function (n = 4). This observation can be explained as having arisen from the umbrella review selection process, but the relatively broad inclusion criteria make it less probable. Instead, two alternatives are suggested; either that less research has been performed on the other clinical outcomes of interest or that the evidence may not have already been synthesised due to high between-study heterogeneity, for example, in outcome definitions and measurements. Additional factors, such as the comparative ease of measuring certain outcomes over others, may also influence which types of studies are performed. However, it is not possible to conclusively account for this asymmetry of evidence from this overview alone.

Generally, it cannot be ignored that most of our findings were based on empirical studies of poor quality, as reported by the authors of the primary reviews. Equally important were the critical flaws we found in the methodological quality of five of the selected primary reviews [36, 41, 48–50] consequently rated as 'low' in overall confidence in their results. Upon inspection, there was no clear boundary of review quality based on the type of empirical research that was initially selected. For example, all the syntheses that included only RCTs did not consistently get higher AMSTAR-2 ratings than those of only observational studies. As such, it can be argued that cohort studies have an essential role in filling the gap left by insufficient randomised trials on this topic.

It should be noted that there was a tendency for papers with the lowest ratings on the AMSTAR-2 checklist to have little overlap of empirical studies with other higher-rated primary reviews. This could be explained by differences in the types of outcomes reported in these syntheses; for instance, one would expect minimal overlap between fractures and cardiovascular mortality. Collectively, the reviews included in this overview had slight overlap, but as Hennesy and Johnson [31] contend, such an observation can be attributed to the breadth of the literature. This is especially true if only a small set of identical studies is shared across the included syntheses, or the overlap is highly outcome-dependent [31]. In these cases, the overall CCA would obscure the true level of overlap.

Strengths and limitations

To the authors' knowledge, this is the first umbrella review on this topic. This overview was conducted in a systematic manner and comprehensive searches were performed to identify the

synthesised literature on the impact of the management of SCH on long-term clinical outcomes. The database searches–including grey literature, to minimise the effects of publication bias [57]–were updated in the course of the review. Screening, data extraction and quality appraisal were all done in duplicate. Furthermore, the intended aim of the umbrella review to compare the synthesised literature on this topic was achieved, even though a secondary meta-analysis was not feasible.

Nonetheless, it is crucial to consider the limitations of this review which relied exclusively on the availability, methods and quality of existing systematic reviews and meta-analyses. Of note, it was not possible to re-analyse and pool all primary review findings due to the variety of selection criteria and outcome definitions. Combining the findings of the included reviews in spite of these differences–and potential confounders–would result in biased and misleading inferences [58].

Also, an inherent limitation of the umbrella review methodology is the limited capacity to conduct detailed evaluations of empirical studies when dealing with synthesised literature. This was particularly challenging when evaluating overlap across the included reviews, as it may have been influenced by factors such as study scope and eligibility criteria. On the other hand, because this type of review was performed, it was possible to examine a wide variety of outcomes for SCH and treatment status within our specific resource constraints. The comprehensive nature of umbrella reviews has been recommended for controversial topics [19]. Furthermore, in this overview, we included IPD meta-analyses, which have been described as beneficial for analysing long-term patient outcomes [59].

Another limitation was scope mismatch between the umbrella review and the included primary reviews, for example, in cases where a selected systematic review included patients with subclinical hypo- and hyperthyroidism. This problem is commonly encountered in overviews [18], and we opted to include such papers for two key reasons. First, a preliminary literature search yielded few results with precisely the same research questions. Second, for an unresolved topic such as this, it was anticipated that the exclusion of these reviews would severely restrict this synthesis by omitting potentially relevant findings. Therefore, to limit this type of bias, inclusion in the umbrella review was based on the availability of results for which treatment status was explicitly stated.

It was also not possible to calculate overlap for the included primary reviews subdivided by their reported outcomes because most included both treated and untreated SCH patients. Consequently, assessing overlap in this way would require a detailed inspection of all their primary studies to identify the exact data sources for the respective subgroup analyses. These activities were considered to be burdensome and beyond the scope of this umbrella review, given that systematic reviews and meta-analyses were the principal units of analysis. Even so, to visualise overlap, we created a citation matrix and presented the results of the pairwise calculations, most of which were in the 'slight' band.

Conclusion

Through this umbrella review, we systematically gathered the existing synthesised literature on the impact of the management of subclinical hypothyroidism on clinical outcomes. Our findings seem to indicate that treatment may be beneficial for SCH patients younger than 70 years due to the higher risk of all-cause mortality and cardiovascular events. In addition, untreated SCH patients with TSH levels above 10mIU/L may be at higher risk of developing cardiovascular diseases than the euthyroid population. However, more robust evidence is needed on stroke, fractures, quality of life and cognitive function in SCH. The main challenge in investigating long-term outcomes is the need for large, adequately powered and timed randomised

trials. This overview further highlights this need, given that majority of the significant findings were based on very few empirical studies often deemed to be of poor quality by the primary reviewers. Future work in observational studies may also be instrumental in strengthening the evidence base.

Supporting information

S1 Appendix. MEDLINE search strategy.
(PDF)

S2 Appendix. AMSTAR-2 checklist.
(PDF)

S3 Appendix. Excluded papers.
(PDF)

S4 Appendix. PRISMA checklist.
(PDF)

S5 Appendix. PRISMA abstract checklist.
(PDF)

Acknowledgments

We wish to thank Vicki Cormie, a senior librarian at the University of St Andrews, who assisted with the development and execution of the database searches.

Author Contributions

Conceptualization: Brenda S. Bauer, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Colin McCowan.

Data curation: Brenda S. Bauer.

Formal analysis: Brenda S. Bauer.

Funding acquisition: Colin McCowan.

Investigation: Brenda S. Bauer, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan.

Methodology: Brenda S. Bauer, Amaya Azcoaga-Lorenzo.

Project administration: Brenda S. Bauer.

Supervision: Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Colin McCowan.

Visualization: Brenda S. Bauer.

Writing – original draft: Brenda S. Bauer.

Writing – review & editing: Brenda S. Bauer, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan.

References

1. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocrine reviews. 2008; 29(1):76–131. https://doi.org/10.1210/er.2006-0043 PMID: 17991805
2. Cooper DS, Biondi B. Subclinical thyroid disease. The Lancet. 2012; 379(9821):1142–54. https://doi.org/10.1016/S0140-6736(11)60276-6 PMID: 22273398

3. Peeters RP. Subclinical hypothyroidism. New England Journal of Medicine. 2017; 376(26):2556–65. https://doi.org/10.1056/NEJMcp1611144 PMID: 28657873

4. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. Jama. 2004; 291(2):228–38. https://doi.org/10.1001/jama.291.2.228 PMID: 14722150

5. Schübel J, Feldkamp J, Bergmann A, Drossard W, Voigt K. Latent hypothyroidism in adults. Deutsches Ärzteblatt International. 2017; 114(25):430. https://doi.org/10.3238/arztebl.2017.430 PMID: 28683860

6. Klaver EI, van Loon HC, Stienstra R, Links TP, Keers JC, Kema IP, et al. Thyroid hormone status and health-related quality of life in the LifeLines Cohort Study. Thyroid. 2013; 23(9):1066–73. https://doi.org/10.1089/thy.2013.0017 PMID: 23530992

7. Surks MI, Goswami G, Daniels GH. The thyrotropin reference range should remain unchanged. The Journal of Clinical Endocrinology & Metabolism. 2005; 90(9):5489–96. https://doi.org/10.1210/jc.2005-0170 PMID: 16148346

8. Biondi B. Natural history, diagnosis and management of subclinical thyroid dysfunction. Best practice & research Clinical endocrinology & metabolism. 2012; 26(4):431–46. https://doi.org/10.1016/j.beem.2011.12.004 PMID: 22863386

9. Jasim S, Abdi H, Gharib H, Biondi B. A Clinical Debate: Subclinical Hypothyroidism. International Journal of Endocrinology and Metabolism. 2021; 19(3):e115948. https://doi.org/10.5812/ijem.115948 PMID: 34567140

10. Vaidya B, Ukoumunne OC, Shuttleworth J, Bromley A, Lewis A, Hyde C, et al. Variability in thyroid function test requests across general practices in south-west England. Qual Prim Care. 2013; 21(3):143–8. PMID: 23968263

11. Wintemute K, Greiver M, McIsaac W, Del Giudice ME, Sullivan F, Aliarzadeh B, et al. Choosing Wisely Canada campaign associated with less overuse of thyroid testing: Retrospective parallel cohort study. Canadian Family Physician. 2019; 65(11):e487–e96. PubMed Central PMCID: PMC6853352. PMID: 31722930

12. Birk-Urovitz E, Del Giudice ME, Meaney C, Grewal K. Use of thyroid-stimulating hormone tests for identifying primary hypothyroidism in family medicine patients. Canadian Family Physician. 2017; 63(9):e389–e94. PubMed Central PMCID: PMC5597029. PMID: 28904049

13. UK National Guideline Centre. Thyroid disease: assessment and management 2019 [updated 20 November 2019]. Available from: https://www.nice.org.uk/guidance/ng145.

14. Bekkering G, Agoritsas T, Lytvyn L, Heen A, Feller M, Moutzouri E, et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. Bmj. 2019; 365. https://doi.org/10.1136/bmj.l2006 PMID: 31088853

15. Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RG, Mooijaart SP, et al. Thyroid hormone therapy for older adults with subclinical hypothyroidism. New England Journal of Medicine. 2017; 376(26):2534–44. https://doi.org/10.1056/NEJMoa1603825 PMID: 28402245

16. Vasileiou M, Gilbert J, Fishburn S, Boelaert K. Thyroid disease assessment and management: summary of NICE guidance. BMJ. 2020; 368:m41. https://doi.org/10.1136/bmj.m41 PMID: 31996347

17. Feller M, Snel M, Moutzouri E, Bauer DC, de Montmollin M, Aujesky D, et al. Association of Thyroid Hormone Therapy With Quality of Life and Thyroid-Related Symptoms in Patients With Subclinical Hypothyroidism: A Systematic Review and Meta-analysis. JAMA. 2018; 320(13):1349–59. https://doi.org/10.1001/jama.2018.13770 PMID: 30285179.

18. Ballard M, Montgomery P. Risk of bias in overviews of reviews: a scoping review of methodological guidance and four-item checklist. Research Synthesis Methods. 2017; 8(1):92–108. https://doi.org/10.1002/jrsm.1229 PMID: 28074553

19. Papatheodorou S. Umbrella reviews: what they are and why we need them. European journal of epidemiology. 2019; 34(6):543–6. https://doi.org/10.1007/s10654-019-00505-6 PMID: 30852716

20. Bauer BS, Azcoaga-Lorenzo A, Agrawal U, McCowan C. Management strategies for patients with subclinical hypothyroidism: a protocol for an umbrella review. Systematic reviews. 2021; 10(1):1–6. https://doi.org/10.1186/s13643-021-01842-y PMID: 33388080

21. Onishi A, Furukawa TA. State-of-the-art reporting. Umbrella reviews. 2016:189–202. https://dx.doi.org/10.1007/978-3-319-25655-9_13.

22. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017; 358:j4008. https://doi.org/10.1136/bmj.j4008 PMID: 28935701

23. Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. Journal of clinical epidemiology. 2016; 69:225–34. https://doi.org/10.1016/j.jclinepi.2015.06.005 PMID: 26092286

24. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. Journal of clinical epidemiology. 2019; 108:26–33. https://doi.org/10.1016/j.jclinepi.2018.12.004 PMID: 30543911

25. Perry R, Whitmarsh A, Leach V, Davies P. A Comparison of Two Assessment Tools Used in Overviews of Systematic Reviews: ROBIS versus AMSTAR-2. Systematic Reviews. 2021; 10(273). https://doi.org/10.1186/s13643-021-01819-x PMID: 34696810

26. Gates M, Gates A, Duarte G, Cary M, Becker M, Prediger B, et al. Quality and risk of bias appraisals of systematic reviews are inconsistent across reviewers and centers. Journal of clinical epidemiology. 2020; 125:9–15. https://doi.org/10.1016/j.jclinepi.2020.04.026 PMID: 32416337

27. Santaguida PL, Keshavarz H, Carlesso LC, Lomotan M, Gross A, MacDermid JC, et al. Suppl 4: A Description of the Methodology Used in an Overview of Reviews to Evaluate Evidence on the Treatment, Harms, Diagnosis/Classification, Prognosis and Outcomes Used in the Management of Neck Pain. The open orthopaedics journal. 2013; 7:461. https://doi.org/10.2174/1874325001307010461 PMID: 24115970

28. Hartling L, Chisholm A, Thomson D, Dryden DM. A descriptive analysis of overviews of reviews published between 2000 and 2011. PloS one. 2012; 7(11):e49667. https://doi.org/10.1371/journal.pone.0049667 PMID: 23166744

29. Pieper D, Antoine S-L, Mathes T, Neugebauer EA, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. Journal of clinical epidemiology. 2014; 67(4):368–75. https://doi.org/10.1016/j.jclinepi.2013.11.007 PMID: 24581293

30. Lunny C, Pieper D, Thabet P, Kanji S. Managing overlap of primary study results across systematic reviews: practical considerations for authors of overviews of reviews. BMC Medical Research Methodology. 2021; 21(140). https://doi.org/10.1186/s12874-021-01269-y PMID: 34233615

31. Hennessy EA, Johnson BT. Examining overlap of included studies in meta-reviews: Guidance for using the corrected covered area index. Research synthesis methods. 2020; 11(1):134–45. https://doi.org/10.1002/jrsm.1390 PMID: 31823513

32. Chaker L, Baumgartner C, Ikram MA, Dehghan A, Medici M, Visser WE, et al. Subclinical thyroid dysfunction and the risk of stroke: a systematic review and meta-analysis. European journal of epidemiology. 2014; 29(11):791–800. https://doi.org/10.1007/s10654-014-9946-8 PMID: 25179793

33. Helfand M, Redfern CC. Screening for thyroid disease: an update. Annals of internal medicine. 1998; 129(2):144–58. https://doi.org/10.7326/0003-4819-129-2-199807150-00020 PMID: 9669977.

34. Ochs N, Auer R, Bauer DC, Nanchen D, Gussekloo J, Cornuz J, et al. Meta-analysis: subclinical thyroid dysfunction and the risk for coronary heart disease and mortality. Annals of internal medicine. 2008; 148(11):832–45. https://doi.org/10.7326/0003-4819-148-11-200806030-00225 PMID: 18490668

35. Du Puy RS, Poortvliet RK, Mooijaart SP, Den Elzen WP, Jagger C, Pearce SH, et al. Outcomes of thyroid dysfunction in people aged eighty years and older: an individual patient data meta-analysis of four prospective studies (Towards Understanding Longitudinal International Older People Studies Consortium). Thyroid. 2021; 31(4):552–62. https://doi.org/10.1089/thy.2020.0567 PMID: 33012278

36. Helfand M. Screening for Thyroid Disease. Systematic Evidence Reviews [Internet]. 2004; No. 23. Available from: https://www.ncbi.nlm.nih.gov/books/NBK42812/. PMID: 20722119

37. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews [Internet]. Vol. 372, The BMJ. BMJ Publishing Group; 2021.

38. Reyes Domingo F, Avey MT, Doull M. Screening for thyroid dysfunction and treatment of screen-detected thyroid dysfunction in asymptomatic, community-dwelling adults: a systematic review. Systematic Reviews. 2019; 8(1):260. https://doi.org/10.1186/s13643-019-1181-7 PMID: 31735166.

39. 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. https://doi.org/10.7326/M14-1456 PMID: 25347444.

40. Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. Cochrane Database of Systematic Reviews. 2007;(3):CD003419. https://doi.org/10.1002/14651858.CD003419.pub2 PMID: 17636722.

41. Dhital R, Poudel DR, Tachamo N, Gyawali B, Basnet S, Shrestha P, et al. Ischemic Stroke and Impact of Thyroid Profile at Presentation: A Systematic Review and Meta-analysis of Observational Studies. Journal of Stroke & Cerebrovascular Diseases. 2017; 26(12):2926–34. https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.07.015 PMID: 28821377.

42. Peng CC, Huang HK, Wu BB, Chang RH, Tu YK, Munir KM. Association of Thyroid Hormone Therapy with Mortality in Subclinical Hypothyroidism: A Systematic Review and Meta-Analysis. Journal of

Clinical Endocrinology & Metabolism. 2021; 106(1):292–303. https://doi.org/10.1210/clinem/dgaa777 PMID: 33107557.

43. Sun J, Yao L, Fang Y, Yang R, Chen Y, Yang K, et al. Relationship between subclinical thyroid dysfunction and the risk of cardiovascular outcomes: a systematic review and meta-analysis of prospective cohort studies. International journal of endocrinology. 2017; 2017. https://doi.org/10.1155/2017/8130796 PMID: 29081800

44. Wirth CD, Blum MR, da Costa BR, Baumgartner C, Collet T-H, Medici M, et al. Subclinical thyroid dysfunction and the risk for fractures: a systematic review and meta-analysis. Annals of internal medicine. 2014; 161(3):189–99. https://doi.org/10.7326/M14-0125 PMID: 25089863

45. Blum MR, Bauer DC, Collet T-H, Fink HA, Cappola AR, da Costa BR, et al. Subclinical Thyroid Dysfunction and Fracture Risk: A Meta-analysis. JAMA. 2015; 313(20):2055–65. https://doi.org/10.1001/jama.2015.5161 PMID: 26010634

46. Razvi S, Shakoor A, Vanderpump M, Weaver JU, Pearce SH. The influence of age on the relationship between subclinical hypothyroidism and ischemic heart disease: a metaanalysis. The Journal of clinical endocrinology and metabolism. 2008; 93(8):2998–3007. https://doi.org/10.1210/jc.2008-0167 PMID: 18505765

47. Rodondi N, Aujesky D, Vittinghoff E, Cornuz J, Bauer DC. Subclinical hypothyroidism and the risk of coronary heart disease: a meta-analysis. American Journal of Medicine. 2006; 119(7):541–51. https://doi.org/10.1016/j.amjmed.2005.09.028 PMID: 16828622.

48. Singh S, Duggal J, Molnar J, Maldonado F, Barsano CP, Arora R. Impact of subclinical thyroid disorders on coronary heart disease, cardiovascular and all-cause mortality: a meta-analysis. International Journal of Cardiology. 2008; 125(1):41–8. https://doi.org/10.1016/j.ijcard.2007.02.027 PMID: 17434631.

49. Yan Z, Huang H, Li J, Wang J. Relationship between subclinical thyroid dysfunction and the risk of fracture: a meta-analysis of prospective cohort studies. Osteoporosis International. 2016; 27(1):115–25. https://doi.org/10.1007/s00198-015-3221-z PMID: 26223189

50. Yang G, Wang Y, Ma A, Wang T. Subclinical thyroid dysfunction is associated with adverse prognosis in heart failure patients with reduced ejection fraction. BMC cardiovascular disorders. 2019; 19(1):83. https://doi.org/10.1186/s12872-019-1055-x PMID: 30947691

51. Baumgartner C, da Costa BR, Collet TH, Feller M, Floriani C, Bauer DC, et al. Thyroid Function Within the Normal Range, Subclinical Hypothyroidism, and the Risk of Atrial Fibrillation. Circulation. 2017; 136(22):2100–16. https://doi.org/10.1161/CIRCULATIONAHA.117.028753 PMID: 29061566.

52. Chaker L, Baumgartner C, den Elzen WP, Ikram MA, Blum MR, Collet TH, et al. Subclinical Hypothyroidism and the Risk of Stroke Events and Fatal Stroke: An Individual Participant Data Analysis. Journal of Clinical Endocrinology & Metabolism. 2015; 100(6):2181–91. https://doi.org/10.1210/jc.2015-1438 PMID: 25856213.

53. Collet TH, Bauer DC, Cappola AR, Asvold BO, Weiler S, Vittinghoff E, et al. Thyroid antibody status, subclinical hypothyroidism, and the risk of coronary heart disease: an individual participant data analysis. The Journal of clinical endocrinology and metabolism. 2014; 99(9):jc20141250. https://dx.doi.org/10.1210/jc.2014-1250.

54. Gencer B, Collet TH, Virgini V, Bauer DC, Gussekloo J, Cappola AR, et al. Subclinical thyroid dysfunction and the risk of heart failure events an individual participant data analysis from 6 prospective cohorts. Circulation. 2012; 126(9):1040–9. https://doi.org/10.1161/CIRCULATIONAHA.112.096024 PMID: 22821943

55. Rodondi N, den Elzen WP, Bauer DC, Cappola AR, Razvi S, Walsh JP, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. JAMA. 2010; 304(12):1365–74. https://doi.org/10.1001/jama.2010.1361 PMID: 20858880.

56. Weetman AP. Graves' disease. New England Journal of Medicine. 2000; 343(17):1236–48. https://doi.org/10.1056/NEJM200010263431707 PMID: 11071676

57. Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Chapter 10: Umbrella Reviews: JBI; 2020. Available from: https://jbi-global-wiki.refined.site/space/MANUAL/3283910768/Chapter+10%3A+Umbrella+reviews.

58. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC medical research methodology. 2011; 11(1):1–6. https://doi.org/10.1186/1471-2288-11-15 PMID: 21291558

59. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Evaluation & the health professions. 2002; 25(1):76–97. https://doi.org/10.1177/0163278702025001006 PMID: 11868447

2.4 Recent publications

I repeated a section of the literature searches in August 2024 (only on MEDLINE and Epistemonikos) to identify relevant systematic reviews and meta-analyses published since July 2021. Using the same search strategy used in the umbrella review, I identified four publications that would have otherwise met the inclusion criteria.

A meta-analysis of 13 observational studies compared untreated and euthyroid participants and reported no significant association between untreated SCH and all-cause mortality for those over 60 years old (pooled HR 1.22, 95% CI 0.87 to 1.70), but treatment was associated with higher risk (pooled HR 1.15, 95% CI 1.01 to 1.30) [121]. For cardiovascular mortality, the results were not statistically significant for untreated vs euthyroid (pooled HR 0.95, 95% CI 0.44 to 2.04) and for the treated vs euthyroid analysis (pooled HR 1.07, 95% CI 0.93 to 1.24) [121]. These findings partly correspond to the umbrella review – in both, treated patients had a higher risk of all-cause mortality than euthyroid participants, but the review by Yang et al. [122] found evidence that untreated individuals were at higher risk than those who were euthyroid. The cardiovascular mortality estimates were similar to one primary review. It should also be noted that six of the 13 included studies were represented in the umbrella review, so overlap was unavoidable.

Holley at al. [123] investigated the impact of treatment vs non-treatment of SCH on cardiovascular outcomes and bone health among older participants (>50 years) and reported no significant association. These findings correspond with the umbrella review, but it must be noted that four of their seven included studies overlapped with the selected primary reviews.

In a systematic review and meta-analysis of 15 studies (1,199 individuals with SCH) on cognitive function in people older than 60, Akintola et al. [124] found no evidence that SCH is associated with cognitive impairment. One key distinction between this and the four studies included in the umbrella review for the cognitive function outcome is that this study also compared untreated SCH to euthyroidism. Even so, the findings across all the reviews (and comparisons) were consistent.

Similarly, van Vliet et al. [125] analysed individual participant data for the association between thyroid dysfunction and cognitive function. In a sensitivity analysis of SCH excluding patients on treatment, they found no significant difference between untreated and euthyroid participants in global cognitive function (SMD 0.04, 95% CI -0.02 to 0.10) and memory (SMD 0.08, 95% CI -0.01 to 0.17) but those with SCH performed better in executive function (SMD 0.09, 95% CI 0.01 to 0.16). However, these results were based on various tests, which limited their comparability to the umbrella review.

2.5 Additional material

A cursory search of the published literature on SCH and long-term outcomes at the start of this PhD revealed that several systematic reviews had been published, thereby introducing the challenge of exploring the breadth of and summarising their results. I therefore took this opportunity to perform an umbrella review in the interest of: (i) presenting an overview of the SCH literature, (ii) identifying priorities for the studies to follow, and (iii) broadening my methodological expertise in evidence synthesis.

Umbrella review methodology was developed in light of the rapid and continuous expansion of systematic reviews and meta-analyses to allow for the clustering of the evidence covered by the same heading [126] or, as described by Choi and Kang [127], the 'pouring rain of evidence'. Indeed, it has been reported that there was a more than 20-fold increase in published systematic reviews between 2000 and 2019 [128], and in 2022 alone, there were over 19,000 [129]. Analogously, Slim and Marquillier [130] ran a MEDLINE search on May 1, 2021, showing 1,999 articles indexed under the Umbrella Review heading, which indicated their increasing popularity.

As recommended for all types of systematic reviews, I wrote a protocol detailing the research questions and Population, Intervention, Comparator, Outcome (PICO)-based inclusion criteria, the search strategy and relevant electronic databases and quality assessment for the umbrella review. The protocol was registered on PROSPERO [131] and subsequently peer-reviewed and published [119]. Although several umbrella reviews have been published in the last decade, and guidelines have been developed for how they should be conducted [132–134], as with any relatively novel research methodology, problems persist [135]. The section below describes my main takeaways from conducting the umbrella review in this thesis.

2.5.1 Search strategy

The search strategies for umbrella reviews do not substantially differ from those of the 'traditional' systematic review. In addition to standard electronic databases, I also used Epistemonikos [136], a unique resource specialising in identifying, screening and storing systematic reviews for health research. The availability of a systematic-reviews database was beneficial during the searches to ensure that as many relevant studies as possible were identified. The use of controlled vocabulary and subject headings may not always yield exhaustive search results, hence the need to conduct searches on both principal and supplementary resources [137]. The best combination of electronic databases for searching for systematic reviews has been found to be MEDLINE and Epistemonikos [138].

2.5.2 Article word limits

Many journals impose word limits on published articles, a lingering element from the old practice of producing journals only in print. Despite the ubiquitousness of online journals in recent years, the word counts permitted for systematic reviews and meta-analyses do not always allow for comprehensive reporting of methods and results. As a result, it was sometimes challenging to identify crucial aspects such as inclusion criteria and quality assessments, particularly for older studies that often did not include Supplementary Material and where the corresponding authors' details were no longer accurate. Thus, some reviews were necessarily excluded based on missing data.

Though it is possible to refer to the empirical studies in the systematic review as suggested by some authors [139], this introduces additional complexity to the process by potentially altering the evidence. After all, the unit of analysis for an umbrella review is the systematic review, not the primary studies [140]. The option to redo the analyses in the systematic

reviews would only be feasible where the number of included studies is low to avoid scope creep. In my case, the levels of overlap and heterogeneity across the included systematic reviews negated the need to repeat the reported analyses.

2.5.3 Study quality

The assessment of systematic review quality is a well-researched topic that has resulted in the widespread use of tools such as AMSTAR-2 and ROBIS [141–143]. On the other hand, when assessing the quality of a particular piece of clinical research, even the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework [144, 145], a tool widely used for developing clinical guidelines, is not infallible [146, 147]. For instance, personal and professional values can be expected to vary between practitioners, so objectivity is challenging. Also, it is problematic to assume higher quality or confidence purely based on study design (e.g. RCTs) because relying on the cumulative value of the evidence, i.e. 'how many studies support this finding?', risks pooling the biases from selected studies, whether or not they are RCTs [148].

One of the findings from the umbrella review was that for the topic of SCH and its association with long-term outcomes, many empirical studies have been of poor quality [120]. This is concerning because the same studies have then been used in several systematic reviews, echoing potentially misleading results. This problem likely ties in with the discordance and controversy regarding SCH and its management under the reasonable assumption that synthesised evidence can only be as good as the data used to obtain it. Even so, the studies whose quality is less than optimal justify further investigation into this topic, including a project such as this thesis.

2.5.4 Study overlap

I used the corrected covered area (CCA) method [149–151] to document overlap in the umbrella review on SCH and long-term outcomes. In general, assessing overlap in umbrella reviews directly depends on the quality of reporting in selected systematic reviews and meta-analyses. It was not always clear which data from the primary studies was used, such as review papers which reported 'data not shown' for subgroup analyses. As described in Chapter 3, it is also likely that the overall CCA for the umbrella review was underestimated because it was impossible to group the included reviews by outcome for the assessment.

There is some guidance on managing overlap in umbrella reviews besides ensuring that it is evaluated and reported for transparency [149, 150, 152]. The options include selecting the most recent review, updating the selected primary reviews or using a quality threshold [149]. However, CCA thresholds require further validation [151], and these strategies can conflict with each other, for example, (i) when analysing several outcomes, some of which are restricted to poor quality systematic reviews and meta-analyses, or (ii) when the most extensive or most recent eligible review is found to be of lower quality than smaller or older syntheses. In the interest of improving the conduct of umbrella reviews as well as avoiding research waste [153], better recommendations are needed, particularly if "the large majority of produced systematic reviews and meta-analyses are unnecessary, misleading, and/or conflicted" [154].

2.6 Conclusion

True to the name, this umbrella review provided an overview of the extant literature on how clinical outcomes are affected by the management of SCH, crucially comparing treatment to non-treatment and SCH to euthyroidism. The key messages from the results were that treatment might be associated with lower all-cause mortality and ischemic heart disease for patients under 70 years old, but the degree of overlap across the reviews could obfuscate the strength of these associations. Evidence on the other outcomes was mixed but predominantly not significant.

Another highlight was that though the body of evidence covered all the outcomes of interest, the most frequently evaluated were all-cause mortality and cardiovascular outcomes. There were no systematic reviews comparing stroke, for instance, between treated and untreated individuals with SCH. This provided further justification for investigating a range of outcomes for the trial emulation in Chapter 4 to build the evidence base.

3

Descriptive cohort study

3.1 Introduction

This chapter details the first of two studies in this thesis that used routinely collected heath data – EHR from the SAIL Databank. In the following sections, I describe the benefits and challenges associated with using EHR for research, followed by the data source, SAIL Databank. This is followed by the published article of a descriptive study on SCH in Wales between January 2000 and December 2021. This paper was peer-reviewed and published in PLoS ONE as follows [155]:

1. **Bauer BS**, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C. Subclinical hypothyroidism in Wales from 2000 to 2021: A descriptive cohort study based on electronic health records. PLOS ONE. 2024 May 21;19(5):e0298871. https://doi.org/10.1186/s13643-021-01842-y

 Supplementary materials are included in Appendix II.

3.2 Methods

3.2.1 EHR-based research

The benefits of using EHR for research are that they contain data on populations that tend to be underrepresented in RCTs [156], provide large sample sizes with information on a range of clinical data [157], provide longitudinal coverage and because the data are collected in real time, there is minimal to no risk of recall bias. Data linkage across datasets also allows for a more comprehensive view of a person's medical history.

EHR systems develop as time passes, so the richness of the data collected today is likely higher – and therefore more useful – than the data collected when there were fewer functionalities. On a related note, the dimensionality of EHR data provides opportunities to investigate the impact of a broader range of covariates than may have otherwise been possible [158]; the temporal span of the records enables the performance of longitudinal studies with varying lengths of follow-up [159].

EHR-based research can also be cheaper than bespoke data collection in terms of funding, time and the facilities required [159], but this would likely depend on the size of the respective studies. To put this in context, convenience sampling of patients attending a GP practice for a small cohort study could be more cost-effective than purchasing a multi-study licence from a data provider.

However, (i) EHR data are not collected for research [157] and therefore differ from 'purpose-driven' primary epidemiological data; (ii) larger sample sizes may increase statistical precision but do not eliminate confounding and bias (selection bias, misclassification or measurement bias (ascertainment bias) [160, 161] and (iii) the probability of missing data is high, hence there may be unmeasured confounding [159].

The lack of a diagnostic code in an individual's EHR, for example, does not necessarily mean they do not have a particular condition. Instead, possible alternative explanations

include that: (i) the diagnosis was not recorded, (ii) an incorrect code was used, or (iii) the person did not seek treatment and, therefore, the diagnosis was not possible – this is also referred to as 'informative presence' [162].

It is sometimes challenging to accurately define the length of follow-up [163], that is, baseline, and if or when an individual should be censored from the study because of attrition. The recorded information may not match what happened if, for instance, someone moved away but then forgot to update their GP registration until the next month and vice versa. In this scenario, the former would result in right-censored EHR data, whereas the latter would fit the definition of left-censoring, giving an incomplete timeline of events.

3.2.2 SAIL Databank

Primary care coverage was the top priority when selecting a data source for this project because GPs manage SCH. As such, it was imperative that the selected resource could provide comprehensive patient data comprising primary care records and the ability to link these data to other health records for extensive EHR-based phenotyping. The Secure Anonymised Information Linkage (SAIL) Databank, a repository of health records for the population of Wales, boasted GP practice coverage of ~80%, or population coverage of 83% as of 2021 [164]. Currently, none of the other UK nations have almost complete primary care coverage [165].

According to the 2021 UK census, the population of Wales was 3.1m, representing 1.4% growth since the previous census in 2011 [166]. SAIL provides population-wide data, which was advantageous for the study cohort size because it would allow for large-scale studies on SCH. One of the critical challenges behind the controversy surrounding the management of SCH has been the small sample sizes in the published literature. The temporal coverage of the SAIL datasets was another significant positive because it would provide relatively long follow-up periods, which, as the name suggests, would be vital for evaluating long-term health outcomes.

Finally, my supervisors and I considered the costs of obtaining and accessing EHR data. Due to a heavily subsidised pricing model, SAIL Databank was considerably cheaper than similar data providers such as CPRD. For example, at the time, whereas an annual single-study dataset licence from CPRD would have cost roughly £15,000 (+ VAT) plus additional costs for data linkage, the data costs for this project amounted to £13,000 (+ VAT) for three years' access to SAIL data.

3.2.2.1 Application process

The SAIL Databank Information Governance Review Panel (IGRP) reviews all applications and proposals to access SAIL data to ensure that appropriate information governance controls have been considered and will be maintained. The IGRP comprises representatives from several institutions (Welsh Government, Public Health Wales, Health and Care Research Wales, Digital Health and Care Wales, British Medical Association Cymru Wales, Local Health Boards, academic institutions and lay members) [167, 168]. Applications are also assessed by the Health Information Research Unit (HIRU) for data availability and study feasibility [167].

After a scoping discussion with SAIL Databank staff, I submitted my IGRP application on December 14th, 2021; approval was granted on February 1st, 2022. Subsequently, data provisioning was completed on February 3rd, 2022.

3.2.2.2 Requested datasets

Following IGRP approval, access was granted to the following eight datasets in the SAIL Databank (Table 3.1 & Figure 3.1).

Welsh Demographic Service Dataset (WDSD)

Demographic data is routinely obtained from the Digital Health & Care Wales (DHCW) (formerly NHS Wales Informatics Service), custodians for all identifiable data on Welsh residents. Before 2009, this information was contained in the NHS Wales Administrative Register.

For the thesis, WDSD was used as the population register – a record of people living in Wales who were registered with a Welsh GP at any point during the study period, their address and registration history, dating back to January 2000.

Table 3.1. *Details of the datasets requested from SAIL Databank.*

SAIL Dataset	Setting	Data Collector	Frequency of Updates	Coverage Start Date	Pathway
Annual District Death Extract (ADDE)	All of Wales	ONS	Monthly	01/01/1996	-
Emergency Department Dataset (EDDS)	Accident & Emergency	DHCW	Monthly	01/04/2009	Secondary Care
Outpatient Database for Wales (OPDW)	Outpatient	DHCW	Monthly	01/04/2004	Secondary Care
Outpatient Referral Dataset (OPRD)	Referrals	DHCW	Monthly	01/04/2009	Secondary Care
Patient Episode Dataset for Wales (PEDW)	Hospitals	DHCW	Monthly	01/04/1995	Secondary/Tertiary Care
Welsh Demographic Service Dataset (WDSD)	All of Wales	DHCW	Weekly	01/01/1990	-
Wales Results Reporting Service (WRRS)	Test Clinics	DHCW	Weekly	01/06/1992	Primary/Secondary Care
Welsh Longitudinal General Practice Dataset (WLGP)	GP Practices (83%)/ 86% of the population	EMIS, Vision, Informatica	Monthly	01/01/2000	Primary Care

Note:
Abbreviations: DHCW – Digital Health & Care Wales (formerly NHS Wales Informatics Service); EMIS – (formerly Egton Medical Information Systems); ONS – Office for National Statistics (Source: Health Data Research UK (HDRUK) Gateway).

Welsh Longitudinal General Practice (WLGP)

This dataset contains patient interactions with GPs, including their presenting signs and symptoms, tests performed, diagnoses, prescriptions and referrals. Crucially, it should be noted that this dataset only covers 86% of the Welsh population due to factors such as migration and registration [164]. Only 83% of GP practices in the country provide this data, given that it is an opt-in service. As such, the timespan covered within the records varies, depending on the practice.

Patient-related information is typically recorded at the GP's discretion during the consultation, whereas investigation results are collected directly from the providers' systems. Most of this information is recorded as Read codes, but research has shown discrepancies in their use [169–172].

Given that SCH is diagnosed and managed by GPs, I used this dataset to identify patients with clinically recorded diagnoses, comorbidities, and medical history, both of which were relevant for selection criteria and study follow-up. On advice from SAIL, the length of the study period was based on WLGP – I was informed of a marked improvement in data quality from 2000, which was then selected as the start date (Figure 3.1).

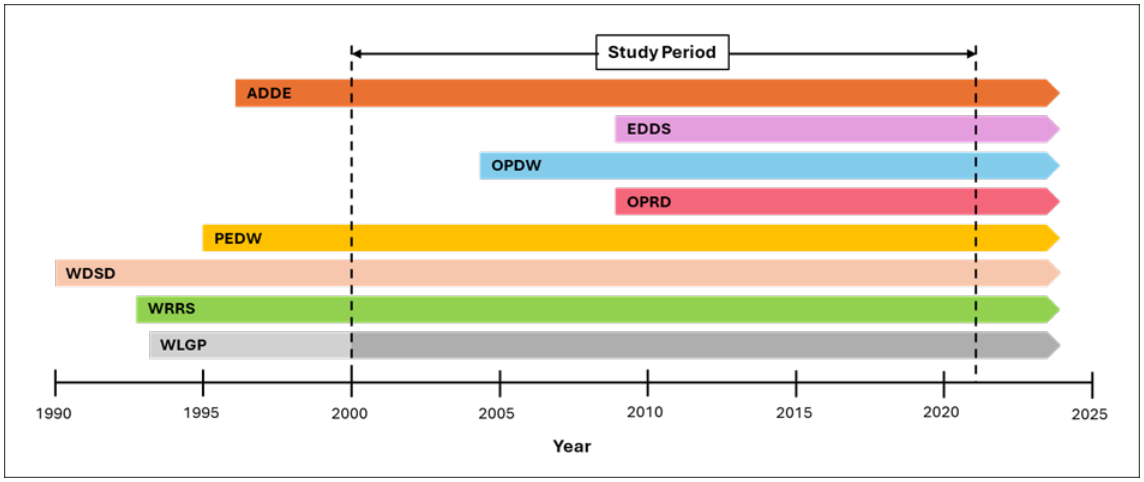


Figure 3.1. *Temporal coverage of the SAIL Databank datasets.*

Annual District Death Extract (ADDE)

ADDE is a single table containing data on all the deaths of people resident in Wales, even if the actual event occurred outside the country, as obtained from the national death registrations. I used this data to identify the date of death where relevant and to confirm the recorded dates in WLGP for patients who passed away during the study period.

I used the following datasets: (i) to identify pre- and post- baseline comorbidities when defining the study cohorts and identifying covariates, (ii) to detect the occurrence of events for the outcomes of interest, and (iii) to ascertain the recorded diagnoses across datasets.

Patient Episode Database for Wales (PEDW)

PEDW contains the records of all NHS Wales hospital day cases and in-hospital admissions, both inpatient and day cases. It also contains data on Welsh residents who receive treatment in England. Each hospital is responsible for the collection and coding of data, so data quality varies within the dataset, particularly regarding secondary diagnoses. ICD-10 diagnostic codes are used in the PEDW tables.

Outpatient Database for Wales (OPDW)

This dataset contains patient data from all hospital outpatient departments, collected separately from each of the hospitals in Wales. All scheduled outpatient appointments are recorded, whether the patient attends the hospital or not. Diagnoses are coded using ICD-10 codes.

Outpatient Referral Dataset (OPRD)

OPRD contains data on outpatient referrals made by primary care practitioners (GPs, General Dental Practitioners (GDPs), A&E Departments) or self-referrals in a single table.

Emergency Department Dataset (EDDS)

Provided by DHCW, this dataset holds the records of persons attending Accident and Emergency departments in Welsh hospitals. It includes records from the All Wales Injury Surveillance Systems (AWISS) dataset but is provided as a single table in which diagnoses are recorded using ICD-10 codes.

Wales Results Reporting Service (WRRS)

The WRRS dataset has all the laboratory results for blood tests requested within Wales, whether from primary or secondary care pathways. The unified Welsh Laboratory Information Management System (WLIMS) was designed to simplify healthcare provider access to the records and patient access to local – and therefore more convenient – testing facilities [173]. The data is structured into three tables (requests, results and reports).

3.2.2.3 Data structure

SAIL records are provided as split files, which can be linked using the Anonymous Linking Field (ALF) generated from NHS registration numbers [167, 174]. SAIL data providers separate the identifiable demographic data from the clinical data in the first instance. The ALF is then added to the datasets to ensure that records from different SAIL datasets can be linked within the secure environment [167]. The unique identifier is also encrypted to form ALF-E (Encrypted) [175].

After IGRP approval, the ALF-E is further encrypted for each project to form ALF_PE [167]. This final step enhances security by deterring cross-linkage of data across different projects. Therefore, two separately approved projects cannot 'share' access to SAIL records because the ALF_PE fields would not match [175, 176].

3.2.2.4 Data access

Per standard practice, my access to the provisioned data views was restricted to the secure SAIL Gateway [175], which required using a Yubikey security token that SAIL provided. Secure login credentials were also provided after my primary supervisor and I signed a Data Access agreement (Appendix II) with no authorisation to share or disclose them to third parties.

3.2.2.5 Data management

Data preparation and cohort definition were performed for the respective studies using the relevant selection criteria described in Chapters 3 and 4.

3.2.2.6 Ethical considerations

This project used only anonymised, routinely collected data from the SAIL Databank, so there was no direct contact with Welsh residents or patients. In order to gain access to the data, an IGRP application was required, in addition to completing a Safe Researcher Training course.

Ethical approval for the study was sought from the University of St Andrews School of Medicine Ethics Committee (SEC). Approval to proceed with the project was granted on February 1st, 2022 from SAIL and March 14th, 2022 from the SEC. A separate application submitted to access the SAIL Gateway was approved on March 1st, 2022. These documents, as well as a signed copy of the SAIL Data Access Agreement, are included in Appendix II.

3.3 Published article

PLOS ONE

RESEARCH ARTICLE

Subclinical hypothyroidism in Wales from 2000 to 2021: A descriptive cohort study based on electronic health records

Brenda S. Bauer[1]*, Amaya Azcoaga-Lorenzo[1,2], Utkarsh Agrawal[3], Adeniyi Francis Fagbamigbe[4], Colin McCowan[1]

1 Division of Population and Behavioural Sciences, University of St Andrews School of Medicine, St Andrews, United Kingdom, **2** Hospital Rey Juan Carlos, Instituto de Investigación Sanitaria Fundación Jimenez Diaz, Red de Investigación en Cronicidad, Atención Primaria y Promoción de la Salud (RICAPPs) ISCIII, Madrid, Spain, **3** Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom, **4** Research & Evaluation Unit, Institute for Health and Wellbeing, Centre for Healthcare and Communities, Coventry University, Coventry, United Kingdom

* bsb1@st-andrews.ac.uk

OPEN ACCESS

Citation: Bauer BS, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C (2024) Subclinical hypothyroidism in Wales from 2000 to 2021: A descriptive cohort study based on electronic health records. PLoS ONE 19(5): e0298871. https://doi.org/10.1371/journal.pone.0298871

Editor: Fateen Ata, Hamad Medical Corporation, QATAR

Received: October 6, 2023

Accepted: February 1, 2024

Published: May 21, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0298871

Data Availability Statement: The data used for this study are third-party data provided by SAIL Databank. Access to the datasets is conditional upon application and review by their Information

Abstract

Background

Subclinical hypothyroidism (SCH) is a biochemical thyroid disorder characterised by elevated levels of Thyroid Stimulating Hormone (TSH) together with normal levels of thyroid hormones. Evidence on the benefits of treatment is limited, resulting in persistent controversies relating to its clinical management.

Aim

This study describes the demographic and clinical characteristics of patients identified as having subclinical hypothyroidism in Wales between 2000 and 2021, the annual cumulative incidence during this period and the testing and treatment patterns associated with this disorder.

Methods

We used linked electronic health records from SAIL Databank. Eligible patients were identified using a combination of diagnostic codes and Thyroid Function Test results. Descriptive analyses were then performed.

Results

199,520 individuals (63.8% female) were identified as having SCH, 23.6% (n = 47,104) of whom received levothyroxine for treatment over the study period. The median study follow-up time was 5.75 person-years (IQR 2.65–9.65). Annual cumulative incidence was highest in 2012 at 502 cases per 100,000 people. 92.5% (n = 184,484) of the study population had TSH levels between the upper limit of normal and 10mIU/L on their first test. 61.9% (n = 5,071) of patients identified using Read v2 codes were in the treated group. 41.9% (n =

Governance Review Panel. Interested parties can access the datasets by following the same application process as the authors, as detailed on the SAIL Databank website (https://saildatabank.com/data/apply-to-work-with-the-data/). The authors did not have any special access privileges that other applicants would not have.

Funding: This study was performed as part of a PhD studentship funded by the School of Medicine, University of St Andrews. The University of St Andrews played no role in the design and conduct of the study.

Competing interests: The authors have declared that no competing interests exist.

19,716) of treated patients had a history of a single abnormal test result before their first prescription.

Conclusion

In Wales, the number of incident cases of SCH has risen unevenly between 2000 and 2021. Most of the study population had mild SCH on their index test, but more than a third of the identified patients received levothyroxine after a single abnormal test result. Patients with clinically recorded diagnoses were more likely to be treated. Given the expectation of steadily increasing patient numbers, more evidence is required to support the clinical management of subclinical hypothyroidism.

Introduction

The thyroid gland is a butterfly-shaped endocrine organ in the neck whose function is regulated by Thyroid Stimulating Hormone (TSH) from the pituitary gland. It produces the thyroid hormones thyroxine (T4) and triiodothyronine (T3), which primarily regulate bodily metabolic function. Hypothyroidism generally refers to a deficiency of these hormones but can be subdivided into overt and subclinical types. The latter, subclinical hypothyroidism (SCH), is a frequently asymptomatic condition in which the thyroid hormone levels are within normal range, but TSH is elevated [1, 2].

Reference ranges for TSH vary due to patient characteristics, particularly age, sex, race, ethnicity and pregnancy status [3, 4]. There is disagreement over what should be considered as the upper limit of normal [5–9], so these values tend to differ between laboratories, given the lack of universally applicable guidelines. However, a commonly used upper cut-off for TSH is 4.5 mIU/L [10], such that patients with measurements between 4.5 and 9.9 mIU/L are said to have mild SCH. On the other hand, TSH levels $\geq$10 mIU/L are classified as severe SCH [2].

Among the causes of SCH, Hashimoto's thyroiditis, an autoimmune thyroid disorder, is the most frequent. Women are more likely to develop SCH than men, regardless of age [2, 11]. Previous studies have also reported that the population prevalence of SCH ranges between 4% and 10%, depending on factors such as age and sex distribution patterns [1, 11–13]. The Colorado Thyroid Disease Prevalence Study determined that in their study population of over 25,000 patients, approximately 9% of subjects not on thyroid medication were found to have SCH [11]. The US National Health and Nutrition Examination Survey (NHANES III) reported SCH prevalence figures of 4.3% in the total population [13].

It has been reported that around 2% to 6% of SCH cases each year experience progression to overt hypothyroidism. This phenomenon occurs more commonly among female patients and those found to test positive for thyroid peroxidase antibodies, for instance, in autoimmune thyroid disease [1, 14]. As reported in some studies, the TSH levels of approximately 60% of patients identified as having mild SCH may later spontaneously revert to normal [2, 15].

The pharmacological treatment for overt and subclinical hypothyroidism is levothyroxine (LT4), a synthetic version of thyroxine. Controversy persists on the clinical management of SCH, specifically around whether to initiate treatment and, in those cases, what level of TSH to use as a threshold. The debate is due to insufficient robust evidence and conflicting study findings on the long-term benefits–and, inversely, potential harms–of treatment for this disorder [1, 10, 16–19]. For example, many studies have investigated the effects of SCH on cardiovascular disease, but while some found that levothyroxine lowered the incidence of myocardial

infarctions, atrial fibrillation and cardiovascular mortality [20], others did not [21, 22]. Similarly, Mooijaart et al. reported that in 2 randomised trials of levothyroxine treatment for patients 80 years and above, there was no significant difference in quality of life (QoL) between treatment and control groups [23]; whereas in a cohort study of 78 patients, Winther et al. found marked improvements in health-related QoL within six months of starting treatment [24]. Even so, current NICE guidelines state that treatment should only be commenced after a repeat abnormal TSH result of ≥10 mIU/L three months after the first [25].

The debate extends to screening for SCH, despite the frequency of no reported symptoms, due to the lack of substantiated evidence of benefits for the majority who might be diagnosed through screening programs [2, 26, 27]. Other issues that are pertinent to the management of SCH are the overuse of thyroid function tests, also known as overtesting, which can potentially increase the detection of elevated TSH [28, 29], and the overuse of levothyroxine for SCH [19, 30], with no clarity on how it affects patients in the long-term.

Our study aims were to (i) describe the demographic and clinical characteristics of patients identified as having SCH in Wales between 2000 and 2021; (ii) estimate the annual cumulative incidence of SCH and the accumulation of patients identified as having SCH during the study period; (iii) characterise thyroid function testing and levothyroxine prescribing over the study period, including the TSH thresholds used to initiate treatment for SCH.

Methods

Study design

This is a retrospective, population-based cohort study using the anonymised, linked electronic health records (EHR) for the population living in Wales between 1 January 2000 and 31 December 2021. These were individuals registered with a General Practitioner (GP) practice contributing to the Secure Anonymised Information Linkage (SAIL) Databank.

Data source

The data source was the SAIL Databank, a repository of anonymised patient records for the population of Wales, representing approximately 5 million individuals between January 2000 and December 2021 [31]. Data is collected from approximately 84% of all GP practices in the country. The development, database structure, policies and requisite procedures governing the use of SAIL data have been detailed previously [32–34]. Approval was granted by the Information Governance Review Panel (IGRP) in February 2022 (ref 1371) for a study on SCH and clinical outcomes to be performed using SAIL datasets. The primary care, hospital, outpatient, emergency department, death, demographic and test result datasets (S1 Appendix) were first accessed on 14 March 2022.

Study cohort

Cohort entry was defined by the presence of one or more of: SCH Read v2 codes, SCH International Classification of Diseases version 10 (ICD-10) codes or test results indicative of SCH within the study timespan. For test results, the respective lab reference ranges were used as recorded to identify the upper limit of TSH and normal levels of T4 (S2 Appendix). Subjects could enter the cohort at any point between 1 January 2000 and 31 December 2021 (Fig 1). Exit from the study occurred at the earliest of: (i) death, (ii) censoring due to the end of GP registration or emigration from Wales or (iii) the end of the study period.

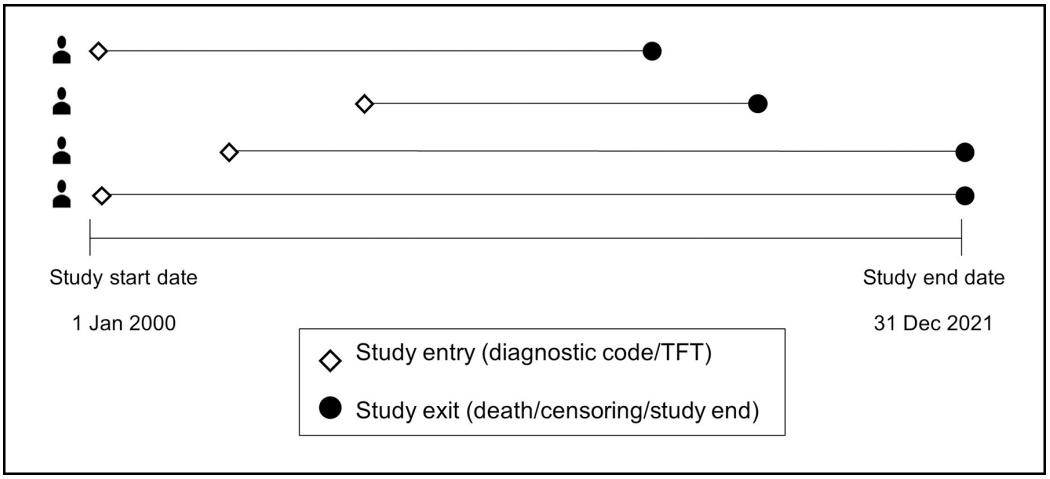


Fig 1. Study design and examples of possible patient pathways.

https://doi.org/10.1371/journal.pone.0298871.g001

Eligibility criteria

To be included in the study, patients must have had a recorded SCH diagnostic code or test result indicative of SCH. They must also have been registered with a GP for at least 12 months before the index SCH code or test result.

The following groups were excluded: women with recorded pregnancies within 12 months of the index SCH code or test, patients under 18 years old, those that had ever been identified as having overt hypothyroidism, those given prescriptions for thyroid-altering medications (amiodarone, lithium and antithyroid drugs) or thyroid hormone replacement (more than 30 days before the index code or test) and those with histories of radioiodine and thyroidectomy, such that there were no other indications for the use of levothyroxine.

Study variables

Baseline characteristics were assessed on the index diagnostic code or test date, which was set as the 'date of identification'. Patients were further classified into two mutually exclusive groups based on whether they received prescriptions for thyroid hormone replacement at any point after the date of identification ('treated') or did not receive any treatment ('untreated').

Sex was recorded as either male or female. Age was calculated from demographic records as the difference in years between the week of birth date and date of identification. This variable was also categorised into bands spanning ten years each. Deprivation scores were enumerated using the 2019 version of the Welsh Index of Multiple Deprivation (WIMD 2019) and assigned based on the recorded home postcode on the date of identification [35, 36]. Length of follow-up was calculated as the difference between the study start date and the exit date or study end date.

Additional SCH-related characteristics such as TSH levels, hypothyroidism codes, and prescribed medications were derived from the primary care or test result datasets. The presence of $\geq$1 thyroid hormone prescription was used as a proxy indicator for treatment status. The number of TFTs performed before and after initiating levothyroxine was also calculated for treated patients. This was done to gauge the frequency with which the NICE treatment

guideline–a repeat abnormal TSH result of ≥10 mIU/L after three months before commencing treatment–was followed.

Where available, thyroid peroxidase (TPO) antibody measurements were extracted from the test results dataset and compared to normal ranges. All references to TFT results in this text required that TSH and FT4 results share the same specimen collection date. Hence, 'normal TFT result(s)' refers to TSH and FT4 having both been within their respective reference ranges, whereas 'abnormal TFT result(s)' represents the elevated TSH and normal FT4 characteristic of SCH. Rather than selecting fixed study thresholds for thyroid hormones and antibodies, the reported lab reference ranges were used to align our identification of SCH with the results clinicians would have received.

Statistical analy5sis

Descriptive statistics were used to summarise baseline demographic data and clinical characteristics. For categorical variables, counts and percentages were used; means and standard deviations were employed for continuous variables.

Estimates of annual cumulative incidence were obtained for each year between 2000 and 2021; this was calculated as the number of newly identified SCH cases between 1 January and 31 December divided by the total number of GP-registered individuals as of 1 July (mid-year population) and multiplied by 100,000. The mid-2011 Welsh population was used to obtain age- and sex-standardised estimates.

The frequencies of normal and abnormal TFT results per patient were classified as prior to or later than the date of identification. We reported the total number of TFTs in the wider GP-registered population to explore if there was a relationship to the number of tests ordered in the same year for the study population. For the former, the study eligibility criteria were applied to all patients who had recorded TFT results between 2000 and 2021 –the number of tests per year was then calculated, irrespective of whether SCH was detected.

The frequency of levothyroxine prescriptions each year and the time between the date of identification and the first prescription were also calculated.

Structured query language (SQL DB2) was used to retrieve and interrogate the SAIL datasets via the SAIL Gateway [37].

Ethical approval

The project was approved in writing by the SAIL Databank IGRP (ref 1371) and by the School of Medicine Ethics Committee, acting on behalf of the University of St Andrews Teaching and Research Ethics Committee (UTREC) (MD16055) in March 2022. As the study data are de-identified, authors had no access to disclosive information and consent from individual patients was not required.

The results are reported using the REporting of studies Conducted using Observational Routinely-collected Data (RECORD) guidelines [38].

Results

Between January 2000 and December 2021, over 5.6 million individuals were identified in the demographic register, 304,148 of whom had recorded diagnostic codes or tests indicative of SCH. After applying the eligibility criteria resulting in the exclusion of 104,628 patients, 199,520 individuals were identified as incident SCH cases over the study period (Fig 2). The proportion of untreated patients (n = 152,416; 76.4%) was more than three times that of those who received levothyroxine over the study period (n = 47,104; 23.6%).

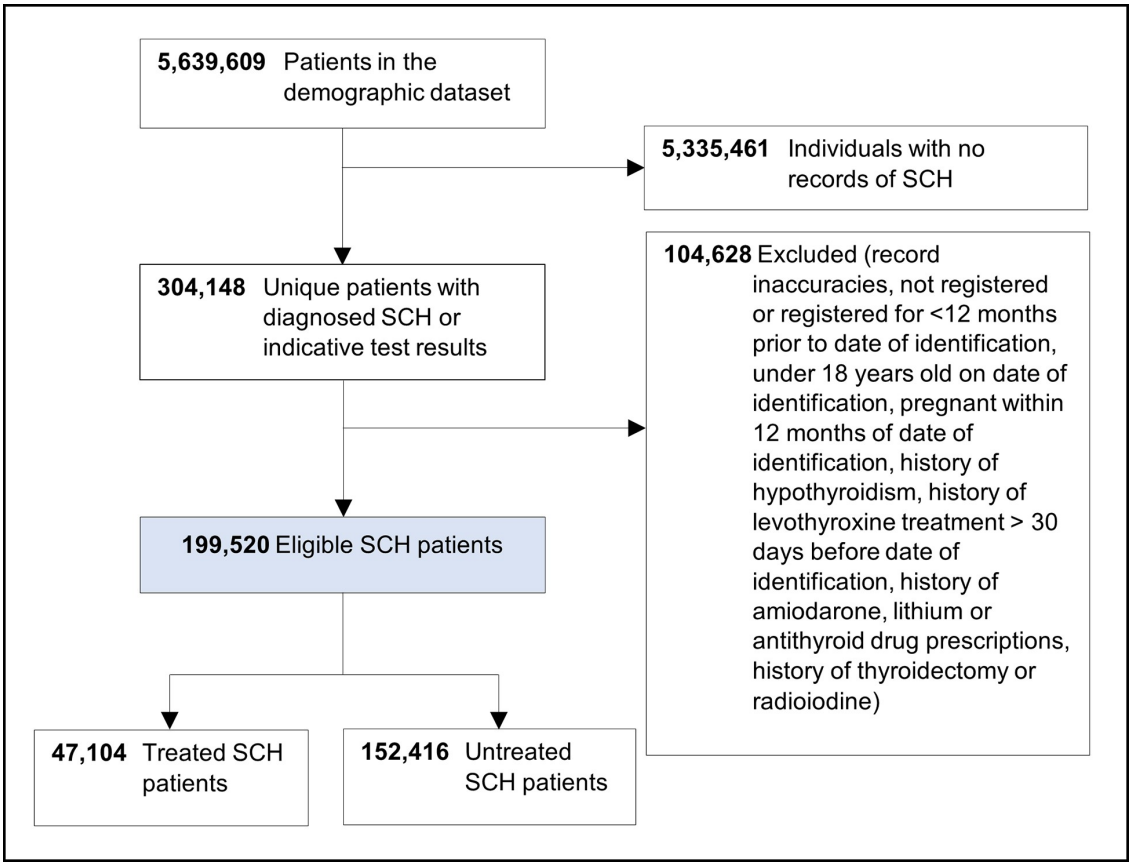


Fig 2. Study flowchart illustrating the selection of eligible SCH patients using SAIL data.

https://doi.org/10.1371/journal.pone.0298871.g002

Patient characteristics

The baseline characteristics of the study population are shown in Table 1. There were more female than male patients in the entire cohort (63.8% vs 36.2%). The mean age of study participants was 57.8 years, with a standard deviation of 17.55 years. The highest patient numbers were observed in the 60–69 age band (n = 39,486; 19.8%). There were more untreated than treated patients across all age bands.

The total length of follow-up for the study population was 1,286,883 person-years, with a median duration of 5.75 person-years (IQR 2.65–9.65). 24.6% of the study population died before the end of the study, as opposed to those censored (n = 7,715; 3.9%) because they moved away from Wales or switched to non-SAIL GP practices. Most patients had records running to the end of the study period, 31 December 2021 (n = 142,687; 71.5%).

There was considerable overlap between categories based on the means of identification from the EHR. 99.2% of patients had recorded TFT results on the date of identification; only 149 patients had a combination of all three criteria (Table 1).

Table 1. Patient characteristics classified by treatment status at baseline and the end of the study period (2000–2021) and the methods used to identify patients using EHR.

		Study population		
		Total (%)[a]	**Treated (%)[b]**	**Untreated (%)[b]**
	N	199,520	47,104 (23.6)	152,416 (76.4)
Sex				
	Male	72,175 (36.2)	12,169 (16.9)	60,006 (83.1)
	Female	127,345 (63.8)	34,935 (27.4)	92,410 (72.6)
Age (years), mean [SD]		57.8 [17.55]	54.7 [16.79]	59.5 [17.71]
Age bands (years)				
	18–29	16,949 (8.5)	3,529 (20.8)	13,420 (79.2)
	30–39	15,478 (7.8)	4,583 (29.6)	10,895 (70.4)
	40–49	25,667 (12.9)	7,878 (30.7)	17,789 (69.3)
	50–59	34,905 (17.5)	9,696 (27.8)	25,209 (72.2)
	60–69	39,486 (19.8)	9,175 (23.2)	30,311 (76.8)
	70–79	35,657 (17.9)	7,263 (20.4)	28,394 (79.6)
	80–89	24,701 (12.4)	4,130 (16.7)	20,571 (83.3)
	90–99	6,522 (3.3)	834 (12.8)	5,688 (87.2)
	100+	155 (0.1)	16 (10.3)	139 (89.7)
Deprivation WIMD 2019				
	Most deprived	39,360 (19.7)	10,022 (25.5)	29,338 (74.5)
	Next most deprived	42,610 (21.4)	10,196 (23.9)	32,414 (76.1)
	Middle deprivation	40,771 (20.4)	9,588 (23.5)	31,183 (76.5)
	Next least deprived	36,806 (18.4)	7,949 (21.6)	28,857 (78.4)
	Least deprived	32,355 (16.2)	7,911 (24.5)	24,444 (75.5)
	Missing	7,618 (3.8)	1,438 (18.9)	6,180 (81.1)
Identification of SCH in EHR[c]				
	Read code	8,195 (4.1)	5,071 (61.9)	3,124 (38.1)
	ICD-10 code	1,097 (0.5)	381 (34.7)	716 (65.3)
	TFT	197,833 (99.2)	46,381 (23.4)	151,452 (76.6)
	Read + ICD-10 + TFT	149 (0.1)	93 (62.4)	56 (37.6)
Length of follow-up (person-years)				
	Total	1,286,883		
	Median (IQR)	5.75 (2.65–9.65)		
Study exit				
	Study ended	142,687 (71.5)	36,047 (25.3)	106,640 (74.7)
	Censored[d]	7,715 (3.9)	1,506 (19.5)	6,209 (80.5)
	Died	49,118 (24.6)	9,551 (19.4)	39,567 (80.6)

[a] Percentage of all eligible patients (n = 199,520);

[b] Percentage of the total in the respective category–row percentage;

[c] These groups add up to more than 100% due to overlap;

[d]Due to the end of GP registration or emigration from Wales.

Abbreviations: *SCH*, Subclinical Hypothyroidism; *SD*, Standard Deviation; *WIMD*, Welsh Index of Multiple Deprivation; *ICD*, International Classification of Diseases

https://doi.org/10.1371/journal.pone.0298871.t001

Annual cumulative incidence

The annual cumulative incidence of SCH was irregular over the study period. The number of new cases was highest at a single historical point: approximately 502 cases per 100,000 people in 2012, following a marked decline in 2010 (297 cases per 100,000 people). Between 2015 and

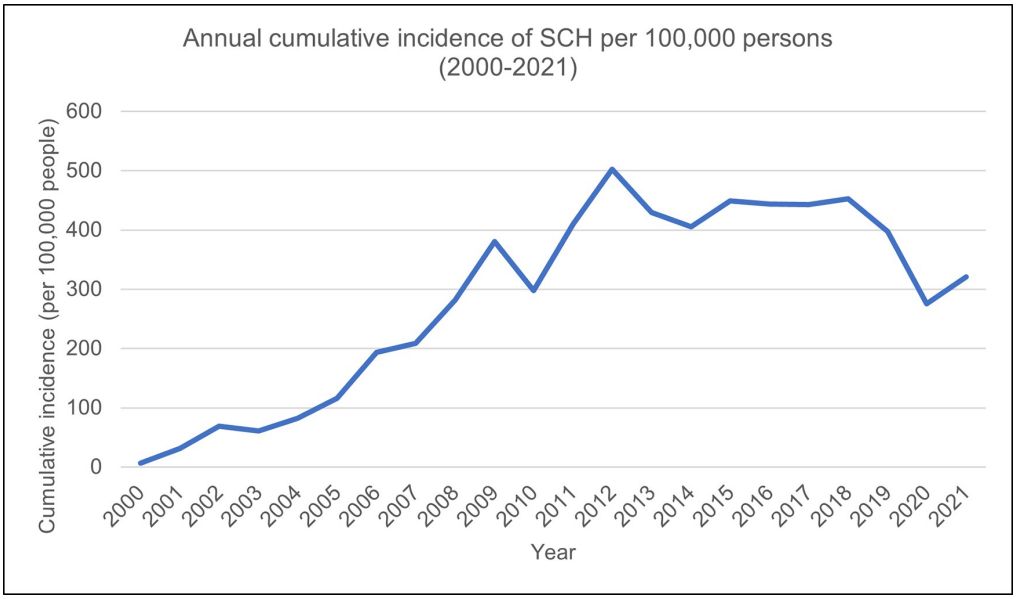


Fig 3. Trend of the annual cumulative incidence of SCH during the study period (2000–2021).

https://doi.org/10.1371/journal.pone.0298871.g003

2018, the cumulative incidence of SCH was steady but fell again until 2020 (275 cases per 100,000 people) (Fig 3).

When stratified by age and sex, the cumulative incidence charts followed a similar trajectory, though it was noted that the number of female patients was higher for both measures (S3 Appendix).

Testing and treatment patterns

The annual overall number of TFTs ordered for SCH patients in this study was 1,046 in 2000, and because more patients were identified as having SCH over the study period, the totals were highest in 2018 (n = 123,448) (Fig 4). There was a drop in 2020 (n = 88,000) compared to the previous year (n = 121,448), coinciding with the COVID-19 pandemic. The plotted graph for tests ordered for the wider GP-registered population appears similar to that for study participants but has corresponding peaks and troughs of a larger magnitude (Fig 4).

On the date of identification, 92.5% (n = 184,484) of the study population had TSH values between the upper limit of normal and 10 mIU/L. Of this group, 39,324 (21.3%) received prescriptions for levothyroxine during the study. However, the split between treated and untreated patients was reversed for patients with TSH levels higher than 10 mIU/L (Table 2).

In total, more than half of patients with severe SCH–TSH levels over 10 mIU/L–on the date of identification were in the treated group (n = 6,891; 53%). There were also more patients with elevated TPO antibodies on the recorded date of the first code or test among the treated (n = 5,428; 59.5%) than the untreated group (n = 3,689; 40.5%), as shown in Table 2.

It was also noted that 6,811 (83.1%) of patients identified using Read v2 codes had mild SCH at the time of identification, but almost two-thirds of these patients subsequently received treatment over the study period (n = 4,140; 60.8%).

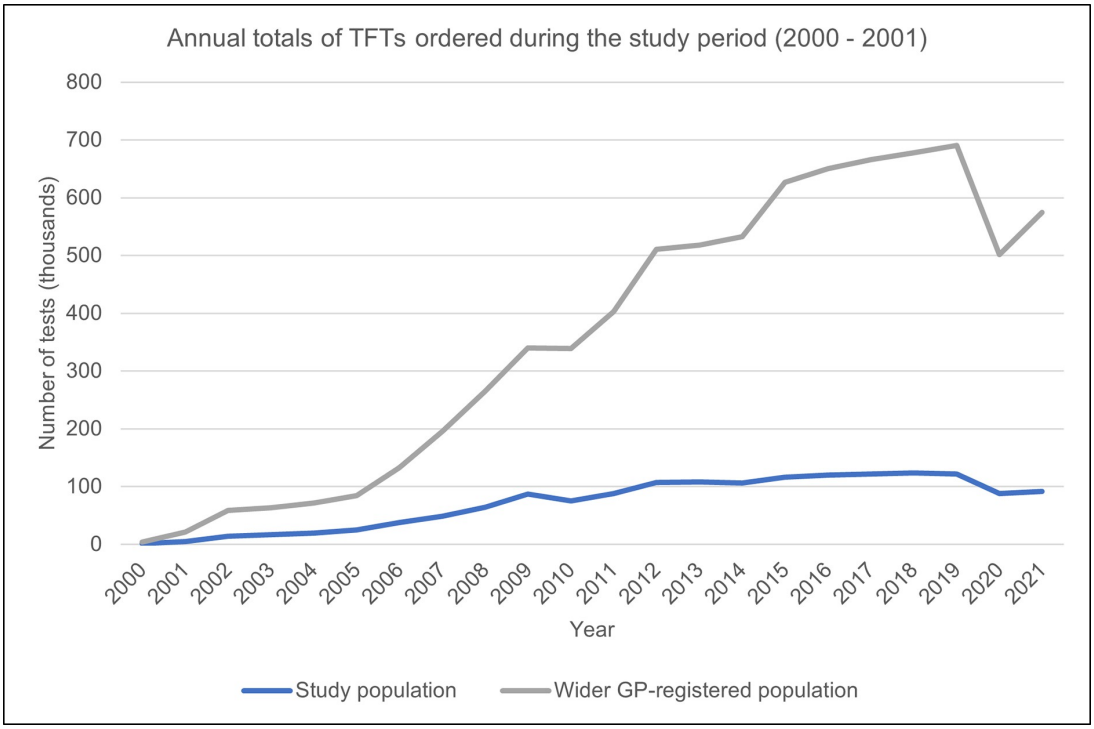


Fig 4. Annual total number of TFTs ordered for the study participants (n = 199,520) and the wider GP-registered population after applying the study eligibility criteria (n = 1,647,510).

https://doi.org/10.1371/journal.pone.0298871.g004

For the treated group, 19,716 (41.9%) patients received their first levothyroxine prescription after only one test result showing raised TSH and normal T4. (Fig 5).

Most treated patients had between 1 and 5 TFTs performed in the first 12 months (n = 31,840; 67.6%), 24 months (n = 32,801; 69.6%) and 36 months (n = 30,348; 64.4%) after treatment with levothyroxine was initially prescribed (Table 3). The proportion of untreated patients with no recorded follow-up TFTs was larger than that of treated patients over the same three-year period after SCH was identified or treated. Over this duration, more in the treated group (n = 7,399; 15.7%) had more than five monitoring tests compared to the untreated (n = 8,167; 5.4%).

The number of patients on treatment rose gradually, from 52 to 33,337 of the existing SCH cases in 2000 and 2021 respectively (Fig 6). Here, 'initiating treatment' refers to the number of patients who received their first prescription for levothyroxine in that year, as opposed to those that had already commenced treatment since the study start date ('continuing treatment'). Except at the start of the study (2000 to 2002), a smaller proportion of patients were started on treatment throughout (Fig 6 and S4 Appendix).

Under one-fifth of patients in the treated group (n = 7,794; 16.5%) received their first levothyroxine prescription within one month of their index test or code for SCH. In contrast, most treated patients (n = 29,818; 63.3%) got their first prescription more than 12 months after SCH was identified (Table 4).

Table 2. Recorded TSH and TPO levels on the date of identification, overall and stratified by treatment status.

		Study population		
		Total (%)[a]	Treated (%)[b]	Untreated (%)[b]
TSH on the date of identification				
	Upto 10 mIU/L	184,484 (92.5)	39,324 (21.3)	145,160 (78.7)
	10–20 mIU/L	10,674 (5.3)	5,669 (53.1)	5,005 (46.9)
	>20 mIU/L	2,328 (1.2)	1,222 (52.5)	1,106 (47.5)
TPO antibodies on the date of identification				
	Normal	1,588 (0.8)	309 (19.5)	1,279 (80.5)
	Elevated	9,117 (4.6)	5,428 (59.5)	3,689 (40.5)
	Missing[d]	188,815 (94.6)	41,367 (21.9)	147,448 (78.1)
Patients with SCH Read v2 codes (n = 8,195)[e]				
TSH on the date of identification				
	Upto 10 mIU/L	6,811 (83.1)	4,140 (60.8)	2,671 (39.2)
	10–20 mIU/L	452 (5.5)	363 (80.3)	89 (19.7)
	>20 mIU/L	53 (0.6)	48 (90.6)	5 (9.4)

[a] Percentage of all eligible patients (n = 199,520);

[b] Percentage of the total in the respective category–row percentage;

[c] These patients did not have TFT results recorded on the index date;

[d] These patients did not have recorded TPO antibody tests on the date of identification;

[e]Not all patients with Read codes also had recorded TFT results.

Abbreviations: *IU*, International Units; *SCH*, Subclinical hypothyroidism; *TSH*, Thyroid Stimulating Hormone; *TPO*, Thyroid Peroxidase

https://doi.org/10.1371/journal.pone.0298871.t002

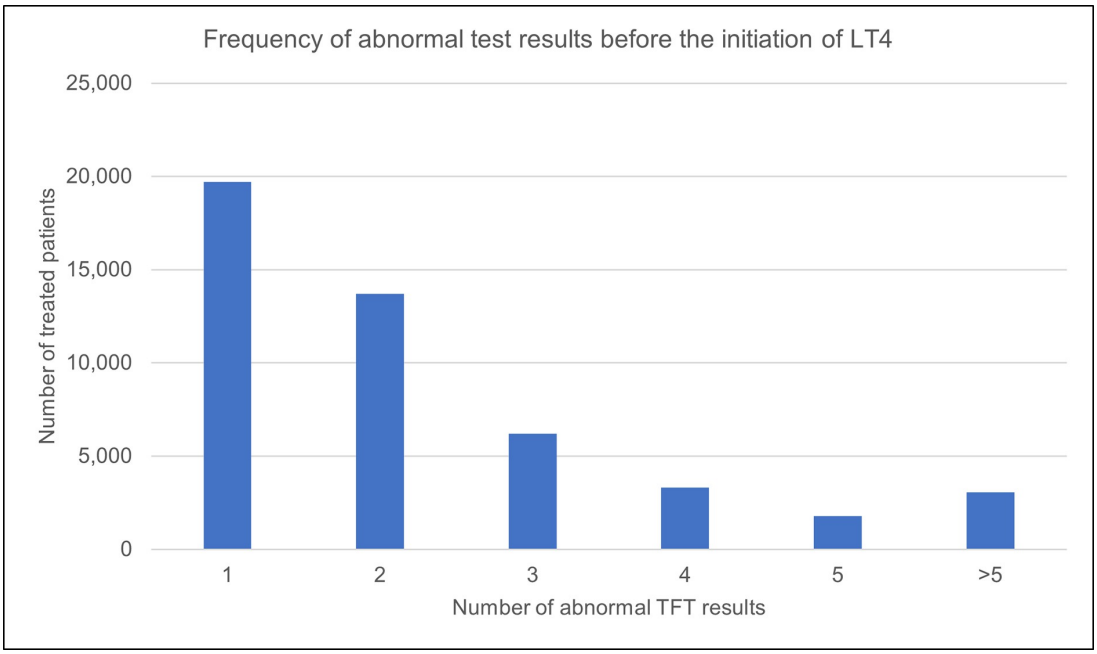


Fig 5. The frequency of abnormal tests before treatment with levothyroxine was initiated for the treated group (n = 47, 104).

https://doi.org/10.1371/journal.pone.0298871.g005

Table 3. The number of follow-up thyroid function tests performed in the immediate period (i) after the first prescription, date for treated patients and (ii) after identification, for untreated patients.

Number of monitoring TFTs performed	Treated patients (n = 47,104)	Untreated patients (n = 152,416)
	After the first prescription (%)	After identification (%)
First year		
None recorded	14,372 (30.5)	65,797 (43.2)
1–5	31,840 (67.6)	85,622 (56.2)
>5	892 (1.9)	997 (0.7)
First two years		
None recorded	10,194 (21.6)	46,650 (30.6)
1–5	32,801 (69.6)	101,807 (66.8)
>5	3,747 (8.0)	3,512 (2.3)
First three years		
None recorded	8,449 (17.9)	39,347 (25.8)
1–5	30,348 (64.4)	103,887 (68.2)
>5	7,399 (15.7)	8,167 (5.4)

https://doi.org/10.1371/journal.pone.0298871.t003

Discussion

Between 2000 and 2021, 199,520 individuals residing in Wales were identified as having SCH. The annual cumulative incidence of SCH was irregular, with a marked drop in 2010 and a prominent peak in 2012 for crude and both age- and sex-standardised estimates. In keeping with other studies on SCH [11, 39], it was found that more females than males were identified

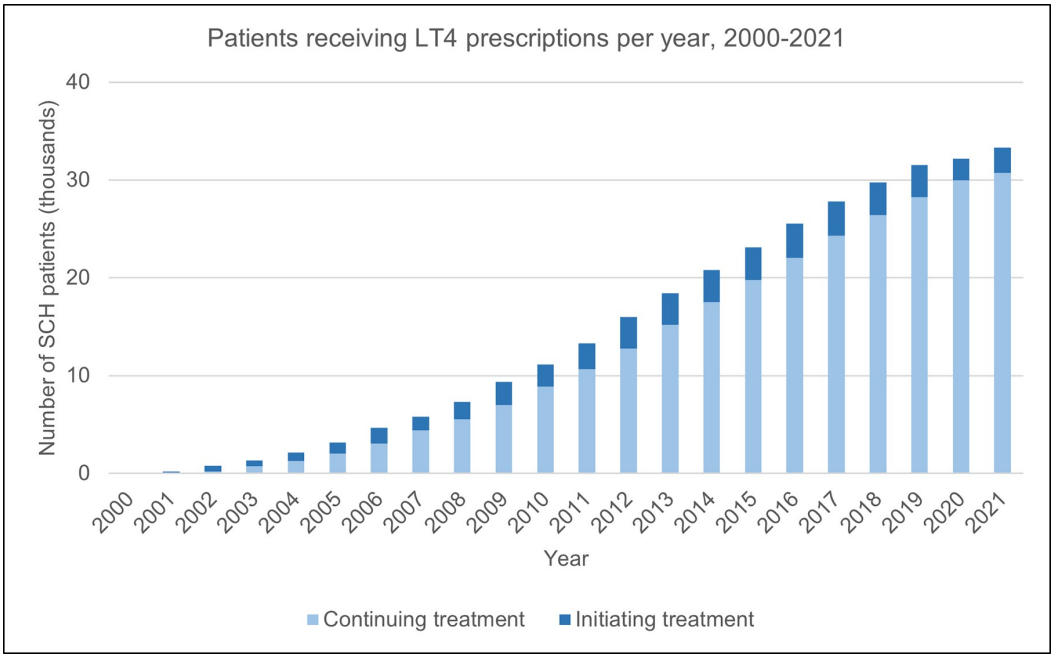


Fig 6. The number of SCH patients with recorded prescriptions for levothyroxine during the study period.

https://doi.org/10.1371/journal.pone.0298871.g006

Table 4. The characteristics of the treatment given to SCH patients in Wales, 2000–2021.

		Treated patients (n = 47,104)
Time from the date of identification to the first LT4 prescription		
	≤1 month	7,794 (16.5)
	1–6 months	6,470 (13.7)
	6–12 months	3,022 (6.4)
	> 1 year	29,818 (63.3)

Abbreviations: *LT4*, Levothyroxine

https://doi.org/10.1371/journal.pone.0298871.t004

with SCH. Less than one-third of the study population received treatment during the study period. Most patients identified using diagnostic codes, specifically Read v2, received levothyroxine over the study period, though an equally large proportion of these patients also had mild SCH on their first test date.

In contrast, 6.5% of the study population had TSH levels higher than 10 mIU/L on their index test results. Levothyroxine was prescribed for 42% of patients in the treated group after a single abnormal test indicative of SCH. The frequency of TFTs for treated patients after their first prescription was higher than tests performed for untreated patients after identifying SCH.

Annual cumulative incidence

The reason for the peak in annual cumulative incidence in 2012 was not immediately apparent. However, most of the patients in this study were identified using TFT results rather than diagnostic codes. A potential reason for the irregular pattern of annual incident cases, therefore, might be altered clinical decision-making as a result of new guidance; for example, the Royal College of Physicians released a statement on the diagnosis and management of primary hypothyroidism in 2008 [40] and updated the guidance in 2011 [41]. The intervening period corresponds to the marked drop in incident cases of SCH in the study population around 2010, but this conclusion cannot be reached based on a single study.

Notably, when standardised to the mid-2011 population of Wales, estimates of the annual cumulative incidence of SCH were broadly similar across the age and sex categories. However, the line representing patients aged 90 and over was flattened compared to the crude annual cumulative incidence plot, suggesting that the proportion that was 90 years or older in the standard population was smaller than among the study population.

Testing and treatment

Overtesting for thyroid function has been widely reported in the literature [28, 29, 42–44]. In contrast, our findings show that in the years following the identification of SCH, most of the untreated patients had fewer tests, which aligns with current guidance to perform repeat TFTs annually or biennially for untreated SCH, depending on the presence of features of underlying thyroid disease [25]. Treated patients, who would typically require frequent monitoring tests–every three months until TSH levels normalise and then annually–were found to have had more TFTs in comparison. Annual monitoring would possibly explain why 64% of treated patients had between one and five tests in the first three years after treatment was started.

Data on TPO antibodies on the date of identification were available for only a small fraction of patients in the cohort. However, this corresponds to the NICE recommendation to consider antibody testing for elevated TSH levels and avoid repeating these specific tests [25]. TPO antibody assays predating these NICE guidelines would also likely have been infrequent.

The plotted rise in the number of TFTs ordered for study participants can be explained by the increasing number of identified SCH cases over the study period. On the other hand, the marked drop observed around 2020 may be attributed to the COVID-19 pandemic, during which mandatory lockdowns were implemented. It is, however, impossible to completely rule out overtesting in this study because inappropriate TFT ordering accompanied by a higher number of patients receiving tests would, unsurprisingly, present a false picture of a steady average among patients with SCH.

Over 90% of the study population had mild SCH–TSH levels between the upper limit of normal and the 10mIU/L threshold–on the date they were identified. This number is higher than that reported from the Colorado Thyroid Disease Prevalence study, 74% [11]. However, the participants in the Colorado study were the attendees of a statewide health fair, representing 25,862 individuals. Unlike this longitudinal study based on SCH cases over 20 years, the fair facilitated a cross-sectional survey. As such, a more extended series of surveys might have approximated our findings more closely.

Patients with a recorded diagnosis were more likely to receive treatment than those identified through test results. Of the latter, less than a third were in the treated group, suggesting that GPs may have been more likely to disregard test results indicative of SCH if, for instance, the TFTs were not directly relevant to their plans for the clinical management of presenting symptoms. In such a scenario, the SCH diagnosis would not be recorded, decreasing the probability of the patient receiving a prescription for levothyroxine. Crucially, the study findings also indicated that the presence of a diagnostic code in the EHR was not an indicator of the severity of SCH–as determined using the 10 mIU/L threshold for TSH. Their index test results showed that most patients with clinically recorded diagnoses had mild SCH in the first instance.

More of the patients who had severe SCH on the index date received treatment during the study, as would be expected. Several treatment guidelines around the world similarly recommend the initiation of levothyroxine for cases with TSH levels above 10mIU/L, including those from the American Thyroid Association [45, 46], European Thyroid Association [47] and Brazilian Society of Endocrinology and Metabolism [48]. A crucial caveat to this guidance is that patient age must be considered, given the physiological increases in TSH levels with age [20, 23]. It is frequently stated that a higher treatment threshold should be applied for older patients, particularly for mild SCH [18, 49, 50]. Even so, a multinational survey on treatment practices for SCH by Razvi et al. [51] found wide variability in the implementation of such recommendations. Potential reasons for the failure to treat cases of severe SCH, as observed for 3.1% of our study population, include patient age, transient increases in TSH or measurement errors [2], which would be lower on subsequent tests.

Levothyroxine is currently the third most frequently prescribed drug by GPs in Wales [52–54]. Previous studies have also described 'overtreatment'–the tendency for clinicians to initiate treatment even when it is not necessarily required and would not benefit the patient [30, 55]. The finding that 42% of all the patients in the treated group had levothyroxine initiated based on a single abnormal test result indicative of SCH may be related to overtreatment. This proportion was notably higher than the number of patients with severe SCH and those who tested positive for TPO antibodies–according to current guidelines, these two groups would have been deemed eligible for immediate treatment. Furthermore, the number of patients who had severe SCH on the date of identification was lower than those who received their first prescription within one month of that date. All these findings point to the potential overuse of levothyroxine over the study period.

Strengths and limitations

The strengths of this study included the use of multiple SAIL datasets to identify patients identified as having SCH and to assess their eligibility for this study; the scale and coverage of the data used, which is almost the whole population of Wales spanning at least two decades, as well as the characterisation of different aspects of SCH and SCH patients to provide a comprehensive description of the disorder and how it was managed clinically.

This study is subject to some limitations. Foremost are the inherent challenges of identifying SCH in EHR: missing data, coding errors and inconsistencies in patient records influenced the inclusion of patients in the study. It was observed that the test results dataset, in particular, was plagued by these issues due to the variety of ways in which test names, codes and values were recorded. Even so, this problem was mitigated by including the primary care and inpatient datasets to identify patients who had been clinically diagnosed with the disorder. SAIL Databank is also recognised as a high-quality primary care data source with a high level of coverage in Wales (84%) [56]. Also, it was not possible to identify the ethnicity of eligible patients as this information was not available in the provisioned datasets.

Related to these points, most of the study population was identified in the EHR using their TFT results. However, these are not error-proof; it has been reported previously that due to the use of population reference ranges, an individual may have a physiologically abnormal thyroid function test (TFT) result that matches what is otherwise considered normal range [1, 57].

It is also essential to note that for the evaluation of treatment, only prescribing data–not dispensing data–were available. It was, therefore, impossible to ascertain whether the levothyroxine prescriptions given were filled (adherence) and the medication taken by the patient as instructed (compliance). This affected the interpretation of treatment duration, as it would otherwise have been possible to explore LT4 use by accounting for the amount of medication given and, possibly, to distinguish between treated patients by dosages and drug formulations. However, it can reasonably be expected that most patients who received prescriptions took their medication; hence, we used prescriptions as a proxy for treatment status.

A key challenge in estimating the true prevalence of SCH is the lack of agreement concerning TSH reference ranges [9]. Diagnosis relies entirely on lab results; therefore, variations in the upper limit of normal for TSH directly influence patient numbers. However, separate constraints arise from using EHR only to measure prevalence. Chief among these is the likelihood of underestimating the number of existing cases because of the complex interplay of factors that affect the decision to seek clinical care [58]. This selection bias arises because patients with more severe symptoms are more likely to visit their GP and have a recorded visit in the EHR [59, 60]. Another factor is that the restriction of study start and stop dates can mimic a closed cohort in which the 'existing patient' count starts at zero. We did not use the study population for prevalence estimates for these reasons.

Finally, the most recently reported SAIL coverage of GP practices in Wales is approximately 84% [56]; therefore, it cannot be assumed that the denominators used to estimate the annual cumulative incidence are directly equal to the respective actual population counts. This challenge was mitigated by performing standardisation of the estimated cumulative incidence, though the plotted graphs were essentially unchanged.

Conclusion

This descriptive study on SCH in Wales shows an uneven rise in the overall number of patients, TFT and levothyroxine use between 2000 and 2021, with the highest annual cumulative incidence in 2012 at 502 cases per 100,000 people.

Compared to patients who only had test results indicative of SCH, those with clinically recorded diagnoses were less likely to meet the 10mIU/L TSH cutoff for severe SCH but were also more likely to be offered treatment. In contrast, patients with mild SCH on their index test were less likely to receive levothyroxine. The clinical management of SCH was inconsistent with the recommendation to consider treatment only if a repeat test reveals TSH levels higher than 10mIU/L, considering that over a third of treated cases had only one prior abnormal test result. However, per the current NICE annual monitoring guidance, TFTs were ordered more frequently for treated than untreated patients.

Our findings show that more robust guidelines are needed to ensure the appropriate clinical management of patients with SCH. Given the potential for a continued rise in patient numbers and conflicting evidence on the impact of SCH on patients' long-term health, more research is needed to inform strategies to improve the use of TFTs and levothyroxine for managing this thyroid disorder.

Supporting information

S1 Appendix. Description of SAIL datasets.
(DOCX)

S2 Appendix. Case definition (SCH patients).
(DOCX)

S3 Appendix. Age- and sex-stratified annual cumulative incidence according to the mid-2011 Welsh census data.
(DOCX)

S4 Appendix. Levothyroxine use during the study period.
(DOCX)

Author Contributions

Conceptualization: Brenda S. Bauer, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan.

Data curation: Brenda S. Bauer.

Formal analysis: Brenda S. Bauer.

Funding acquisition: Colin McCowan.

Investigation: Brenda S. Bauer.

Methodology: Brenda S. Bauer.

Project administration: Brenda S. Bauer.

Resources: Colin McCowan.

Supervision: Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Colin McCowan.

Validation: Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan.

Visualization: Brenda S. Bauer.

Writing – original draft: Brenda S. Bauer.

Writing – review & editing: Brenda S. Bauer, Amaya Azcoaga-Lorenzo, Utkarsh Agrawal, Adeniyi Francis Fagbamigbe, Colin McCowan.

References

1. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocrine reviews. 2008; 29:76–131. https://doi.org/10.1210/er.2006-0043 PMID: 17991805
2. Cooper DS, Biondi B. Subclinical thyroid disease. The Lancet. 2012; 379:1142–54. https://doi.org/10.1016/S0140-6736(11)60276-6 PMID: 22273398
3. Biondi B. The normal TSH reference range: what has changed in the last decade?: Oxford University Press; 2013. p. 3584–7.
4. Dayan C, Panicker V. Interpretation of thyroid function tests and their relationship to iodine nutrition: changes in TSH, free T4, and free T3 resulting from iodine deficiency and iodine excess. in: Preedy V R, et al(Eds), Comprehensive handbook of iodine: nutritional, biochemical, pathological and therapeutic aspects, Academic Press, Boston, USA. 2009:47–54.
5. Kratzsch J, Fiedler GM, Leichtle A, Brügel M, Buchbinder S, Otto L, et al. New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry criteria and regular ultrasonography of the thyroid. Clinical chemistry. 2005; 51:1480–6. https://doi.org/10.1373/clinchem.2004.047399 PMID: 15961550
6. Meyerovitch J, Rotman-Pikielny P, Sherf M, Battat E, Levy Y, Surks MI. Serum thyrotropin measurements in the community: five-year follow-up in a large network of primary care physicians. Archives of Internal Medicine. 2007; 167:1533–8. https://doi.org/10.1001/archinte.167.14.1533 PMID: 17646608
7. Surks MI. TSH reference limits: new concepts and implications for diagnosis of subclinical hypothyroidism. Endocrine Practice. 2013; 19:1066. https://doi.org/10.4158/EP13246.CO PMID: 24014005
8. 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; 84:799–808. https://doi.org/10.1111/cen.12824 PMID: 26010808
9. Hennessey J, Espaillat R. Subclinical hypothyroidism: a historical view and shifting prevalence. International Journal of Clinical Practice. 2015; 69:771–82. https://doi.org/10.1111/ijcp.12619 PMID: 25846327
10. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. Jama. 2004; 291:228–38. https://doi.org/10.1001/jama.291.2.228 PMID: 14722150
11. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. Archives of internal medicine. 2000; 160:526–34. https://doi.org/10.1001/archinte.160.4.526 PMID: 10695693
12. Surks MI, Hollowell JG. Age-specific distribution of serum thyrotropin and antithyroid antibodies in the US population: implications for the prevalence of subclinical hypothyroidism. The Journal of Clinical Endocrinology & Metabolism. 2007; 92:4575–82. https://doi.org/10.1210/jc.2007-1499 PMID: 17911171
13. Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, et al. Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). The Journal of Clinical Endocrinology & Metabolism. 2002; 87:489–99.
14. Díez JJ, Iglesias P. Spontaneous Subclinical Hypothyroidism in Patients Older than 55 Years: An Analysis of Natural Course and Risk Factors for the Development of Overt Thyroid Failure. The Journal of Clinical Endocrinology & Metabolism. 2004; 89:4890–7. https://doi.org/10.1210/jc.2003-032061 PMID: 15472181
15. Díez JJ, Iglesias P, Burman KD. Spontaneous normalization of thyrotropin concentrations in patients with subclinical hypothyroidism. The Journal of Clinical Endocrinology & Metabolism. 2005; 90:4124–7. https://doi.org/10.1210/jc.2005-0375 PMID: 15811925
16. Biondi B. Natural history, diagnosis and management of subclinical thyroid dysfunction. Best practice & research Clinical endocrinology & metabolism. 2012; 26:431–46. https://doi.org/10.1016/j.beem.2011.12.004 PMID: 22863386
17. Calissendorff J, Falhammar H. To treat or not to treat subclinical hypothyroidism, what is the evidence? Medicina. 2020; 56:40. https://doi.org/10.3390/medicina56010040 PMID: 31963883
18. Chrysant SG. The current debate over treatment of subclinical hypothyroidism to prevent cardiovascular complications. International journal of clinical practice. 2020; 74:e13499. https://doi.org/10.1111/ijcp.13499 PMID: 32159256

19. Chu JW, Crapo LM. The treatment of subclinical hypothyroidism is seldom necessary. The Journal of Clinical Endocrinology & Metabolism. 2001; 86:4591–9. https://doi.org/10.1210/jcem.86.10.7961 PMID: 11600508

20. Hennessey JV, Espaillat R. Diagnosis and management of subclinical hypothyroidism in elderly adults: a review of the literature. Journal of the American Geriatrics Society. 2015; 63:1663–73. https://doi.org/10.1111/jgs.13532 PMID: 26200184

21. Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RG, Mooijaart SP, et al. Thyroid hormone therapy for older adults with subclinical hypothyroidism. New England Journal of Medicine. 2017; 376:2534–44. https://doi.org/10.1056/NEJMoa1603825 PMID: 28402245

22. Waring AC, Harrison S, Samuels MH, Ensrud KE, LeBlanc ES, Hoffman AR, et al. Thyroid function and mortality in older men: a prospective study. The Journal of Clinical Endocrinology & Metabolism. 2012; 97:862–70. https://doi.org/10.1210/jc.2011-2684 PMID: 22238396

23. Mooijaart SP, Du Puy RS, Stott DJ, Kearney PM, Rodondi N, Westendorp RG, et al. Association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism. Jama. 2019; 322:1977–86. https://doi.org/10.1001/jama.2019.17274 PMID: 31664429

24. Winther KH, Cramon P, Watt T, Bjorner JB, Ekholm O, Feldt-Rasmussen U, et al. Disease-specific as well as generic quality of life is widely impacted in autoimmune hypothyroidism and improves during the first six months of levothyroxine therapy. PloS one. 2016; 11:e0156925. https://doi.org/10.1371/journal.pone.0156925 PMID: 27257805

25. UK National Guideline Centre. Thyroid disease: assessment and management 2019 [updated 20 November 2019]. https://www.nice.org.uk/guidance/ng145.

26. Gencer B, Rodondi N. Evidence and controversies regarding the screening for subclinical hypothyroidism in patients with cardiovascular disease. Journal of thoracic disease. 2016; 8:E446. https://doi.org/10.21037/jtd.2016.04.20 PMID: 27294253

27. Cappola AR, Cooper DS. Screening and treating subclinical thyroid disease: getting past the impasse. Annals of internal medicine. 2015; 162:664–5. https://doi.org/10.7326/M15-0640 PMID: 25799007

28. Vaidya B, Ukoumunne OC, Shuttleworth J, Bromley A, Lewis A, Hyde C, et al. Variability in thyroid function test requests across general practices in south-west England. Qual Prim Care. 2013; 21:143–8. PMID: 23968263

29. Werhun A, Hamilton W. Thyroid function testing in primary care: overused and under-evidenced? A study examining which clinical features correspond to an abnormal thyroid function result. Family practice. 2015; 32:187–91. https://doi.org/10.1093/fampra/cmv010 PMID: 25782692

30. Somwaru LL, Arnold AM, Joshi N, Fried LP, Cappola AR. High frequency of and factors associated with thyroid hormone over-replacement and under-replacement in men and women aged 65 and over. The Journal of Clinical Endocrinology & Metabolism. 2009; 94:1342–5. https://doi.org/10.1210/jc.2008-1696 PMID: 19126628

31. Akbari A, Bedston S, Abbasizanjani H, Davies G, Fry R, Lowthian E, et al. Developing a population-scale harmonised ethnicity-spine in Wales. International Journal of Population Data Science. 2022; 7.

32. Ford DV, Jones KH, Verplancke J-P, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC health services research. 2009; 9:1–12.

33. Jones KH, Ford DV, Thompson S, Lyons R. A profile of the SAIL databank on the UK secure research platform. International Journal of Population Data Science. 2019; 4. https://doi.org/10.23889/ijpds.v4i2.1134 PMID: 34095541

34. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford DV, et al. The SAIL databank: linking multiple health and social care datasets. BMC medical informatics and decision making. 2009; 9:1–8.

35. Cyngor Gwynedd. Welsh Index of Multiple Deprivation [cited 2023 05 May]. https://www.gwynedd.llyw.cymru/en/Council/Key-statistics-and-data/Welsh-Index-of-Multiple-Deprivation.aspx.

36. Jones H. Where are the most deprived areas in Wales?: Senedd Research, National Assembly for Wales; 2019 [updated 27 May 2021]. https://research.senedd.wales/research-articles/where-are-the-most-deprived-areas-in-wales/.

37. Jones KH, Ford DV, Jones C, Dsilva R, Thompson S, Brooks CJ, et al. A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: a privacy-protecting remote access system for health-related research and evaluation. Journal of biomedical informatics. 2014; 50:196–204. https://doi.org/10.1016/j.jbi.2014.01.003 PMID: 24440148

38. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS medicine. 2015; 12:e1001885. https://doi.org/10.1371/journal.pmed.1001885 PMID: 26440803

39. Vanderpump MP, Tunbridge WMG. Epidemiology and prevention of clinical and subclinical hypothyroidism. Thyroid. 2002; 12:839–47. https://doi.org/10.1089/105072502761016458 PMID: 12487765

40. Allahabadia A, Razvi S, Abraham P, Franklyn J. Diagnosis and treatment of primary hypothyroidism. BMJ: British Medical Journal (Online). 2009; 338. https://doi.org/10.1136/bmj.b725 PMID: 19325179

41. Royal College of Physicians. The diagnosis and management of primary hypothyroidism—Revised 2011. 2011.

42. Dalal S, Bhesania S, Silber S, Mehta P. Use of electronic clinical decision support and hard stops to decrease unnecessary thyroid function testing. BMJ Open Quality. 2017; 6:u223041. w8346. https://doi.org/10.1136/bmjquality.u223041.w8346 PMID: 28469901

43. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function tests in patients with stable untreated subclinical hypothyroidism. Thyroid. 2008; 18:303–8. https://doi.org/10.1089/thy.2007.0241 PMID: 18303960

44. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function in subclinical hypothyroidism: importance of clinical follow-up and therapy. European Journal of Endocrinology. 2011; 164:317. https://doi.org/10.1530/EJE-10-1021 PMID: 21208994

45. Jonklaas J, Bianco AC, Bauer AJ, Burman KD, Cappola AR, Celi FS, et al. Guidelines for the treatment of hypothyroidism: prepared by the american thyroid association task force on thyroid hormone replacement. thyroid. 2014; 24:1670–751. https://doi.org/10.1089/thy.2014.0028 PMID: 25266247

46. Garber JR, Cobin RH, Gharib H, Hennessey JV, Klein I, Mechanick JI, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. Thyroid. 2012; 22:1200–35. https://doi.org/10.1089/thy.2012.0205 PMID: 22954017

47. Pearce SH, Brabant G, Duntas LH, Monzani F, Peeters RP, Razvi S, et al. 2013 ETA guideline: management of subclinical hypothyroidism. European thyroid journal. 2013; 2:215–28. https://doi.org/10.1159/000356507 PMID: 24783053

48. Sgarbi JA, Teixeira PF, Maciel LM, Mazeto GM, Vaisman M, Montenegro Junior RM, et al. The Brazilian consensus for the clinical approach and treatment of subclinical hypothyroidism in adults: recommendations of the thyroid Department of the Brazilian Society of Endocrinology and Metabolism. Arquivos Brasileiros de Endocrinologia & Metabologia. 2013; 57:166–83. https://doi.org/10.1590/s0004-27302013000300003 PMID: 23681263

49. Leng O, Razvi S. Hypothyroidism in the older population. Thyroid research. 2019; 12:1–10.

50. Biondi B, Cappola AR. Subclinical hypothyroidism in older individuals. The Lancet Diabetes & Endocrinology. 2022; 10:129–41. https://doi.org/10.1016/S2213-8587(21)00285-0 PMID: 34953533

51. Razvi S, Arnott B, Teare D, Hiu S, O'Brien N, Pearce SH. Multinational survey of treatment practices of clinicians managing subclinical hypothyroidism in older people in 2019. European thyroid journal. 2021; 10:330–8. https://doi.org/10.1159/000509228 PMID: 34395305

52. Statistics for Wales. Prescriptions in Wales: April 2019 to March 2020. 2020 23 September 2020. Report No.

53. Statistics for Wales. Prescriptions in Wales: April 2021 to March 2022. 2022 7 September 2022. Report No.

54. Statistics for Wales. Prescriptions in Wales: April 2020 to March 2021. 2021 8 September 2021. Report No.

55. El Kawkgi OM, Brito JP. Screening for thyroid dysfunction: prevention of overdiagnosis and overtreatment. CMAJ. 2019; 191:E1260–E1. https://doi.org/10.1503/cmaj.191437 PMID: 31944183

56. SAIL Databank. Welsh Longitudinal General Practice Dataset (WLGP)—Welsh Primary Care: Health Data Research Innovation Gateway; [updated 05 April 2023]. https://web.www.healthdatagateway.org/dataset/33fc3ffd-aa4c-4a16-a32f-0c900aaea3d2.

57. Andersen S, Pedersen KM, Bruun NH, Laurberg P. Narrow individual variations in serum T4 and T3 in normal subjects: a clue to the understanding of subclinical thyroid disease. The Journal of Clinical Endocrinology & Metabolism. 2002; 87:1068–72.

58. Bagley SC, Altman RB. Computing disease incidence, prevalence and comorbidity from electronic medical records. Journal of biomedical informatics. 2016; 63:108–11. https://doi.org/10.1016/j.jbi.2016.08.005 PMID: 27498067

59. Rusanov A, Weiskopf NG, Wang S, Weng C. Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research. BMC medical informatics and decision making. 2014; 14:1–9.

60. Edmondson ME, Reimer AP. Challenges frequently encountered in the secondary use of electronic medical record data for research. CIN: Computers, Informatics, Nursing. 2020; 38:338–48. https://doi.org/10.1097/CIN.0000000000000609 PMID: 32149742

3.4 Conclusion

In this chapter, I used routinely collected data from SAIL Databank to describe the patterns of incidence, testing and treatment of SCH in Wales between the start of the year 2000 and the end of 2021. The frequency of TFTs, and therefore the incidence of SCH, consistently rose until the COVID-19 pandemic. Testing and prescriptions for levothyroxine were inconsistent with the general guidance around confirmatory tests and TSH thresholds. This study provided a nationwide profile of SCH, demonstrating the impact of inconclusive guidance on clinical practice. Chapter 4 uses observational data to investigate the causal effects of treating SCH on patient health in the long term. The results of the trial emulation will contribute to the evidence base by applying an analytical framework that aims to limit bias.

4

Target trial emulation

5

Discussion

6

References

1. Boelaert K, Franklyn J. Thyroid hormone in health and disease. Journal of Endocrinology. 2005;187:1–15.

2. Bliss D. Thyroid gland. 2001;1127x1072.

3. Dev N, Sankar J, Vinay MV. Functions of thyroid hormones. In: Imam SK, Ahmad SI, editors. Thyroid disorders: Basic science and clinical practice. Cham: Springer International Publishing; 2016. p. 11–25.

4. Felsenfeld AJ, Levine BS. Calcitonin, the forgotten hormone: Does it deserve to be forgotten? Clinical kidney journal. 2015;8:180–7.

5. Ross DS. Treating hypothyroidism is not always easy: When to treat subclinical hypothyroidism, TSH goals in the elderly, and alternatives to levothyroxine monotherapy. Journal of Internal Medicine. 2022;291:128–40.

6. Keestra S, Högqvist Tabor V, Alvergne A. Reinterpreting patterns of variation in human thyroid function: An evolutionary ecology perspective. Evolution, medicine, and public health. 2021;9:93–112.

7. Fekete C, Lechan RM. Central regulation of hypothalamic-pituitary-thyroid axis under physiological and pathophysiological conditions. Endocrine reviews. 2014;35:159–94.

8. Duntas LH. Environmental endocrinology and the hypothalamus-pituitary-thyroid axis. In: Pivonello R, Diamanti-Kandarakis E, editors. Environmental endocrinology and endocrine disruptors: Endocrine and endocrine-targeted actions and related human diseases. Cham: Springer International Publishing; 2020. p. 1–19.

9. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, et al. Subclinical thyroid disease: Scientific review and guidelines for diagnosis and management. Jama. 2004;291:228–38.

10. Razvi S, Weaver JU, Butler TJ, Pearce SH. Levothyroxine treatment of subclinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. Archives of internal medicine. 2012;172:811–7.

11. Klaver EI, Loon HC van, Stienstra R, Links TP, Keers JC, Kema IP, et al. Thyroid hormone status and health-related quality of life in the LifeLines cohort study. Thyroid. 2013;23:1066–73.

12. Reyes Domingo F, Avey MT, Doull M. Screening for thyroid dysfunction and treatment of screen-detected thyroid dysfunction in asymptomatic, community-dwelling adults: A systematic review. Systematic Reviews. 2019;8:260.

13. Laurberg P, Andersen S, Carlé A, Karmisholt J, Knudsen N, Pedersen IB. The TSH upper reference limit: Where are we at? Nature Reviews Endocrinology. 2011;7:232–9.

14. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. Endocrine reviews. 2008;29:76–131.

15. Cooper DS, Biondi B. Subclinical thyroid disease. The Lancet. 2012;379:1142–54.

16. Peeters RP. Subclinical hypothyroidism. New England Journal of Medicine. 2017;376:2556–65.

17. Meyerovitch J, Rotman-Pikielny P, Sherf M, Battat E, Levy Y, Surks MI. Serum thyrotropin measurements in the community: Five-year follow-up in a large network of primary care physicians. Archives of Internal Medicine. 2007;167:1533–8.

18. Hennessey J, Espaillat R. Subclinical hypothyroidism: A historical view and shifting prevalence. International Journal of Clinical Practice. 2015;69:771–82.

19. Bensenor IM, Olmos RD, Lotufo PA. Hypothyroidism in the elderly: Diagnosis and management. Clinical interventions in aging. 2012;7:97–111.

20. Mendes D, Alves C, Silverio N, Marques FB. Prevalence of undiagnosed hypothyroidism in europe: A systematic review and meta-analysis. European Thyroid Journal. 2019;8:130–43.

21. Garmendia Madariaga A, Santos Palacios S, Guillén-Grima F, Galofré JC. The incidence and prevalence of thyroid dysfunction in europe: A meta-analysis. The Journal of Clinical Endocrinology & Metabolism. 2014;99:923–31.

22. Razvi S, Bhana S, Mrabeti S. Challenges in interpreting thyroid stimulating hormone results in the diagnosis of thyroid dysfunction. Journal of thyroid research. 2019;2019:4106816.

23. Bell RJ, Rivera-Woll L, Davison SL, Topliss DJ, Donath S, Davis SR. Well-being, health-related quality of life and cardiovascular disease risk profile in women with subclinical thyroid disease - a community-based study. Clinical endocrinology. 2007;66:548–56.

24. Yoo WS, Chung HK. Subclinical hypothyroidism: Prevalence, health impact, and treatment landscape. Endocrinology and Metabolism. 2021;36:500–13.

25. Tunbridge W, Evered D, Hall R, Appleton D, Brewis M, Clark F, et al. The spectrum of thyroid disease in a community: The whickham survey. Clinical endocrinology. 1977;7:481–93.

26. Drinka P, Nolten W. Prevalence of previously undiagnosed hypothyroidism in residents of a midwestern nursing home. Southern medical journal. 1990;83:1259–61, 1265.

27. Konno N, Yuri K, Taguchi H, Miura K, Taguchi S, Hagiwara K, et al. Screening for thyroid diseases in an iodine sufficient area with sensitive thyrotrophin assays, and

serum thyroid autoantibody and urinary iodide determinations. Clinical endocrinology. 1993;38:273–81.

28. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The colorado thyroid disease prevalence study. Archives of internal medicine. 2000;160:526–34.

29. Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, et al. Serum TSH, T4, and thyroid antibodies in the united states population (1988 to 1994): National health and nutrition examination survey (NHANES III). The Journal of Clinical Endocrinology & Metabolism. 2002;87:489–99.

30. Hoogendoorn EH, Hermus AR, De VegT F, Ross HA, Verbeek AL, Kiemeney LA, et al. Thyroid function and prevalence of anti-thyroperoxidase antibodies in a population with borderline sufficient iodine intake: Influences of age and sex. Clinical chemistry. 2006;52:104–11.

31. Kim WG, Kim WB, Woo G, Kim H, Cho Y, Kim TY, et al. Thyroid stimulating hormone reference range and prevalence of thyroid dysfunction in the korean population: Korea national health and nutrition examination survey 2013 to 2015. Endocrinology and Metabolism. 2017;32:106–14.

32. Pearce SH, Brabant G, Duntas LH, Monzani F, Peeters RP, Razvi S, et al. 2013 ETA guideline: Management of subclinical hypothyroidism. European thyroid journal. 2013;2:215–28.

33. Livingston EH. Subclinical hypothyroidism. JAMA. 2019;322:180.

34. Chiovato L, Magri F, Carlé A. Hypothyroidism in context: Where we've been and where we're going. Advances in therapy. 2019;36:47–58.

35. Kim TH, Kim KW, Ahn HY, Choi HS, Won H, Choi Y, et al. Effect of seasonal changes on the transition between subclinical hypothyroid and euthyroid status. The Journal of Clinical Endocrinology & Metabolism. 2013;98:3420–9.

36. Díez JJ, Iglesias P, Burman KD. Spontaneous normalization of thyrotropin concentrations in patients with subclinical hypothyroidism. The Journal of Clinical Endocrinology & Metabolism. 2005;90:4124–7.

37. Javed Z, Sathyapalan T. Levothyroxine treatment of mild subclinical hypothyroidism: A review of potential risks and benefits. Therapeutic advances in endocrinology and metabolism. 2016;7:12–23.

38. Vaidya B, Ukoumunne OC, Shuttleworth J, Bromley A, Lewis A, Hyde C, et al. Variability in thyroid function test requests across general practices in south-west england. 2013;21:143–8.

39. Bashkin A, Yaakobi E, Nodelman M, Ronen O. Is routine measurement of TSH in hospitalized patients necessary? Endocrine connections. 2018;7:567–72.

40. Dayan CM. Interpretation of thyroid function tests. The Lancet. 2001;357:619–24.

41. Carvalho GA de, Perez CLS, Ward LS. The clinical use of thyroid function tests. Arquivos Brasileiros de Endocrinologia & Metabologia. 2013;57:193–204.

42. Sheehan MT. Biochemical testing of the thyroid: TSH is the best and, oftentimes, only test needed–a review for primary care. Clinical medicine & research. 2016;14:83–92.

43. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function tests in patients with stable untreated subclinical hypothyroidism. Thyroid. 2008;18:303–8.

44. Abbey EJ, McGready J, Sokoll LJ, Simonsick EM, Mammen JS. Free thyroxine distinguishes subclinical hypothyroidism from other aging-related changes in those with isolated elevated thyrotropin. Frontiers in Endocrinology. 2022;13:858332.

45. Centre UNG. Thyroid disease: Assessment and management. 2019.

46. Razvi S, Jabbar A, Addison C, Vernazza J, Syed A, Soran H, et al. Variation in the reference range limits of thyroid function tests and association with the prevalence of levothyroxine treatment. European Journal of Endocrinology. 2023;188:K5–9.

47. Surks MI. TSH reference limits: New concepts and implications for diagnosis of subclinical hypothyroidism. Endocrine Practice. 2013;19:1066–9.

48. Surks MI, Hollowell JG. Age-specific distribution of serum thyrotropin and antithyroid antibodies in the US population: Implications for the prevalence of subclinical hypothyroidism. The Journal of Clinical Endocrinology & Metabolism. 2007;92:4575–82.

49. Yamada S, Horiguchi K, Akuzawa M, Sakamaki K, Yamada E, Ozawa A, et al. The impact of age-and sex-specific reference ranges for serum thyrotropin and free thyroxine on the diagnosis of subclinical thyroid dysfunction: A multicenter study from japan. Thyroid. 2023;33:428–39.

50. Mammen JS. Interpreting elevated TSH in older adults. Current opinion in endocrine and metabolic research. 2019;5:68–73.

51. Razvi S. Thyroid-function reference ranges in the diagnosis of thyroid dysfunction in adults. Nature Reviews Endocrinology. 2024;20:253–4.

52. Kuś A, Sterenborg RB, Haug EB, Galesloot TE, Visser WE, Smit JW, et al. Towards personalized TSH reference ranges: A genetic and population-based approach in three independent cohorts. Thyroid. 2024;34:969–79.

53. Spencer CA, Takeuchi M, Kazarosyan M. Current status and performance goals for serum thyrotropin (TSH) assays. Clinical chemistry. 1996;42:140–5.

54. Barth JH, Luvai A, Jassam N, Mbagaya W, Kilpatrick ES, Narayanan D, et al. Comparison of method-related reference intervals for thyroid hormones: Studies from a prospective reference population and a literature review. Annals of clinical biochemistry. 2018;55:107–12.

55. Beckett G, MacKenzie F. Thyroid guidelines-are thyroid-stimulating hormone assays fit for purpose? Annals of clinical biochemistry. 2007;44:203–8.

56. Horowitz GL. Estimating reference intervals. American Journal of Clinical Pathology. 2010;133:175–7.

57. Waise A, Price HC. The upper limit of the reference range for thyroid-stimulating hormone should not be confused with a cut-off to define subclinical hypothyroidism. Annals of clinical biochemistry. 2009;46:93–8.

58. Kalaria T, Sanders A, Fenn J, Ashby HL, Mohammed P, Buch HN, et al. The diagnosis and management of subclinical hypothyroidism is assay-dependent–implications for clinical practice. Clinical Endocrinology. 2021;94:1012–6.

59. Brodersen J, Schwartz LM, Heneghan C, O'Sullivan JW, Aronson JK, Woloshin S. Overdiagnosis: What it is and what it isn't. 2018;23:1–3.

60. Roelfsema F, Veldhuis JD. Thyrotropin secretion patterns in health and disease. Endocrine reviews. 2013;34:619–57.

61. Razvi S, Ryan V, Ingoe L, Pearce SH, Wilkes S. Age-related serum thyroid-stimulating hormone reference range in older patients treated with levothyroxine: A randomized controlled feasibility trial (SORTED 1). European Thyroid Journal. 2020;9:40–8.

62. Boucai L, Hollowell JG, Surks MI. An approach for development of age-, gender-, and ethnicity-specific thyrotropin reference limits. Thyroid. 2011;21:5–11.

63. Symonds C, Kline G, Gjata I, Rose MS, Guo M, Cooke L, et al. Levothyroxine prescribing and laboratory test use after a minor change in reference range for thyroid-stimulating hormone. CMAJ. 2020;192:E469–75.

64. Calissendorff J, Falhammar H. To treat or not to treat subclinical hypothyroidism, what is the evidence? Medicina. 2020;56:40.

65. Rahim A, Goundrey-Smith S. Thyroid replacement therapy: The hidden challenges. Journal of Endocrinology and Metabolism. 2014;3:127–31.

66. Meizlik P, Cucchiara A, Kannan L, Mays J, Jolivert J, Scattergood T, et al. Physiologic

effects of levothyroxine and liothyronine in the in older individuals with persistent subclinical hypothyroidism: A randomized, double-blind, cross-over study. Journal of the Endocrine Society. 2021;5 Suppl 1:A976–77.

67. Biondi B. Natural history, diagnosis and management of subclinical thyroid dysfunction. Best practice & research Clinical endocrinology & metabolism. 2012;26:431–46.

68. Duntas LH, Jonklaas J. Levothyroxine dose adjustment to optimise therapy throughout a patient's lifetime. Advances in therapy. 2019;36:30–46.

69. Flynn RW, Bonellie SR, Jung RT, MacDonald TM, Morris AD, Leese GP. Serum thyroid-stimulating hormone concentration and morbidity from cardiovascular disease and fractures in patients on long-term thyroxine therapy. The Journal of Clinical Endocrinology & Metabolism. 2010;95:186–93.

70. Benvenga S. When thyroid hormone replacement is ineffective? Current Opinion in Endocrinology, Diabetes and Obesity. 2013;20:467–77.

71. Ramadhan A, Tamilia M. Treatment-refractory hypothyroidism. CMAJ. 2012;184:205–9.

72. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function in subclinical hypothyroidism: Importance of clinical follow-up and therapy. European Journal of Endocrinology. 2011;164:317.

73. Bekkering G, Agoritsas T, Lytvyn L, Heen A, Feller M, Moutzouri E, et al. Thyroid hormones treatment for subclinical hypothyroidism: A clinical practice guideline. BMJ. 2019;365:l2006.

74. Vasileiou M, Gilbert J, Fishburn S, Boelaert K. Thyroid disease assessment and management: Summary of NICE guidance. BMJ. 2020;368:m41.

75. Dave JA, Klisiewicz A, Bayat Z, Mohamed NA, Stevens Z, Mollentze WF, et al. SEMDSA/ACE-SA guideline for the management of hypothyroidism in adults. Journal of Endocrinology, Metabolism and Diabetes of South Africa. 2015;20:18–26.

76. El Kawkgi OM, Brito JP. Screening for thyroid dysfunction: Prevention of overdiagnosis and overtreatment. CMAJ. 2019;191:E1260–1.

77. Ku EJ, Yoo WS, Chung HK. Management of subclinical hypothyroidism: A focus on proven health effects in the 2023 korean thyroid association guidelines. Endocrinology and Metabolism. 2023;38:381–91.

78. Garber JR, Cobin RH, Gharib H, Hennessey JV, Klein I, Mechanick JI, et al. Clinical practice guidelines for hypothyroidism in adults: Cosponsored by the american association of clinical endocrinologists and the american thyroid association. Thyroid.

2012;22:1200–35.

79. Brenta G, Vaisman M, Sgarbi JA, Bergoglio LM, Andrada NC de, Bravo PP, et al. Clinical practice guidelines for the management of hypothyroidism. Arquivos Brasileiros de Endocrinologia & Metabologia. 2013;57.

80. Chung HK, Ku EJ, Yoo WS, Kang YE, Kim KJ, Kim BH, et al. 2023 korean thyroid association management guidelines for patients with subclinical hypothyroidism. ijt. 2023;16:32–50.

81. 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;162:35–45.

82. Birk-Urovitz E, Del Giudice ME, Meaney C, Grewal K. Use of thyroid-stimulating hormone tests for identifying primary hypothyroidism in family medicine patients. Canadian Family Physician. 2017;63:e389–94.

83. Werhun A, Hamilton W. Thyroid function testing in primary care: Overused and under-evidenced? A study examining which clinical features correspond to an abnormal thyroid function result. Family practice. 2015;32:187–91.

84. Wintemute K, Greiver M, McIsaac W, Del Giudice ME, Sullivan F, Aliarzadeh B, et al. Choosing wisely canada campaign associated with less overuse of thyroid testing: Retrospective parallel cohort study. 2019;65:e487–96.

85. Wootton T, Bates R. Things we do for no reason™: Routine thyroid-stimulating hormone testing in the hospital. Journal of hospital medicine. 2020;15:560–2.

86. Dalal S, Bhesania S, Silber S, Mehta P. Use of electronic clinical decision support and hard stops to decrease unnecessary thyroid function testing. BMJ Open Quality. 2017;6:u223041. w8346.

87. Crampton N, Kalia S, Del Giudice ME, Wintemute K, Sullivan F, Aliarzadeh B, et al. Over-use of thyroid testing in canadian and UK primary care in frequent attenders: A cross-sectional study. International Journal of Clinical Practice. 2021;75:e14144.

88. Family Physicians of Canada C of. Thirteen things physicians and patients should question. 2015.

89. Willis BH. Spectrum bias—why clinicians need to be cautious when applying diagnostic test studies. Family practice. 2008;25:390–6.

90. Scargill JJ, Livingston M, Holland D, Duff CJ, Fryer AA, Heald AH. Monitoring thyroid function in patients on levothyroxine. Assessment of conformity to national guidance and variability in practice. Experimental and Clinical Endocrinology & Diabetes.

2017;125:625–33.

91. Andersen S, Karmisholt J, Bruun NH, Riis J, Noahsen P, Westergaard L, et al. Interpretation of TSH and T4 for diagnosing minor alterations in thyroid function: A comparative analysis of two separate longitudinal cohorts. Thyroid research. 2022;15:19.

92. Javaid U, Kennedy D, Addison C, Tsatlidis V, Razvi S. Frequency, determinants and costs of thyroid function testing in a laboratory serving a large population. European Journal of Endocrinology. 2022;186:553–60.

93. Veltri F, Rocha FO, Willems D, Praet J-P, Grabczan L, Kleynen P, et al. Prevalence of thyroid dysfunction and autoimmunity in the older population and implications of age-specific reference ranges. Clinica chimica acta. 2017;465:34–9.

94. Somwaru LL, Arnold AM, Joshi N, Fried LP, Cappola AR. High frequency of and factors associated with thyroid hormone over-replacement and under-replacement in men and women aged 65 and over. The Journal of Clinical Endocrinology & Metabolism. 2009;94:1342–5.

95. Chu JW, Crapo LM. The treatment of subclinical hypothyroidism is seldom necessary. The Journal of Clinical Endocrinology & Metabolism. 2001;86:4591–9.

96. McDermott MT, Ridgway EC. Subclinical hypothyroidism is mild thyroid failure and should be treated. The journal of clinical endocrinology & metabolism. 2001;86:4585–90.

97. Peeters RP, Brito JP. Subclinical hypothyroidism: To treat or not to treat? European journal of endocrinology. 2020;183:D15–24.

98. Sawka AM, Cappola AR, Peeters RP, Kopp PA, Bianco AC, Jonklaas J. Patient context and thyrotropin levels are important when considering treatment of subclinical hypothyroidism. Thyroid. 2019;29:1359–63.

99. Feller M, Snel M, Moutzouri E, Bauer DC, Montmollin M de, Aujesky D, et al. Association of thyroid hormone therapy with quality of life and thyroid-related symptoms in patients with subclinical hypothyroidism: A systematic review and meta-analysis. JAMA. 2018;320:1349–59.

100. Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RG, Mooijaart SP, et al. Thyroid hormone therapy for older adults with subclinical hypothyroidism. New England Journal of Medicine. 2017;376:2534–44.

101. Taylor P, Bianco AC. Urgent need for further research in subclinical hypothyroidism. Nature Reviews Endocrinology. 2019;15:503–4.

102. Koulouri O, Moran C, Halsall D, Chatterjee K, Gurnell M. Pitfalls in the measurement and interpretation of thyroid function tests. Best practice & research Clinical

endocrinology & metabolism. 2013;27:745–62.

103. Cappola AR, Cooper DS. Screening and treating subclinical thyroid disease: Getting past the impasse. Annals of internal medicine. 2015;162:664–5.

104. Authority NBS. Prescription cost analysis – england 2020/21. Report. 2021.

105. Taylor PN, Iqbal A, Minassian C, Sayers A, Draman MS, Greenwood R, et al. Falling threshold for treatment of borderline elevated thyrotropin levels—balancing benefits and risks: Evidence from a large community-based study. JAMA internal medicine. 2014;174:32–9.

106. Goichot B, Pearce SH. Subclinical thyroid disease: Time to enter the age of evidence-based medicine. Thyroid. 2012;22:765–8.

107. 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 PreCIS database study. Thyroid. 2011;21:837–43.

108. Rodondi N, Elzen WP den, Bauer DC, Cappola AR, Razvi S, Walsh JP, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. JAMA. 2010;304:1365–74.

109. Kong WM, Sheikh MH, Lumb PJ, Freedman DB, Crook M, Doré CJ, et al. A 6-month randomized trial of thyroxine treatment in women with mild subclinical hypothyroidism. The American journal of medicine. 2002;112:348–54.

110. Loh HH, Lim LL, Yee A, Loh HS. Association between subclinical hypothyroidism and depression: An updated systematic review and meta-analysis. BMC psychiatry. 2019;19:1–10.

111. Cappola AR, Fried LP, Arnold AM, Danese MD, Kuller LH, Burke GL, et al. Thyroid status, cardiovascular risk, and mortality in older adults. Jama. 2006;295:1033–41.

112. Parle J, Roberts L, Wilson S, Pattison H, Roalfe A, Haque M, et al. A randomized controlled trial of the effect of thyroxine replacement on cognitive function in community-living elderly subjects with subclinical hypothyroidism: The birmingham elderly thyroid study. The Journal of Clinical Endocrinology & Metabolism. 2010;95:3623–32.

113. Boekholdt SM, Titan SM, Wiersinga WM, Chatterjee K, Basart DC, Luben R, et al. Initial thyroid status and cardiovascular risk factors: The EPIC-norfolk prospective population study. Clinical endocrinology. 2010;72:404–10.

114. Roberts LM, Pattison H, Roalfe A, Franklyn J, Wilson S, Hobbs FR, et al. Is subclinical thyroid dysfunction in the elderly associated with depression or cognitive dysfunction?

Annals of internal medicine. 2006;145:573–81.

115. Moher D. The problem of duplicate systematic reviews. BMJ. 2013;347:f5040.

116. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. American journal of epidemiology. 2016;183:758–64.

117. Danaei G, García Rodríguez LA, Cantero OF, Logan RW, Hernán MA. Electronic medical records can be used to emulate target trials of sustained treatment strategies. Journal of Clinical Epidemiology. 2018;96:12–22.

118. Reviews C for, Dissemination. PROSPERO international prospective register of systematic reviews. 2021;2021.

119. Bauer BS, Azcoaga-Lorenzo A, Agrawal U, McCowan C. Management strategies for patients with subclinical hypothyroidism: A protocol for an umbrella review. Systematic reviews. 2021;10:1–6.

120. Bauer BS, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C. The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review. PLOS ONE. 2022;17:e0268070.

121. Zhong J, Mu D, Zou Y, Li L, Cheng X, Qiu L. High thyrotropin levels and risk of mortality in the elderly with subclinical hypothyroidism: A systematic review and meta-analysis. Endocrine Practice. 2023;29:206–13.

122. Yang G, Wang Y, Ma A, Wang T. Subclinical thyroid dysfunction is associated with adverse prognosis in heart failure patients with reduced ejection fraction. BMC cardiovascular disorders. 2019;19:83.

123. Holley M, Razvi S, Farooq MS, Dew R, Maxwell I, Wilkes S. Cardiovascular and bone health outcomes in older people with subclinical hypothyroidism treated with levothyroxine: A systematic review and meta-analysis. Systematic Reviews. 2024;13:123.

124. Akintola AA, Jansen SW, Van Bodegom D, Van der Grond J, Westendorp RG, Craen AJ de, et al. Subclinical hypothyroidism and cognitive function in people over 60 years: A systematic review and meta-analysis. Frontiers in aging neuroscience. 2015;7:150.

125. Van Vliet NA, Van Heemst D, Almeida OP, Åsvold BO, Aubert CE, Bae JB, et al. Association of thyroid dysfunction with cognitive function: An individual participant data analysis. JAMA internal medicine. 2021;181:1440–50.

126. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. International journal of evidence-based healthcare. 2015;13:132–40.

127. Choi GJ, Kang H. The umbrella review: A useful strategy in the rain of evidence. The Korean Journal of Pain. 2022;35:127–8.

128. Hoffmann F, Allers K, Rombey T, Helbach J, Hoffmann A, Mathes T, et al. Nearly 80 systematic reviews were published each day: Observational study on trends in epidemiology and reporting over the years 2000-2019. Journal of Clinical Epidemiology. 2021;138:1–11.

129. Smela B, Toumi M, Świerk K, Gawlik K, Clay E, Boyer L. Systematic literature reviews over the years. Journal of Market Access & Health Policy. 2023;11:2244305.

130. Slim K, Marquillier T. Umbrella reviews: A new tool to synthesize scientific evidence in surgery. Journal of Visceral Surgery. 2022;159:144–9.

131. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: An international prospective register of systematic reviews. Systematic reviews. 2012;1:1–9.

132. Ballard M, Montgomery P. Risk of bias in overviews of reviews: A scoping review of methodological guidance and four-item checklist. Research Synthesis Methods. 2017;8:92–108.

133. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evidence-based mental health. 2018;21:95–100.

134. Ioannidis JP. Integration of evidence from multiple meta-analyses: A primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. CMAJ. 2009;181:488–93.

135. Gates M, Gates A, Guitard S, Pollock M, Hartling L. Guidance for overviews of reviews continues to accumulate, but important challenges remain: A scoping review. Systematic reviews. 2020;9:1–19.

136. Rada G, Perez D, Araya-Quintanilla F, Avila C, Bravo-Soto G, Bravo-Jeria R, et al. Epistemonikos: A comprehensive database of systematic reviews for health decision-making. BMC medical research methodology. 2020;20:1–7.

137. Gusenbauer M, Haddaway NR. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of google scholar, PubMed, and 26 other resources. Research synthesis methods. 2020;11:181–217.

138. Goossen K, Hess S, Lunny C, Pieper D. Database combinations to retrieve systematic reviews in overviews of reviews: A methodological study. BMC medical research methodology. 2020;20:1–15.

139. Choi GJ, Kang H. Introduction to umbrella reviews as a useful evidence-based

practice. Journal of lipid and atherosclerosis. 2023;12:3–11.

140. Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Chapter 10: Umbrella reviews in JBI manual for evidence synthesis. Book. Joanna Briggs Institute (JBI); 2020.

141. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.

142. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. Journal of clinical epidemiology. 2019;108:26–33.

143. Perry R, Whitmarsh A, Leach V, Davies P. A comparison of two assessment tools used in overviews of systematic reviews: ROBIS versus AMSTAR-2. Systematic Reviews. 2021;10:273.

144. Moberg J, Oxman AD, Rosenbaum S, Schünemann HJ, Guyatt G, Flottorp S, et al. The GRADE evidence to decision (EtD) framework for health system and public health decisions. Health research policy and systems. 2018;16:1–15.

145. Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME, et al. Judging the quality of evidence in reviews of prognostic factor research: Adapting the GRADE framework. Systematic reviews. 2013;2:1–12.

146. Mercuri M, Gafni A. The evolution of GRADE (part 1): Is there a theoretical and/or empirical basis for the GRADE framework? Journal of Evaluation in Clinical Practice. 2018;24:1203–10.

147. Mercuri M, Gafni A. The evolution of GRADE (part 2): Still searching for a theoretical and/or empirical basis for the GRADE framework. Journal of Evaluation in Clinical Practice. 2018;24:1211–22.

148. Borgerson K. Valuing evidence: Bias and the evidence hierarchy of evidence-based medicine. Perspectives in biology and medicine. 2009;52:218–33.

149. Hennessy EA, Johnson BT. Examining overlap of included studies in meta-reviews: Guidance for using the corrected covered area index. Research synthesis methods. 2020;11:134–45.

150. Lunny C, Pieper D, Thabet P, Kanji S. Managing overlap of primary study results across systematic reviews: Practical considerations for authors of overviews of reviews. BMC Medical Research Methodology. 2021;21:140.

151. Pieper D, Antoine S-L, Mathes T, Neugebauer EA, Eikermann M. Systematic review

finds overlapping reviews were not mentioned in every other overview. Journal of clinical epidemiology. 2014;67:368–75.

152. Bougioukas KI, Vounzoulaki E, Mantsiou CD, Savvides ED, Karakosta C, Diakonidis T, et al. Methods for depicting overlap in overviews of systematic reviews: An introduction to static tabular and graphical displays. Journal of Clinical Epidemiology. 2021;132:34–45.

153. Lunny C, Reid EK, Neelakant T, Chen A, Zhang JH, Shinger G, et al. A new taxonomy was developed for overlap across' overviews of systematic reviews': A meta-research study of research waste. Research Synthesis Methods. 2022;13:315–29.

154. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. The Milbank Quarterly. 2016;94:485–514.

155. Bauer BS, Azcoaga-Lorenzo A, Agrawal U, Fagbamigbe AF, McCowan C. Subclinical hypothyroidism in wales from 2000 to 2021: A descriptive cohort study based on electronic health records. PLOS ONE. 2024;19:e0298871.

156. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: A systematic sampling review. Jama. 2007;297:1233–40.

157. Sauer CM, Chen L-C, Hyland SL, Girbes A, Elbers P, Celi LA. Leveraging electronic health records for data science: Common pitfalls and how to avoid them. The Lancet Digital Health. 2022;4:e893–898.

158. Schneeweiss S, Eddings W, Glynn RJ, Patorno E, Rassen J, Franklin JM. Variable selection for confounding adjustment in high-dimensional covariate spaces when analyzing healthcare databases. Epidemiology. 2017;28:237–48.

159. Farmer R, Mathur R, Bhaskaran K, Eastwood SV, Chaturvedi N, Smeeth L. Promises and pitfalls of electronic health record analysis. Diabetologia. 2018;61:1241–8.

160. Young JC, Conover MM, Jonsson Funk M. Measurement error and misclassification in electronic medical records: Methods to mitigate bias. Current epidemiology reports. 2018;5:343–56.

161. Calanzani N, Vojt G, Weller D, Campbell C. Exploring challenges and strategies when using routine data in research: A systematic review. Report. 2017.

162. McGee G, Haneuse S, Coull BA, Weisskopf MG, Rotem RS. On the nature of informative presence bias in analyses of electronic health records. Epidemiology. 2022;33:105–13.

163. Thayer D, Rees A, Kennedy J, Collins H, Harris D, Halcox J, et al. Measuring follow-up time in routinely-collected health datasets: Challenges and solutions. PLoS One. 2020;15:e0228545.

164. 2024;2024.

165. Edwards L, Pickett J, Ashcroft DM, Dambha-Miller H, Majeed A, Mallen C, et al. UK research data resources based on primary care electronic health records: Review and summary for potential users. BJGP Open. 2023;7.

166. Roskams M. Population and household estimates, england and wales: Census 2021. Report. 2022.

167. Ford DV, Jones KH, Verplancke J-P, Lyons RA, John G, Brown G, et al. The SAIL databank: Building a national architecture for e-health research and evaluation. BMC health services research. 2009;9:1–12.

168. Databank S. Information governance. 2024;2024.

169. Tai TW, Anandarajah S, Dhoul N, De Lusignan S. Variation in clinical coding lists in UK general practice: A barrier to consistent data entry? Informatics in primary care. 2007;15:143–50.

170. Stone M, Camosso-Stefinovic J, Wilkinson J, De Lusignan S, Hattersley A, Khunti K. Incorrect and incomplete coding and classification of diabetes: A systematic review. Diabetic Medicine. 2010;27:491–7.

171. Zghebi SS, Reeves D, Grigoroglou C, McMillan B, Ashcroft DM, Parisi R, et al. Clinical code usage in UK general practice: A cohort study exploring 18 conditions over 14 years. BMJ open. 2022;12:e051456.

172. Harkness EF, Grant L, O'Brien SJ, Chew-Graham CA, Thompson DG. Using read codes to identify patients with irritable bowel syndrome in general practice: A database study. BMC family practice. 2013;14:1–10.

173. Davies AA Gareth; Akbari, teams; CC-C. The welsh results reports service (WRRS) data: Wales population-scale pathology data, a national data asset. Report. SAIL Databank; 2022.

174. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford DV, et al. The SAIL databank: Linking multiple health and social care datasets. BMC medical informatics and decision making. 2009;9:1–8.

175. Jones KH, Ford DV, Jones C, Dsilva R, Thompson S, Brooks CJ, et al. A case study of the secure anonymous information linkage (SAIL) gateway: A privacy-protecting remote access system for health-related research and evaluation. Journal of biomedical informatics. 2014;50:196–204.

176. Jones KH, Ford DV, Thompson S, Lyons R. A profile of the SAIL databank on the UK secure research platform. International Journal of Population Data Science. 2019;4:1134.

7

Appendices

Appendix I: Supplementary Material for Chapter 2

Protocol Supplementary Material

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P: Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	☒	☐	1-2
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☒	
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	☒	☐	56
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	☒	☐	4-11
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	☒	☐	324-327
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☒	☐	281-283
Support					
Sources	5a	Indicate sources of financial or other support for the review	☒	☐	319-322
Sponsor	5b	Provide name for the review funder and/or sponsor	☒	☐	319-322
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	☒	☐	319-322
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	☒	☐	61-90

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	☒	☐	91-100
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	☒	☐	112-157
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	☒	☐	159-173
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	☒	☐	Additional file 2
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	☒	☐	177-183
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	☒	☐	177-189
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	☒	☐	191-209
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	☒	☐	193-203
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	☒	☐	134-140
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	☒	☐	211-226
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	☒	☐	228-242
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	☐	☒	
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-	☒	☐	233-236

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
		regression)			
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	☒	☐	237-242
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☒	☐	211-221
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☒	☐	244-249

Additional file 2: MEDLINE search strategy

	Inclusion criteria	Search terms
Population	Patients with subclinical hypothyroidism No age restrictions Limit by pregnancy (screening stage)	1. exp Hypothyroidism/ 2. exp Thyroid diseases/ 3. hypothyroid*.tw 4. (thyroid? adj3 deficien*).tw 5. (thyroid? adj3 insufficien*).tw 6. (thyroid? adj3 failure?).tw 7. (thyroid? adj3 low adj3 product*).tw 8. (thyroid? adj3 under adj3 product*).tw 9. (thyroid? adj3 underactiv*).tw 10. (thyroid? adj3 hypofunction).tw 11. (thyroid? adj3 d?sfunction*).tw 12. or/1-11 13. (mild* OR sub-clinic* OR subclinic*).tw 14. 12 and 13 15. (tsh adj3 (elevat* or increas* or high*)).tw. 16. (SHT OR SCH).tw 17. 14 or 15 or 16
Intervention	Treatment with Levothyroxine	18. exp Hormone replacement therapy/ 19. (thyroid? adj3 (therapy OR treat*)).tw 20. exp Thyroxine/ 21. (thyroxin* or levothyrox* or levo-thyrox* or l-thyrox* or L-T4 or LT4).tw. 22. or/18-21
	Follow-up with no treatment	23. (follow?up OR monitor* OR observ* OR surveil*).tw 24. ((no OR lack*) adj2 (treatment* OR therap* OR intervention*)).tw 25. (untreated OR ('not' adj2 treat*)).tw 26. or/23-25
Comparator	NA	
Outcomes	Cardiovascular outcomes	27. Exp Cardiovascular diseases/ OR exp Heart diseases/ OR exp Myocardial ischemia/ OR exp Vascular diseases/ OR exp Arteriosclerosis/ 28. Carotid Intima Media Thickness/ OR Intima-Media Thickness, Carotid/ OR Atherosclerosis/ OR Atherosclerоses/ OR Atherogenesis/ 29. ((cardiovasc* OR vasc* OR cardio* OR cardia* OR heart* OR coronary* OR myocard* OR pericard* OR isch$em*) adj2 (disease? OR event? OR arrest? OR fail* OR mortality)).tw. 30. (myocardi* adj (infarct* OR revascular* OR re-vascular* OR isch$emi*)).tw. 31. (heart attack* OR angina).tw.

		32. (morbid* adj5 (cardio* OR cardia* OR heart* OR coronary* OR myocard* OR pericard* OR isch$em*)).tw. 33. peripheral arter* disease*.tw. 34. (emboli* OR arrhythmi* OR thrombo* OR atrial fibrillat* OR atrial flutter* OR tachycardi* OR endocardi* OR (sick adj sinus)).tw. 35. (isch$emi* adj2 (vascular OR heart)).tw. 36. (flow-mediated vasodilat* OR flow-mediated dilat* OR endothelial-dependent vasodilat* OR endothelial-dependent dilat* OR endothelial function$ OR carotid intima-media thickness OR intima-media thickness OR carotid-wall thickness OR carotid atherosclerosis OR C-IMT).tw. 37. (lipid$ OR cholesterol OR triglyceride$ OR LDL OR HDL).tw 38. or/27-37
	QoL	39. exp Quality of Life/ 40. quality of life.tw. 41. (QoL OR HRQoL).tw 42. or/39-41
	Cerebrovascular outcomes	43. exp Stroke/ 44. exp Ischemic Attack, Transient/ 45. (stroke$ OR apoplexy).tw 46. ((cerebrovasc* OR cerebral vascular OR brain) adj2 (disease? OR event? OR arrest? OR fail* OR mortality OR accident* OR death*)).tw 47. ((brain* OR cerebral OR lacunar) adj2 infarct*).tw 48. (isch$emi* adj2 (transient OR attack* OR cerebral or brain)).tw 49. or/43-48
	Frailty fractures	50. exp Frailty/ 51. frail*.tw 52. exp Fractures, bone/ OR fracture$.tw 53. (50 or 51) and 52
	Mortality	54. (mortality OR death*).tw
Study design	Systematic reviews	(SIGN Sys Reviews Filter) 55. Meta-Analysis as Topic/ 56. meta analy$.tw. 57. metaanaly$.tw. 58. Meta-Analysis/ 59. (systematic adj (review$1 or overview$1)).tw. 60. exp Review Literature as Topic/ 61. or/55-60 62. cochrane.ab. 63. embase.ab. 64. (psychlit or psyclit).ab. 65. (psychinfo or psycinfo).ab. 66. (cinahl or cinhal).ab. 67. science citation index.ab.

		68. bids.ab. 69. cancerlit.ab. 70. or/62-69 71. reference list$.ab. 72. bibliograph$.ab. 73. hand-search$.ab. 74. relevant journals.ab. 75. manual search$.ab. 76. or/71-75 77. selection criteria.ab. 78. data extraction.ab. 79. 77 or 78 80. Review/ 81. 79 and 80 82. Comment/ 83. Letter/ 84. Editorial/ 85. animal/ 86. human/ 87. 85 not (85 and 86) 88. or/82-84,87 89. 61 or 70 or 76 or 81 90. 89 not 88
PICOS		91. 17 and (22 or 26) and (38 or 42 or 49 or 53 or 54)
PICOS + Filter		92. 90 and 91

Data Extraction Template

Review citation details

first author (year)

journal

title

Review purpose

question(s)

aim/objectives

Review methods

protocol registered

reporting guideline used (e.g. PRISMA)

sources/databases searched

date range of searches

date of last search (if repeated)

synthesis methods/analysis

inclusion criteria: PICO – participants, intervention(s), comparator, outcomes of interest, setting/context

definition/description of the outcome(s) – (by review authors)

exclusion criteria (if explicitly stated)

Reviews results

number of included studies

study designs included (e.g. RCT, cohort etc)

country of origin of included studies

total number of participants

study details

- author, year of publication, title, study type
- country
- PICO
- number of participants
- number of controls
- time of follow-up

Quality appraisal

- quality appraisal tool used
- quality appraisal rating for included studies
- risk of bias assessment

Quality of evidence (e.g. GRADE)

Review outcomes

relevant study

meta-analysis (Y/N)

- number of studies included in meta-analysis
- sub-group analysis criteria

effect size

- measure of effect size {eg, 'Hedge's g', incidence rate ratio (IRR), odds ratio (OR), risk ratio (RR)} of association (preferably unadjusted)}
- the outcome and its CI (eg, '1.35 (95% CI 0.97 to 1.73)')
- statistical significance ($P < 0.05$)

Significance/direction

- p value; I^2; prediction interval; CI of the largest study; equivalent OR; Egger test; excess significance test

Overall review findings/conclusions

Heterogeneity across included studies

Review limitations/potential biases

Additional notes/comments

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:
- ☐ Population
- ☐ Intervention
- ☐ Comparator group
- ☐ Outcome

Optional (recommended)
- ☐ Timeframe for follow-up

☐ Yes
☐ No

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:
The authors state that they had a written protocol or guide that included ALL the following:
- ☐ review question(s)
- ☐ a search strategy
- ☐ inclusion/exclusion criteria
- ☐ a risk of bias assessment

For Yes:
As for partial yes, plus the protocol should be registered and should also have specified:
- ☐ a meta-analysis/synthesis plan, if appropriate, *and*
- ☐ a plan for investigating causes of heterogeneity
- ☐ justification for any deviations from the protocol

☐ Yes
☐ Partial Yes
☐ No

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:
- ☐ *Explanation for* including only RCTs
- ☐ OR *Explanation for* including only NRSI
- ☐ OR *Explanation for* including both RCTs and NRSI

☐ Yes
☐ No

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):
- ☐ searched at least 2 databases (relevant to research question)
- ☐ provided key word and/or search strategy
- ☐ justified publication restrictions (e.g. language)

For Yes, should also have (all the following):
- ☐ searched the reference lists / bibliographies of included studies
- ☐ searched trial/study registries
- ☐ included/consulted content experts in the field
- ☐ where relevant, searched for grey literature
- ☐ conducted search within 24 months of completion of the review

☐ Yes
☐ Partial Yes
☐ No

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:
- ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- ☐ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

☐ Yes
☐ No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers achieved consensus on which data to extract from included studies
- ☐ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

☐ Yes
☐ No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- ☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- ☐ Justified the exclusion from the review of each potentially relevant study

☐ Yes
☐ Partial Yes
☐ No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):

- ☐ described populations
- ☐ described interventions
- ☐ described comparators
- ☐ described outcomes
- ☐ described research designs

For Yes, should also have ALL the following:

- ☐ described population in detail
- ☐ described intervention in detail (including doses where relevant)
- ☐ described comparator in detail (including doses where relevant)
- ☐ described study's setting
- ☐ timeframe for follow-up

☐ Yes
☐ Partial Yes
☐ No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from

- ☐ unconcealed allocation, *and*
- ☐ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- ☐ allocation sequence that was not truly random, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only NRSI

NRSI

For Partial Yes, must have assessed RoB:

- ☐ from confounding, *and*
- ☐ from selection bias

For Yes, must also have assessed RoB:

- ☐ methods used to ascertain exposures and outcomes, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only RCTs

10. Did the review authors report o n the sources of funding for the studies included in the review?

For Yes

- ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

☐ Yes
☐ No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
  - ☐ AND investigated the causes of any heterogeneity

☐ Yes
☐ No
☐ No meta-analysis conducted

For NRSI

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present
  - ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available
  - ☐ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review

☐ Yes
☐ No
☐ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.

☐ Yes
☐ No
☐ No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results

☐ Yes
☐ No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes:

- ☐ There was no significant heterogeneity in the results
- ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

☐ Yes
☐ No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

☐ Yes
☐ No
☐ No meta-analysis conducted

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	
For Yes: ☐ The authors reported no competing interests OR ☐ The authors described their funding sources and how they managed potential conflicts of interest	☐ Yes ☐ No

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

Umbrella Review Supplementary Material

Additional File 1. MEDLINE Search Strategy

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to February 17, 2021>

Search Strategy:

---

1 exp Hypothyroidism/ (33056)

2 exp Thyroid Diseases/ (149791)

3 hypothyroid*.tw. (35912)

4 (thyroid? adj3 deficien*).tw. (1536)

5 (thyroid? adj3 insufficien*).tw. (488)

6 (thyroid? adj3 failure?).tw. (549)

7 (thyroid? adj3 low adj3 product*).tw. (7)

8 (thyroid? adj3 under adj3 product*).tw. (4)

9 (thyroid? adj3 underactiv*).tw. (27)

10 (thyroid? adj3 hypofunction).tw. (265)

11 (thyroid? adj3 d?sfunction).tw. (5642)

12 or/1-11 (162773)

13 (mild* or sub-clinic* or subclinic*).tw. (453666)

14 12 and 13 (8752)

15 (tsh adj3 (elevat* or increas* or high*)).tw. (6791)

16 (SHT or SCH).tw. (8664)

17 14 or 15 or 16 (22418)

18 exp Hormone Replacement Therapy/ (25185)

19 (thyroid? adj3 (therap* or treat*)).tw. (12185)

20 exp Thyroxine/ (48783)

21 (thyroxin* or levothyrox* or levo-thyrox* or l-thyrox* or L-T4 or LT4).tw. (34600)

22 or/18-21 (94909)

23 (follow?up or monitor* or observ* or surveil*).tw. (4488288)

24 ((no or lack*) adj2 (treatment* or therap* or intervention*)).tw. (137605)

25 (untreated or ('not' adj2 treat*)).tw. (264180)

26 or/23-25 (4777064)

27 exp Cardiovascular diseases/ or exp Heart diseases/ or exp Myocardial ischemia/ or exp Vascular diseases/ or exp Arteriosclerosis/ (2435564)

28 Carotid Intima Media Thickness/ or Intima-Media Thickness, Carotid/ or Atherosclerosis/ or Atheroscleroses/ or Atherogenesis/ (41075)

29 ((cardiovasc* or vasc* or cardio* or cardia* or heart* or coronary* or myocard* or pericard* or isch$em*) adj2 (disease? or event? or arrest? or fail* or mortality)).tw. (745362)

30 (lipid$ or cholesterol or TC or triglyceride$ or LDL or HDL).tw. (756725)

31 (myocardi* adj (infarct* or revascular* or re-vascular* or isch$emi*)).tw. (225198)

32 (heart attack* or angina).tw. (58869)

33 (morbid* adj5 (cardio* or cardia* or heart* or coronary* or myocard* or pericard* or isch$em*)).tw. (25873)

34 peripheral arter* disease*.tw. (14503)

35 (emboli* or arrhythmi* or thrombo* or atrial fibrillat* or atrial flutter* or tachycardi* or endocardi* or (sick adj sinus)).tw. (690392)

36 (flow-mediated vasodilat* or flow-mediated dilat* or endothelial-dependent vasodilat* or endothelial-dependent dilat* or endothelial function$ or carotid intima-media thickness or intima-media thickness or carotid-wall thickness or carotid atherosclerosis or C-IMT).tw. (35362)

37 (isch$emi* adj2 (vascular or heart)).tw. (34721)

38 or/27-37 (3563857)

39 exp Quality of Life/ (204526)

40 quality of life.tw. (295840)

41 (QoL or HRQoL).tw. (56831)

42 or/39-41 (359236)

43 exp Stroke/ (140280)

44 exp Ischemic Attack, Transient/ (20638)

45 (stroke$ or apoplexy).tw. (263191)

46 ((cerebrovasc* or cerebral vascular or brain) adj2 (disease? or event? or arrest? or fail* or mortality or accident* or death*)).tw. (62563)

47 ((brain* or cerebral or lacunar) adj2 infarct*).tw. (27518)

48 (isch$emi* adj2 (transient or attack* or cerebral or brain)).tw. (57232)

49 or/43-48 (387778)

50 exp Frailty/ (3592)

51 frail*.tw. (24209)

52 exp Fractures, Bone/ or fracture$.tw. (310642)

53 (50 or 51) and 52 (1293)

54 (mortality or death*).tw. (1479781)

55 Meta-Analysis as Topic/ (19028)

56 meta analy$.tw. (194305)

57 metaanaly$.tw. (2211)

58 Meta-Analysis/ (126788)

59 (systematic adj (review$1 or overview$1)).tw. (197114)
60 exp Review Literature as Topic/ (15372)
61 or/55-60 (335876)
62 cochrane.ab. (94602)
63 embase.ab. (105156)
64 (psychlit or psyclit).ab. (915)
65 (psychinfo or psycinfo).ab. (40574)
66 (cinahl or cinhal).ab. (32163)
67 science citation index.ab. (3252)
68 bids.ab. (563)
69 cancerlit.ab. (633)
70 or/62-69 (170652)
71 reference list$.ab. (19037)
72 bibliograph$.ab. (19191)
73 hand-search$.ab. (7333)
74 relevant journals.ab. (1216)
75 manual search$.ab. (4784)
76 or/71-75 (46231)
77 selection criteria.ab. (31649)
78 data extraction.ab. (23699)
79 77 or 78 (52968)
80 Review/ (2767148)
81 79 and 80 (29899)
82 Comment/ (894822)
83 Letter/ (1124632)
84 Editorial/ (558685)
85 animal/ (6756006)
86 human/ (19036693)
87 85 not (85 and 86) (4755937)
88 or/82-84,87 (6623165)
89 61 or 70 or 76 or 81 (402770)
90 89 not 88 (382792)
91 17 and (22 or 26) and (38 or 42 or 49 or 53 or 54) (2382)
92 90 and 91 (96)

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:
- ☐ Population
- ☐ Intervention
- ☐ Comparator group
- ☐ Outcome

Optional (recommended)
- ☐ Timeframe for follow-up

☐ Yes
☐ No

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:
The authors state that they had a written protocol or guide that included ALL the following:
- ☐ review question(s)
- ☐ a search strategy
- ☐ inclusion/exclusion criteria
- ☐ a risk of bias assessment

For Yes:
As for partial yes, plus the protocol should be registered and should also have specified:
- ☐ a meta-analysis/synthesis plan, if appropriate, *and*
- ☐ a plan for investigating causes of heterogeneity
- ☐ justification for any deviations from the protocol

☐ Yes
☐ Partial Yes
☐ No

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:
- ☐ *Explanation for* including only RCTs
- ☐ OR *Explanation for* including only NRSI
- ☐ OR *Explanation for* including both RCTs and NRSI

☐ Yes
☐ No

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):
- ☐ searched at least 2 databases (relevant to research question)
- ☐ provided key word and/or search strategy
- ☐ justified publication restrictions (e.g. language)

For Yes, should also have (all the following):
- ☐ searched the reference lists / bibliographies of included studies
- ☐ searched trial/study registries
- ☐ included/consulted content experts in the field
- ☐ where relevant, searched for grey literature
- ☐ conducted search within 24 months of completion of the review

☐ Yes
☐ Partial Yes
☐ No

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:
- ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- ☐ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer.

☐ Yes
☐ No

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers achieved consensus on which data to extract from included studies
- ☐ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

☐ Yes
☐ No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- ☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- ☐ Justified the exclusion from the review of each potentially relevant study

☐ Yes
☐ Partial Yes
☐ No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (ALL the following):

- ☐ described populations
- ☐ described interventions
- ☐ described comparators
- ☐ described outcomes
- ☐ described research designs

For Yes, should also have ALL the following:

- ☐ described population in detail
- ☐ described intervention in detail (including doses where relevant)
- ☐ described comparator in detail (including doses where relevant)
- ☐ described study's setting
- ☐ timeframe for follow-up

☐ Yes
☐ Partial Yes
☐ No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from

- ☐ unconcealed allocation, *and*
- ☐ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- ☐ allocation sequence that was not truly random, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only NRSI

NRSI

For Partial Yes, must have assessed RoB:

- ☐ from confounding, *and*
- ☐ from selection bias

For Yes, must also have assessed RoB:

- ☐ methods used to ascertain exposures and outcomes, *and*
- ☐ selection of the reported result from among multiple measurements or analyses of a specified outcome

☐ Yes
☐ Partial Yes
☐ No
☐ Includes only RCTs

10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes

- ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

☐ Yes
☐ No

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
  - ☐ AND investigated the causes of any heterogeneity

☐ Yes
☐ No
☐ No meta-analysis conducted

For NRSI

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
  - ☐ AND they used an appropriate weighted technique to combine study results, adjusting for heterogeneity if present
  - ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available
  - ☐ AND they reported separate summary estimates for RCTs and NRSI separately when both were included in the review

☐ Yes
☐ No
☐ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.

☐ Yes
☐ No
☐ No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

For Yes:

- ☐ included only low risk of bias RCTs
- ☐ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results

☐ Yes
☐ No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes:

- ☐ There was no significant heterogeneity in the results
- ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

☐ Yes
☐ No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes:

- ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

☐ Yes
☐ No
☐ No meta-analysis conducted

AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

For Yes:

- ☐ The authors reported no competing interests OR
- ☐ The authors described their funding sources and how they managed potential conflicts of interest

☐ Yes
☐ No

To cite this tool: Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, Moher D, Tugwell P, Welch V, Kristjansson E, Henry DA. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

List of papers excluded after full-text screening

Not a systematic review or meta-analysis (n = 23)

1. Leng O, Razvi S. Treatment of subclinical hypothyroidism: assessing when treatment is likely to be beneficial. Expert Review of Endocrinology & Metabolism. 2021 Mar 4;16(2):73-86.
2. Panday P, Franchini AP, Iskander B, Anwer F, Oliveri F, Kakargias F, Hamid P. Subclinical hypothyroidism in geriatric population and its association with heart failure. Cureus. 2021 Apr 5;13(4).
3. Apostu D, Lucaciu O, Oltean-Dan D, Mureșan AD, Moisescu-Pop C, Maxim A, Benea H. The influence of thyroid pathology on osteoporosis and fracture risk: A review. Diagnostics. 2020 Mar;10(3):149.
4. Biondi B, Palmieri EA, Lombardi G, Fazio S. Effects of subclinical thyroid dysfunction on the heart. Annals of internal medicine. 2002 Dec 3;137(11):904-14.
5. Delitala AP, Scuteri A, Maioli M, Mangatia P, Vilardi L, Erre GL. Subclinical hypothyroidism and cardiovascular risk factors. Minerva Medica. 2019 Nov 11;110(6):530-45.
6. Chrysant SG. The current debate over treatment of subclinical hypothyroidism to prevent cardiovascular complications. International Journal of Clinical Practice. 2020 Jul;74(7):e13499.
7. Tognini S, Pasqualetti G, Calsolaro V, Polini A, Monzani F. Cognitive function and quality of life in mild thyroid hormone deficiency. Recent patents on endocrine, metabolic & immune drug discovery. 2014 May 1;8(2):124-34.
8. Thvilum M, Brandt F, Brix TH, Hegedüs L. A review of the evidence for and against increased mortality in hypothyroidism. Nature Reviews Endocrinology. 2012 Jul;8(7):417-24.
9. Triggiani V, Angelo Giagulli V, De Pergola G, Licchelli B, Guastamacchia E, Iacoviello M. Mechanisms explaining the influence of subclinical hypothyroidism on the onset and progression of chronic heart failure. Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders). 2016 Mar 1;16(1):2-7.
10. Surks MI, Ortiz E, Daniels GH, Sawin CT, Col NF, Cobin RH, Franklyn JA, Hershman JM, Burman KD, Denke MA, Gorman C. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. Jama. 2004 Jan 14;291(2):228-38.
11. Helfand M. Screening for subclinical thyroid dysfunction in nonpregnant adults: a summary of the evidence for the US Preventive Services Task Force. Annals of internal medicine. 2004 Jan 20;140(2):128-41.
12. Sgarbi JA, Teixeira PF, Maciel LM, Mazeto GM, Vaisman M, Montenegro Junior RM, Ward LS. The Brazilian consensus for the clinical approach and treatment of subclinical hypothyroidism in adults: recommendations of the thyroid Department of the Brazilian Society of Endocrinology and Metabolism. Arquivos Brasileiros de Endocrinologia & Metabologia. 2013;57:166-83.
13. Allan GM, Morros MP, Young J. Subclinical hypothyroidism and TSH screening. Canadian Family Physician. 2020 Mar 1;66(3):188-188.

14. Mariotti S, Cambuli VM. Cardiovascular risk in elderly hypothyroid patients. Thyroid. 2007 Nov 1;17(11):1067-73.
15. Marrakchi S, Kanoun F, Idriss S, Kammoun I, Kachboura S. Arrhythmia and thyroid dysfunction. Herz. 2015 Apr;40(2):101-9.
16. Hennessey JV, Espaillat R. Reversible morbidity markers in subclinical hypothyroidism. Postgraduate Medicine. 2015 Jan 2;127(1):78-91.
17. Velkeniers B, Van MA, Unuane D, Haentjens P. A critical synopsis of meta-analysis in the field of subclinical thyroid disease. In Endocrine Abstracts 2010 Apr 1 (Vol. 22). Bioscientifica.
18. Johnson BE. Thyroid hormone therapy does not improve QoL or symptoms in subclinical hypothyroidism. Annals of Internal Medicine. 2019 Feb;170(4):JC17.
19. Lee, J. H. Early detection and management of SCH is important. Journal of Family Practice. 2006; 55 (6):543
20. Lucas-Martín AM. Hipotiroidismo subclínico: tratar o no tartar. Medicina Clínica. 2004; 122(5):182-183. doi:https://doi.org/10.1016/S0025-7753(04)74187-1.
21. Rodondi N, Maisonneuve P, Razvi S, Elzen WD, Gussekloo J, Iervasi G, Asvold BO, Imaizumi M, Vanderpump M, Westendorp RG, Franklyn JA. Subclinical Hypothyroidism and the Risk of Coronary Heart Disease and Mortality: An Individual Participant Data Analysis from Nine Prospective Cohort Studies. Journal of General Internal Medicine. 2010 Jun 30;25:394-5.
22. Gencer B, Collet TH, Virgini V, Auer R, Rodondi N. Subclinical thyroid dysfunction and cardiovascular outcomes among prospective cohort studies. Endocrine, Metabolic & Immune Disorders-Drug Targets (Formerly Current Drug Targets-Immune, Endocrine & Metabolic Disorders). 2013 Mar 1;13(1):4-12.
23. Rugge B, Balshem H, Sehgal R, et al. Screening and Treatment of Subclinical Hypothyroidism or Hyperthyroidism. Agency for Healthcare Research and Quality (US), Rockville (MD); 2011. PMID: 22299183.

No reported management (treated or not treated) (n = 7)

1. 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 Sep 1;159(3):329-41.
2. Zhu H, Zhang J, Wang J, Zhao X, Gu M. Association of subclinical thyroid dysfunction with bone mineral density and fracture: a meta-analysis of prospective cohort studies. Endocrine. 2020 Mar;67(3):685-98.
3. Tsai TY, Tu YK, Munir KM, Lin SM, Chang RH, Kao SL, Loh CH, Peng CC, Huang HK. Association of hypothyroidism and mortality in the elderly population: a systematic review and meta-analysis. The Journal of Clinical Endocrinology & Metabolism. 2020 Jun 1;105(6):2068-80.
4. Larsson SC, Allara E, Mason AM, Michaëlsson K, Burgess S. Thyroid function and dysfunction in relation to 16 cardiovascular diseases: a Mendelian randomization study. Circulation: Genomic and Precision Medicine. 2019 Mar;12(3):e002468.

5. Ning Y, Cheng YJ, Liu LJ, Sara JD, Cao ZY, Zheng WP, Zhang TS, Han HJ, Yang ZY, Zhang Y, Wang FL. What is the association of hypothyroidism with risks of cardiovascular events and mortality? A meta-analysis of 55 cohort studies involving 1,898,314 participants. BMC medicine. 2017 Dec;15(1):1-5.
6. Moon S, Kim MJ, Yu JM, Yoo HJ, Park YJ. Subclinical hypothyroidism and the risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. Thyroid. 2018 Sep 1;28(9):1101-10.
7. Brenta G, Vaisman M, Sgarbi JA, Bergoglio LM, Andrada NC, Bravo PP, Orlandi AM, Graf H. Clinical practice guidelines for the management of hypothyroidism. Arquivos Brasileiros de Endocrinologia & Metabologia. 2013;57:265-91.

No primary outcome (n = 5)

1. He W, Li S, Zhang JA, Zhang J, Mu K, Li XM. Effect of levothyroxine on blood pressure in patients with subclinical hypothyroidism: a systematic review and meta-analysis. Frontiers in endocrinology. 2018:454.
2. Gómez-Izquierdo J, Filion KB, Boivin JF, Azoulay L, Pollak M, Yu OH. Subclinical hypothyroidism and the risk of cancer incidence and cancer mortality: a systematic review. BMC endocrine disorders. 2020 Dec;20(1):1-0.
3. Gibbons VA. The Epidemiology and Management of Hypothyroidism in General Practice 2011 (Doctoral dissertation, University of Auckland).
4. Burgos N, Toloza FJ, Singh Ospina NM, Brito JP, Salloum RG, Hassett LC, Maraka S. Clinical outcomes after discontinuation of thyroid hormone replacement: a systematic review and meta-analysis. Thyroid. 2021 May 1;31(5):740-51.
5. Segna D, Bauer DC, Feller M, Schneider C, Fink HA, Aubert CE, Collet TH, da Costa BR, Fischer K, Peeters RP, Cappola AR. Association between subclinical thyroid dysfunction and change in bone mineral density in prospective cohorts. Journal of internal medicine. 2018 Jan;283(1):56-72.

Reported markers/midpoints rather than events (n = 3)

1. Chen Y, Tai HY. Levothyroxine in the treatment of overt or subclinical hypothyroidism: a systematic review and meta-analysis. Endocrine Journal. 2020:EJ19-0583.
2. Faber J, Galløe AM. Changes in bone mass during prolonged subclinical hyperthyroidism due to L-thyroxine treatment: a meta-analysis. European journal of Endocrinology. 1994 Apr 1;130(4):350-6.
3. Peng L, GU MJ. Influence of thyroxine treatment on serum lipid levels in patients with subclinical hypothyroidism: a Meta-analysis. Academic Journal of Second Military Medical University. 1985.

Wrong population (n = 4)

1. Gawlik A, Such K, Dejner A, Zachurzok A, Antosz A, Malecka-Tendera E. Subclinical hypothyroidism in children and adolescents: is it clinically relevant?. International Journal of Endocrinology. 2015 Mar 29;2015.
2. Monzani A, Prodam F, Rapa A, Moia S, Agarla V, Bellone S, Bona G. Endocrine disorders in childhood and adolescence. Natural history of subclinical hypothyroidism in children and adolescents and potential effects of replacement therapy: a review. European Journal of Endocrinology. 2012 Dec 10;168(1):R1-1.
3. Salerno M, Improda N, Capalbo D. Management of endocrine disease subclinical hypothyroidism in children. European Journal of Endocrinology. 2020 Aug 1;183(2):R13-28.
4. Bona G, Prodam F, Monzani A. Subclinical hypothyroidism in children: natural history and when to treat. Journal of clinical research in pediatric endocrinology. 2013 Mar;5(Suppl 1):23.

Not subclinical hypothyroidism (n = 3)

1. Chaker L, Baumgartner C, Den Elzen WP, Collet TH, Ikram MA, Blum MR, Dehghan A, Drechsler C, Luben RN, Portegies ML, Iervasi G. Thyroid function within the reference range and the risk of stroke: an individual participant data analysis. The Journal of Clinical Endocrinology & Metabolism. 2016 Nov 1;101(11):4270-82.
2. Yang LB, Jiang DQ, Qi WB, Zhang T, Feng YL, Gao L, Zhao J. Subclinical hyperthyroidism and the risk of cardiovascular events and all-cause mortality: an updated meta-analysis of cohort studies. European Journal of Endocrinology. 2012 Jul 1;167(1):75.
3. Wiersinga WM. Should we treat mild subclinical/mild hyperthyroidism? Yes. European Journal of Internal Medicine. 2011 Aug 1;22(4):324-9.

Full review published in a separate paper (n = 2)

1. Collet TH, Aujesky D, Vittinghoff E, Bauer D, Gussekloo J, Cappola AR, den Elzen WP, Sgarbi J, Cornuz J, Bremner AP, Maciel RM. Auto-immunity, subclinical hypothyroidism and the risk of coronary heart disease and mortality. Journal of General Internal Medicine. 2012 Jul 31;27:S130-S130.
2. Chaker, L.; Baumgartner, C.; Den Elzen, W. J. P.; Ikram, M. A.; Blum, M. R.; Bakker, S. J. L.; Dehghan, A.; Drechsler, C.; Luben, R. N.; Hofman, A.; Portegies, M. L. P.; Medici, M.; Iervasi, G.; Collet, T. H.; Brenmer, A.; Wanner, C.; Iacoviello, M.; Dullaart, R. P.; Sgarbi, J. A.; Ceresini, G.; Westendorp, R. G.; Jukema, J. W.; Imaizumi, M.; Franklyn, J. A.; Bauer, D. C.; Cappola, A. R.; Walsh, J. P.; Razvi, S.; Khaw, K. T.; Volzke, H.; Franco, O. H.; Gussekloo, J.; Rodondi, N.; Peeters, R. P. Subclinical hypothyroidism and the risk of non-fatal and fatal stroke: An individual participant analysis. European Thyroid Journal. 2014 August 3; 96 doi: http://dx.doi.org/10.1159/000365244

Full data unavailable (n = 2)

1. Du Puy RS, Poortvliet RK, Mooijaart SP, Den Elzen WP, Jagger C, Pearce SH, Arai Y, Hirose N, Teh R, Menzies O, Rolleston A. Outcomes of thyroid dysfunction in people aged eighty years and older: an individual patient data meta-analysis of four prospective studies (Towards Understanding Longitudinal International Older People Studies Consortium). Thyroid. 2021 Apr 1;31(4):552-62.
2. Ochs N, Auer R, Bauer DC, Nanchen D, Gussekloo J, Cornuz J, Rodondi N. Meta-analysis: subclinical thyroid dysfunction and the risk for coronary heart disease and mortality. Annals of internal medicine. 2008 Jun 3;148(11):832-45.

Reviews were updated (n = 2)

1. Chaker L, Baumgartner C, Ikram MA, Dehghan A, Medici M, Visser WE, Hofman A, Rodondi N, Peeters RP, Franco OH. Subclinical thyroid dysfunction and the risk of stroke: a systematic review and meta-analysis. European journal of epidemiology. 2014 Nov;29(11):791-800.
2. Helfand M, Redfern CC. Screening for thyroid disease: an update. Annals of internal medicine. 1998 Jul 15;129(2):144-58.

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Pg 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Pg 2 - 4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pg 5 - 7
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pg 7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Pg 8 - 9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Pg 8
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	S1 Appendix
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 8-9
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Pg 10
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Pg 9, Protocol
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Pg 10, Protocol
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 10 - 11
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	NA
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Pg 12 - 13
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	NA
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Pg 12-13
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Pg 12-13
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	NA
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	NA
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	NA
Certainty	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	NA

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Pg 13, Fig 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	S3
Study characteristics	17	Cite each included study and present its characteristics.	Table 2
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Pg 37 - 39, Table 10
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Pg 25 - 36
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Pg 13-25
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Pg 12-13
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	NA
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pg 39-42
	23b	Discuss any limitations of the evidence included in the review.	Pg 44
	23c	Discuss any limitations of the review processes used.	Pg 43
	23d	Discuss implications of the results for practice, policy, and future research.	Pg 44-5
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 4
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 7
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Pg 8
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	NA
Competing interests	26	Declare any competing interests of review authors.	NA
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	NA

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/

PRISMA 2020 for Abstracts Checklist

Section and Topic	Item #	Checklist item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesise results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	No
Registration	12	Provide the register name and registration number.	Yes

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

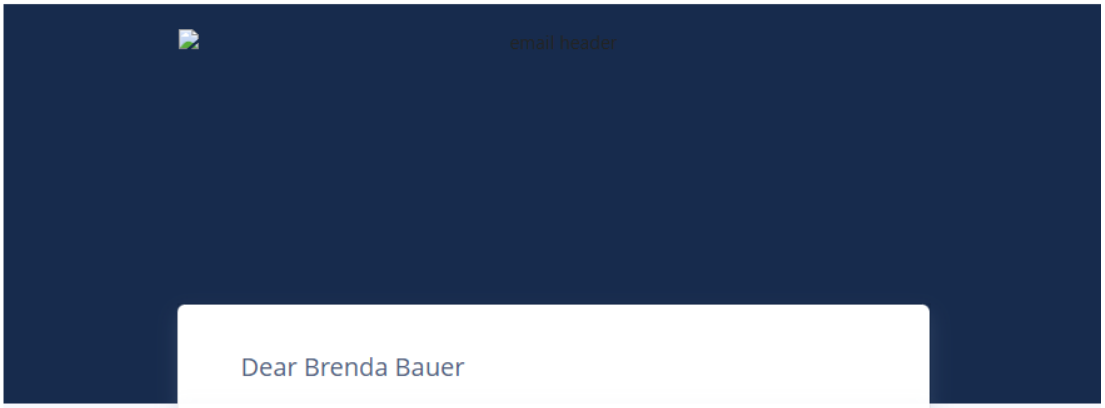
For more information, visit: http://www.prisma-statement.org/

Appendix II: Supplementary Material for Chapter 3

Application 1371 - Subclinical hypothyroidism effects on clinical outcomes has been Approved

SAIL IGRP Team <igrp@chi.swan.ac.uk>
Tue 01/02/2022 10:26 AM

To: Brenda Bauer <bsb1@st-andrews.ac.uk>
Cc: c.r.arkley@swansea.ac.uk <c.r.arkley@swansea.ac.uk>; c.l.mcnerney@swansea.ac.uk <c.l.mcnerney@swansea.ac.uk>; sailbusinessadmin@swansea.ac.uk <sailbusinessadmin@swansea.ac.uk>; e.a.halfpenny@swansea.ac.uk <e.a.halfpenny@swansea.ac.uk>; c.n.batchelder@swansea.ac.uk <c.n.batchelder@swansea.ac.uk>; Cathrine.E.Richards@Swansea.ac.uk <Cathrine.E.Richards@Swansea.ac.uk>



Dear Brenda Bauer

1371 - Subclinical hypothyroidism effects on clinical outcomes

Your proposal to use the SAIL databank has been **approved** by the SAIL Information Governance Review Panel (IGRP). The membership of the IGRP is comprised of representatives from:

- British Medical Association Cymru (BMA)
- National Research Ethics Service (NRES)
- Public Health Wales
- NHS Wales Informatics Service (NWIS)
- Health Boards
- The Public

Your application has now been handed to our data provisioning team and we will inform you as soon as your data is available to you. The data provisioning process takes approximately 4 weeks. We will inform you when your data views have been created and are available for use.

Any new user who requires access to the data via the SAIL Gateway will need to apply for an account following the process outlined here:https://saildatabank.com/application-process/following-approval/#remote-access

As part of your permissions of use authors who use data from the SAIL Databank in any presentations, reports or other publications must acknowledge SAIL using the following wording *"This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers who make anonymised data available for research."* and should cite the relevant primary SAIL publications.

If you have any questions please contact https://help.saildatabank.com

Regards SAIL IGRP Team

Secure Anonymised Information Linkage (SAIL)

DATA ACCESS AGREEMENT

VERSION	9.0
CLASSIFICATION	Open
OWNER	Cynthia McNerney
DATE ISSUED	01/08/2021
STATUS	Final Version
REFERENCE	Data Access Agreement

THE SAIL DATABANK
DATA ACCESS AGREEMENT

This agreement governs the terms on which access will be granted to the data stored in the SAIL databank.

In signing this agreement, you are agreeing to be bound by the terms and conditions of access set out in this agreement.

The terms of access set out in this agreement apply both to the User and the User's Authorising Institution. User and User's Authorising Institution are referred to within the agreement as "You" and "Your Institution" and shall be construed accordingly.

Definitions:

Anonymous means data from which personal identifiers have been removed and that the removed identifiers are not held, in any form or place by the organisation holding the remaining data.

Authorising Institution is the organisation to which the user is affiliated or employed and which signs the data access agreement.

Data means all data made available from SAIL.

Data Provider means the organisation that has agreed to share data with SAIL.

Disciplinary action means a sanction applied by SAIL against a *User* approved under the terms of this Agreement

Information means any knowledge, insights or opinions that have been informed by the use of the SAIL data.

Publication means, without limitation, articles published in print journals, electronic journals, reviews, books, posters and other written and verbal presentations of research.

SAIL means the data linkage infrastructure and governance system housing the SAIL Databank; more details can be found on the SAIL website http://www.saildatabank.com.

SAIL Databank means the database that contains all the data obtained via the Secure Anonymised Information Linkage (SAIL) system[1].

User Institution means the organisation at which the User is employed, affiliated or enrolled.

[1] Ford DV, Jones KH, Verplancke JP, Lyons RA, John G, Brown G, Brooks CJ, Thompson S, Bodger O, Couch T, Leake K. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Health Services Research 2009;9:157 http://www.biomedcentral.com/1472-6963/9/

User means a researcher whose Authorising Institution has previously completed this Data Access Agreement and has received acknowledgement of its acceptance.

Terms and Conditions:

In signing this Agreement:

1. You agree to use the data only according to the study outline in the SAIL application form and understand that any amendments to this protocol must be passed through the independent Information Governance Review Panel (IGRP) for approval. Use of the data for any unapproved purposes is strictly prohibited and will result in disciplinary action.
2. You agree to the statement of the procedures, as given in the National Statistics Code of Practice: Protocol on Data Access and Confidentiality[2], adopted by Welsh Government to protect the confidentiality of personal data, and confirm that in any use made of this data, the User will follow these procedures and will adhere to the provisions of the General Data Protection Regulation 2016/679, the UK GDPR under the provisions of the Data Protection Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019 and the Data Protection Act 2018, both in letter and spirit, to the maximum extent that they apply. Breaches will be recorded on a risk log and reported to relevant external bodies. SAIL will work with all custodians to assist them enact any measures, including penalties, that apply.
3. Any disputes arising between the providing and beneficiary organisations will be resolved initially informally by reference to the principals to this agreement. Any outstanding issues will be referred to the Information Governance Review Panel, whose determination will be final in the matter.
4. You accept that the data may be protected by the General Data Protection Regulation 2016/679, the UK Data Protection Act 2018[3], the UK GDPR, the common law duty of confidentiality and the Human Rights Act 1998 and that you are responsible for ensuring compliance with any such applicable la. The Information Governance Review Panel reserves the right to request and inspect data security and management documentation and arrangements as required to ensure compliance.
5. You accept that access cannot be granted to anyone currently being investigated for any data protection contravention, or anyone who has been found to have been breach of any relevant policy or law.
6. This agreement shall be construed, interpreted and governed by the laws of England and Wales and shall be subject to the exclusive jurisdiction of the courts of England and Wales.
7. You agree to follow the principles

[2]http://doc.ukdataservice.ac.uk/doc/8097/mrdoc/pdf/8097_ons_protocol_for_data_access_and_confidentiality.pdf
[3] The Data Protection Act 2018. http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted

- *You will use the information entrusted to you for the public good - to improve health and health (and allied) services.*
- *You will acknowledge the origins of the information you use and will respect the rights of patients and their information guardians.*
- *You will only handle information that is anonymised to the point where an individual cannot be identified by information held by SAIL or any information available to SAIL.*
- *You will never make deliberate attempts to discover the identity of any individual from the information to which you have access.*
- *You will never make public the results of our analysis that might result in an individual, or small groups of individuals, being identified.*
- *You will abide by all relevant laws and codes of practice current at the time.*
- *You will only use SAIL data for genuine research for the public good, in line with your project scope agreed by the SAIL Information Governance Review Panel. You will never use SAIL data for commercial purposes (e.g. market research).*
- *You will treat the information you may view responsibly and take proper precautions with regard to the security of the information.*
- *You will conduct your analysis thoughtfully and only make public results that have been carefully judged to be fair and unbiased.*
- *You will use the results of our analysis fairly and equitably, and will not enter into campaigns or arguments that are not pursued solely in the public interest.*
- *You will try and help improve the quality of routinely collected health information, by passing on the lessons you learn.*
- *You will not allow the information to be used to embarrass, coerce or performance manage organisations or the departments or individuals that work within them.*

8. You agree to preserve, at all times, the confidentiality of information and data pertaining to any identifiable individual. In particular, you undertake not to use, or attempt to use the data to compromise or otherwise infringe the confidentiality of information relating to any individual and their right to privacy.
9. You will be personally liable for any breach of data protection laws or wider obligation of confidentiality where you have acted outside the remit of the relevant IGRP approval or the reasonable instructions of your User Institution.
10. You will consult SAIL in writing before taking any step that could put at risk the confidentiality or security of the data.
11. You will not attempt to match or link SAIL data to any data from another source that may lead to the disclosure of information about individual units, households or businesses.
12. You understand and acknowledge that use of the data granted under this Agreement should not be construed as conferring ownership of the data, which are protected by copyright and other intellectual property rights.

13. Unless conflicting factors are reported to SAIL, you agree to share all computer code, statistical scripts and similar material developed while using SAIL data with the SAIL user community, through the channels SAIL make available for this purpose
14. You understand that no data, results, or any product derived from the data, may be copied, by any means and for any purposes, from the SAIL Databank without it having first been submitted for approval by a designated SAIL member. Data must not be released as part of a study output that could identify an individual or could be used to identify an individual. This includes all statistical findings, summary figures, counts, percentages, presentations, publications, papers and analysis. The final decision in respect of release of outputs remains with SAIL such permission shall not be unreasonably withheld or denied. You will proactively bring to the attention of SAIL any outputs that you consider may identify an individual or individuals.
15. You are required to provide a copy of all final publications (e.g. papers, reports, abstracts and posters) to SAIL as soon as they become available in accordance with SAIL's Impact Assessment process. Details of publications and other outputs should be submitted via the SAIL portal by clicking the 'Update Outcomes' link beside your project title.
16. You agree to abide by the terms outlined in the SAIL Publications Policy along with any specific conditions prescribed by the data provider and the requirements as defined within your IGRP approval.
17. Any reference to the use of SAIL data in any publication, poster, or presentation prior to approval by the Information Governance Review Panel, must be clearly described as "Pending approval".
18. You agree to acknowledge SAIL in any published paper or presentation, which is based wholly or partly on the data; you agree to acknowledge the source of the data and the role of SAIL in making this available. A suitable wording is provided in the SAIL Publications Policy.
19. Your use of SAIL data is at your own risk in respect of data accuracy and completeness. You accept that SAIL, the original data creators or guardians, depositors or copyright holders, or the funders of the data or any part of the data supplied: a) bears no legal responsibility for the accuracy or comprehensiveness of the data; and b) accept no liability for indirect, consequential, or incidental, damages or losses arising from use of the data, or from the unavailability of, or break in access to, the data for whatever reason.
20. You agree that you will submit a report through the Information Governance Review Panel, if requested, on completion of the agreed purpose. The SAIL Information Governance Review Panel agrees to treat the report and all information, data, results, and conclusions contained within such report as confidential information belonging to the User Institution.
21. You agree that you and your collaborators will make any reasonable changes to the products of your work involving the SAIL data, as requested by the Information Governance Review

Panel. The SAIL Information Governance Review Panel will only request such changes if, in its opinion, your work infringes the spirit of the agreement or the principles of collaboration that SAIL has with its data provider organisations.

22. You accept that data will be refreshed from time to time, with notice given.
23. You accept that it may be necessary for SAIL or its appointed agent to alter the terms of this agreement from time to time. In this event, SAIL or its appointed agent will contact you to inform you of any changes.
24. You accept that this agreement will terminate immediately if you breach any term of this agreement. In this event, you will be required to destroy any analysis and products derived from this data and confirm that this has been done.
25. You accept that the user account created for you will be used by yourself only and no other individual. By logging on you reaffirm your agreement to uphold data confidentiality and security in terms of this Agreement.
26. You will take reasonable steps to ensure that, when accessing SAIL data, your display screen is not being overlooked by unauthorised persons and that you have taken all reasonable physical and other security precautions to maintain the security of the data supplied.
27. You agree that data from the SAIL Databank shall not be electronically copied, disseminated or distributed in its raw form.
28. You agree to ensure that your computer operating system, used to access the SAIL Gateway, is updated with the latest security patches, and that you will run reliable, effective and up-to-date anti-virus software. Failure to do this will result in your access privileges being revoked.
29. You agree that if you suspect that the availability, integrity or security of the SAIL system is compromised in any way you will immediately notify the helpdesk. You will also inform the helpdesk if you detect weakness in the SAIL system which could result in it being compromised. The SAIL Helpdesk can be reached via https://help.saildatabank.com/ or saildatabank@swansea.ac.uk.
30. You accept that all projects which have been IGRP approved and their users may be subject to an audit at any time to ensure that you are following SAIL policy, the specifications of the approved application and the conditions of this agreement. Such audit shall be at a reasonable interval of no more than once in any twelve month period.
31. You agree to demonstrate that you have satisfactorily completed an approved safe researcher or other relevant training course. You will re-attend training if expiry of your certificate occurs within the time period of your study.
32. Disciplinary Action - Offences and Penalties

SAIL reserves the right to suspend access to data if it considers that any User is perpetrating or attempting to perpetrate any breach of the terms of this Agreement.

SAIL retains discretionary powers over the application of penalties for self-reported breaches.

- The following are illustrative of the type of penalty that may be applied for a breach that is not self-reported using the service and/or data for unapproved commercial purposes – permanent suspension
- Infringing SAIL security protocols - first offence - 6 months suspension
- Transferring log in details to another user – more than one offence - permanent suspension

The above list is illustrative and not exhaustive. SAIL also reserves the right to report any action in breach of data protection laws to the Information Commissioners Office at its discretion. The User retains legal liability for any action that arises as a result of their actions that is in breach of this Agreement.

There shall be a right of appeal to a SAIL Director from the initial decision to apply a sanction. The decision of the SAIL director on appeal is final.

33. SAIL holds non identifiable data that is outside the remit of the Data Protection Act. However Section 171 of the Act provides for the offence of re-identification of de-identified data and s170 of the Data Protection Act 2018 states that the knowing or reckless obtaining or disclosure of personal data without the consent of the data controller is a criminal offence. If SAIL believes that a breach of relevant sections of the Act has occurred it shall report accordingly.

SAIL DATABANK
DATA ACCESS AGREEMENT

I have read the terms and conditions set out in the SAIL Data Access Agreement version 9.0 and agree to be bound by them. I declare that I am not currently being investigated under the Data Protection Act and have not been found to be in breach of the Act.

Name: Ms Brenda Bauer

Job Title: PhD Candidate

Organisation: University of St Andrews

Signature: 

Date: _____15 February 2022_____

Head of Department/Authorising Institution Approval

I confirm that I am happy for this applicant to have access to SAIL data and I am aware of the penalties arising from breaches of these terms and conditions.

Name: Colin McCowan

Job Title: Professor in Health Data Science

Organisation: University of St Andrews

Signature: Colin Mc Cowan

Date: _23/02/2022_____

t958EnREwh

Descriptive Study

S1 Appendix. Description of SAIL datasets

Demographic data, including residential periods in Wales, as well as indices of deprivation, are available from WDSD. All relevant GP records within the study timespan are obtained from WLGP – individual visits, investigations, diagnoses and prescribed medications per patient. Similarly, PEDW contains data on secondary care from all hospitals in Wales, precisely hospital admissions, diagnoses and interventions. Lab test result data across Wales are available from WRRS, regardless of whether the originator of the pathology request works in primary or secondary care. Outpatient referrals from primary care and the details of outpatient encounters are recorded in OPRD and OPDW, respectively. In EDDS, data are routinely collected on interactions with hospital emergency departments – this dataset was included to lessen the likelihood that patient outcomes would be missed. Deaths are recorded in multiple datasets, including WDSD, which contains demographic records and ADDE, a standalone register of all deaths in Wales. All EHR collected between 1 January 2000 and 31 December 2021 were subsequently provisioned.

S1 Table. Datasets provided by SAIL Databank for this study

SAIL Datasets	Description
Annual District Death Extract (ADDE)	Death records
Emergency Department Dataset (EDDS)	Accident & Emergency records
Outpatient Database for Wales (OPDW)	Outpatient attendance records
Outpatient Referrals Dataset (OPRD)	Outpatient referral records
Patient Episode Database for Wales (PEDW)	Hospital admission records
Welsh Demographic Service Dataset (WDSD)	Demographic data
Welsh Longitudinal General Practice (WLGP)	Primary care/GP records
Welsh Results Reports Service (WRRS)	Laboratory test records

S2 Appendix. Case definition (SCH patients)

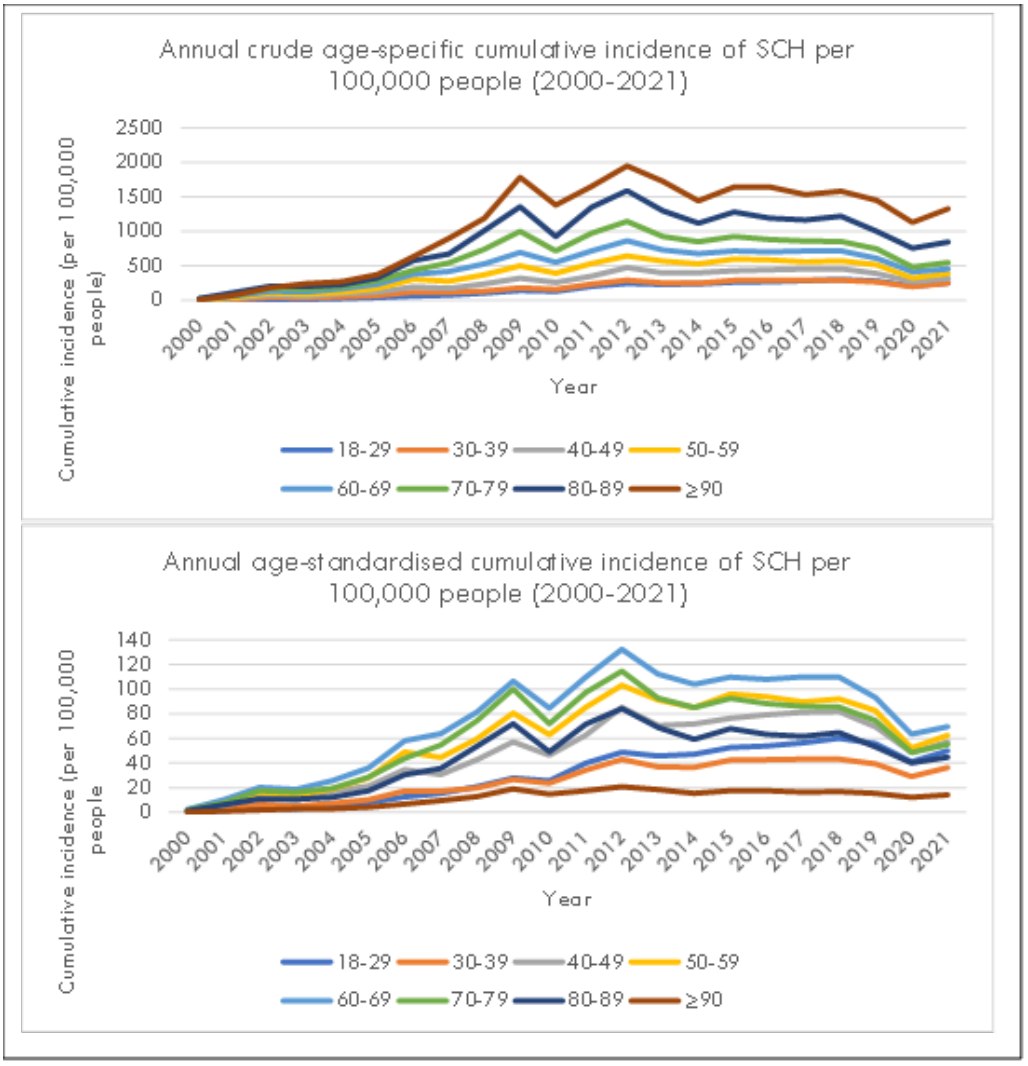
Eligible patients were identified using a combination of Read v2 and International Classification of Diseases version 10 (ICD-10) diagnostic codes in the WLGP and PEDW/OPDW datasets, respectively (Supplementary Table 2). Office for Population Censuses Surveys version 4 (OPCS-4) codes were also used in PEDW to identify interventions such as investigations and surgeries that patients had received.

S2 Table. Diagnostic codes used to identify patients with subclinical hypothyroidism

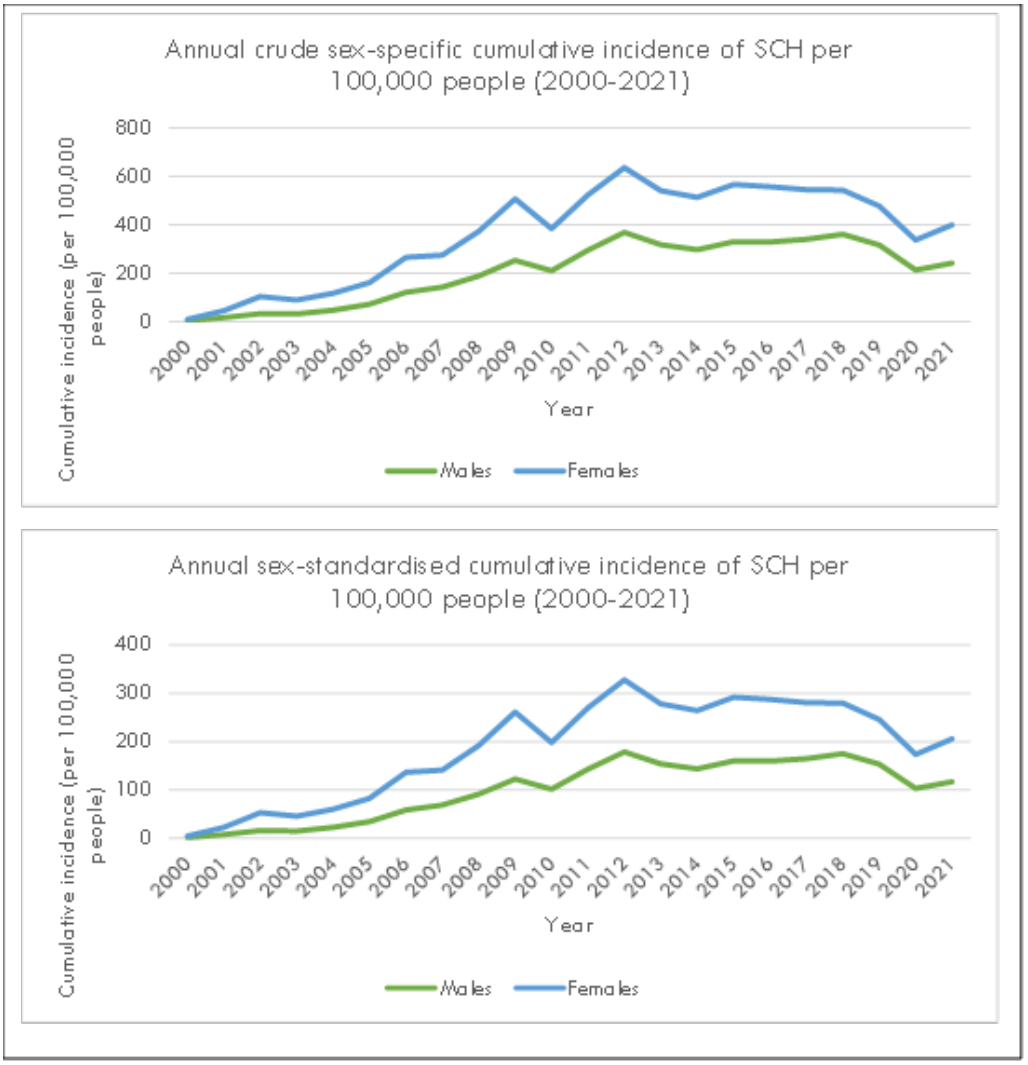
Code	Type	Description
C047.00	Read v2	Subclinical hypothyroidism
C0A5.00	Read v2	Subclinical iodine-deficiency hypothyroidism
E02X	ICD-10	Subclinical iodine-deficiency hypothyroidism
E038	ICD-10	Borderline hypothyroidism

Because SCH is diagnosed due to thyroid function tests, WRRS was also checked for patients meeting the criteria of high TSH and normal FT4 from blood specimens collected on the same day. The respective lab reference ranges were used, as recorded alongside the test results. The assumption was that these patients had tests indicative of SCH, regardless of whether their corresponding GP or hospital records contained the appropriate diagnostic codes. It was not possible, however, to include levothyroxine treatment as a criterion for case identification because its use is not restricted to SCH.

S3 Appendix. Age- and sex-stratified annual cumulative incidence according to the mid-2011 Welsh census data



Supplementary Fig 1. Age-standardised cumulative incidence of SCH over the study period (2000-2021)



Supplementary Fig 2. Sex- standardised cumulative incidence of SCH over the study period (2000-2021)

S4 Appendix. Levothyroxine use during the study period

S3 Table. Frequency counts of patients receiving levothyroxine over the study period (2000-2021)

Year	Cumulative total of SCH patients	Number of patients receiving LT4 prescriptions		
		Total (%)[a]	New users	Existing users
2000	200	52 (26.0)	52	0
2001	1,143	187 (16.4)	140	47
2002	3,235	773 (23.9)	590	183
2003	5,103	1,333 (26.1)	599	734
2004	7,669	2,126 (27.7)	858	1,268
2005	11,307	3,180 (28.1)	1,150	2,030
2006	17,392	4,673 (26.9)	1,640	3,033
2007	23,969	5,815 (24.3)	1,393	4,422
2008	32,892	7,302 (22.2)	1,780	5,522
2009	44,962	9,340 (20.8)	2,342	6,998
2010	54,396	11,160 (20.5)	2,272	8,888
2011	67,382	13,321 (19.8)	2,674	10,647
2012	83,353	16,014 (19.2)	3,274	12,740
2013	97,011	18,398 (19.0)	3,199	15,199
2014	109,942	20,775 (18.9)	3,253	17,522
2015	124,291	23,105 (18.6)	3,348	19,757
2016	138,509	25,549 (18.4)	3,535	22,014
2017	152,754	27,784 (18.2)	3,479	24,305
2018	167,333	29,764 (17.8)	3,374	26,390
2019	180,184	31,534 (17.5)	3,276	28,258
2020	189,103	32,185 (17.0)	2,239	29,946
2021	199,520	33,337 (16.7)	2,637	30,700

[a] Percentage of all SCH patients in the respective year; [b] Calculated as the number of prescriptions divided by the total number of patients that received them.

Abbreviations: *LT4* levothyroxine; *SCH* subclinical hypothyroidism.

Appendix III: Supplementary Material for Chapter 4