# Assessing the Cardiovascular Effects of Levothyroxine Use in an Ageing United Kingdom Population with Subclinical Hypothyroidism: Emulated Target Trial (ACEL-UK-ETT)

Mia Holley[1*], Salman Razvi[2], Ian Maxwell[1], Rosie Dew[1], Scott Wilkes[1]
1 School of Medicine, Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK
2 Translational and Clinical Research Institute, Newcastle University, Newcastle-upon-Tyne, UK
*Correspondence: mia.holley@sunderland.ac.uk

**Abstract**

**Background:** Thyrotropin levels increase with age, but standard reference intervals do not account for this, potentially leading to overdiagnosis of subclinical hypothyroidism (SCH) and overuse of levothyroxine in older adults.

**Methods:** Using data from The Health Improvement Network, this observational emulated target trial study assessed 10-year outcomes in adults over 50 years with SCH (thyrotropin 4·1-10·0mU/L, free thyroxine 10·0–24·0 pmol/L) who were prescribed levothyroxine versus those who were not. Subgroup analyses were limited to patients with age-specific thyrotropin levels. The primary outcome was cardiovascular events (angina, myocardial infarction, peripheral vascular disease, stent procedures, or stroke). Secondary outcomes included bone events (fragility fractures or osteoporosis) and all-cause mortality. Hazard ratios, adjusted through inverse probability of treatment weighting (IPTW) for age, sex assigned at birth, body mass index, Charlson comorbidity index, total cholesterol, hypertension, thyrotropin, hormonal medications, and smoking, were estimated.

**Findings:** Between January 1, 2006, and January 1, 2022, 22,621 patients (median age [IQR] 66 [59-75] years, 76·7% female) were identified; 62% received levothyroxine and 38% did not. Levothyroxine was associated with reduced cardiovascular (IPTW-adjusted hazard ratio (aHR) 0·82; 95% CI: 0·74-0·91; p<0·0001) and all-cause mortality (aHR: 0·71; CI: 0·67-0·75; p<0·0001), with no adverse effects on bone (aHR: 1·04; CI: 0·93-1·17; p=0·45). Cardiovascular benefits were limited to patients with thyrotropin levels above the age-specific range and after at least five years of treatment.

**Interpretation:** Long-term levothyroxine use in older adults with SCH was associated with lower long-term cardiovascular and all-cause mortality risks, with no significant harm to bone health. Age-specific thyrotropin levels should guide treatment decisions.

## Background

The interpretation of thyrotropin (TSH) levels in older adults is debated, as conventional reference ranges (0·4–4·0/4·5 mU/L) may not reflect age-related physiological changes. Studies show that TSH rises naturally with age, despite normal free thyroxine (fT4).[1–3] Longitudinal data suggests that the upper limit may reach 8·0 mU/L in nonagenarians.[4–6] This has led to recognition that age-adjusted thresholds may better define thyroid dysfunction in older populations.[3]

Large studies, including the Scottish Thyroid Epidemiology, Audit, and Research Study (TEARS) (n = 153,127) and the United States National Health and Nutrition Examination Survey (n = 16,533), show the 97·5th percentile of TSH increases with age, exceeding 5·9 mU/L in those >80 years.[1,3] TSH concentrations exceeding 10·0 mU/L have consistently been linked to increased cardiovascular risk,[7] while mild elevations (4·5–7·0 mU/L) have been associated with improved mobility and reduced mortality in older age groups.[8,9]

Given these findings, there is increasing debate about the appropriateness of initiating levothyroxine (LT4) in older adults with mildly elevated, age-specific TSH levels. The role of LT4 in the treatment of subclinical hypothyroidism (SCH), defined as elevated TSH with normal fT4, remains uncertain due to limited high-quality, long-term evidence.[10] In older adults, LT4 carries a recognised risk of overtreatment, which has been associated with increased incidence of cardiovascular events and fractures.[11,12] Despite this, LT4 is frequently prescribed in this population, often resulting in suppressed TSH levels.[13] Although European guidelines recommend LT4 initiation in adults with TSH concentrations exceeding 10 mU/L with symptoms,[14] real-world adherence to this is inconsistent.[13]

Evidence on the cardiovascular and skeletal effects of LT4 treatment in older adults with SCH remains limited, primarily due to a lack of adequately powered randomised controlled trials (RCTs). The Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism (TRUST) trial aimed to investigate cardiovascular, but was underpowered due to recruitment challenges and revised its primary outcome to quality of life.[15] An upcoming trial in China will evaluate LT4 in older adults with SCH using age-specific TSH reference ranges, with 254 participants and a 48-week follow-up.[16] Consequently, fundamental questions regarding the long-term impact of LT4 in this population remain unresolved, with no registered trials addressing this gap. A recent United Kingdom cohort study (n = 53,899) reported cardiovascular benefits of LT4 but increased risks to bone health and all-cause mortality.[17] Systematic reviews have also highlighted the absence of evidence on long-term cardiovascular and bone health outcomes in older adults with SCH over 50 years, emphasising the need for further research.[18,19]

To address this evidence gap, this emulated target trial was developed.[20] By emulating an RCT with observational primary care data,[21] this study aimed to provide more causally robust evidence than traditional observational designs. Notably, the ACEL UK ETT study uniquely evaluated LT4 treatment in individuals with age-specific mildly elevated TSH levels, as defined by the TEARS study, offering clearer insight into its effects in this group.

## Materials and Methods

*Study Design and Setting*

This study employed an emulated target trial design using observational data to replicate the randomisation and prospective follow-up of an RCT. The study drew upon data from The Health Improvement Network (THIN), which includes electronic healthcare records from approximately 6% of the UK population. THIN contains clinical information such as diagnoses, laboratory test results, and prescription data, enabling longitudinal tracking of patient health and treatments. This database has shown generalisability to the UK population.[22] The study adhered to the Hernan and Robins framework (Table 1) to ensure robustness, with the protocol published in November 2023·[20]

*Study Population*

Data were extracted from the THIN database for patients aged over 50 years on January 1, 2006, with at least one TSH reading exceeding 4·1 mU/L between January 1, 2006, and January 1, 2022. The inclusion and exclusion criteria were then applied to this dataset.

Inclusion criteria:

1. Adults registered on the THIN database aged 50 years or older as of January 1, 2006·
2. TSH level between 4·01 mU/L and 10·0 mU/L at baseline.
3. FT4 level between 10·0pmol/L and 24·0pmol/L at baseline.[23]

Exclusion criteria:

1. History of thyroid cancer, hyperthyroidism or pituitary disease.
2. Use of drugs affecting thyroid function (liothyronine, lithium, or amiodarone).
3. Previous thyroid surgery or radioiodine treatment.
4. Pre-existing cardiovascular conditions or bone disease, depending on the outcome assessment.

The population was subsequently divided into three cohorts based on the findings of the TEARS study's TSH levels (Table 2).[3]

1. Cohort 1: Patients with a mildly elevated TSH level.
2. Cohort 1a: Patients within Cohort 1 with an age-specific normal TSH level.
3. Cohort 1b: Patients within Cohort 1 with an age-specific mildly elevated TSH level.

*Treatment Strategies*

The two primary exposures in this emulated target trial were:

1. **LT4 treatment:** Participants in this group received LT4 at baseline.
2. **No treatment:** Participants in this group did not receive LT4 at baseline.

A trial emulation design using an intention-to-treat approach was employed to assess the effect of LT4 on the time to first outcome event or death. Patients who had not previously used LT4 were assigned to the treatment group at the time of their first LT4 prescription, and they simultaneously met the eligibility criteria. For patients in the comparison (no treatment) group, the index date was established when they met eligibility criteria, with follow-up starting at eligibility for both groups.

Patient follow-up occurred over a 10-year period, between January 1, 2006, and January 1, 2022. The follow-up period for each patient began at their respective index date and continued until one of the following occurred: outcome event, deregistration from the database, death, or completion of ten years of follow-up. There was no grace period or minimum follow-up period implemented in the study design.[24]

*Outcome Measures*

The primary outcome of this study was cardiovascular events, which included angina, myocardial infarction, peripheral vascular disease, stent procedures, and stroke. Secondary outcomes included bone health-related events (osteoporosis and fragility fractures) and all-cause mortality. Composite outcomes were identified using the International Classification of Diseases, 10th Edition (ICD-10) and Read Codes (Supplementary Data, Table 1). Mortality was determined based on the recorded date of death. The first occurrence of any outcome event was considered the event of interest for each patient.

*Statistical Analysis*

Baseline covariates were balanced between treatment groups using inverse probability of treatment weighting (IPTW). [25,26] Weights were calculated from logistic regression models predicting treatment assignment, then stabilised for overall exposure and truncated at the 1st and 99th percentiles to ensure robustness. Covariates included age, sex assigned at birth, body mass index (BMI), Charlson Comorbidity Index (CCI),[27] total cholesterol, hypertension, concomitant hormonal medications (oestrogen, testosterone), TSH, and smoking status. E-values were calculated to assess the impact of unmeasured confounding.[28] Ethnicity was excluded from the analysis due to >50% missing data (Supplementary Data, Table 2).

For raw outcome comparisons, frequencies and percentages were presented. The primary analysis used weighted Cox proportional hazard ratios (HRs) with 95% confidence intervals (CIs) to assess associations between LT4 exposure and outcomes. A significance level of 0·01 was selected, following a Bonferroni correction from an initial significance level of 0·05.[29] Kaplan-Meier survival curves visualised survival probability differences between treatment groups. Missing data were addressed using multiple imputations with predictive mean matching and polytomous regression, depending on variable type.

Sensitivity analyses used alternative upper limit TSH cutoffs (4·5 and 5·0 mU/L) to account for variations in levels and laboratory ranges.[23] Additionally, cholesterol was removed from IPTW models, given its potential role as a mediator between SCH and cardiovascular disease. Subgroup analyses examined age, sex assigned at birth, smoking status, and baseline fT4 (above or below the population median). All analyses were performed using R version 4·4·2. [30]

Based on the Kaplan-Meier curves suggesting time-varying cardiovascular effects, post-hoc analyses evaluated LT4 impacts over time: (1) piecewise Cox regression with a five-year split, (2) landmark analysis restricting follow-up to patients event-free at five years, and (3) restricted mean survival time (RMST) comparing event-free survival up to five years and between five and ten years. Stabilised IPTW weights were used; all analyses were exploratory and not pre-specified.

*Ethical Approval*

This study was approved by the University of Sunderland Research Ethics Group (ID 011081) and the THIN Scientific Review Committee (protocol 22-003). THIN database has Health Research Authority ethical approval for research (reference 20/SC/0011).

*Data Sharing*

The data that supports the findings of this study are available from THIN, a wholly owned subsidiary of Cegedim SA, which owns the proprietary rights to THIN data. Restrictions apply to the availability of these data, which were used under license for the current study and are not publicly available. Data are, however, available from the authors upon reasonable request and with the permission of THIN.

**Findings**

There were 282,036 initial patient records received from THIN. After applying the study criteria for cardiovascular outcomes, 264,281 patients were removed (Supplementary Data, Figure 1). There were 22,621 patients included in the analyses for Cohort 1, with 14,064 (62·2%) taking LT4 and 8,557 (37·8%) receiving no treatment. Cohort 1a comprised 4,951 patients, with 1,398 (28·2%) receiving treatment and 3,553 (71·2%) not. Cohort 1b comprised 17,670 patients, with 12,666 (71·7%) prescribed LT4 and 5,004 (28·3%) not. The median (interquartile range) follow-up for cardiovascular outcomes was 10·0 (10·0 – 10·0) years for all three groups. Among treated patients in Cohort 1, the median baseline and follow-up LT4 dose was 50 [IQR 25–50] mcg/day. At baseline, treated patients were more likely to have fT4 values below the cohort median compared with untreated patients across cohorts (Cohort 1: 57·5% vs 33·9%; Cohort 1a: 50·7% vs 32·2%; Cohort 1b: 58·2% vs 35·1%), with untreated patients more frequently having fT4 values above the median. The study population had a higher proportion of females than males, with the majority aged 61–70 years. Ethnicity was predominantly White, with a small representation of Black patients (0·5%). Baseline characteristics were balanced after weighting (Table 3), as presented by standardised mean differences less than 0.1 (Supplementary Data, Table 3).

For bone health outcomes, 256,984 patients were excluded based on the eligibility criteria, leaving 25,052 patients for weighting. Among these, 15,216 (60·7%) were prescribed LT4, while 9,836 (39·2%) were not, comprising Cohort 1. Cohort 1a included 5,816 patients, with 27·7% treated and 72·2% untreated. Cohort 1b consisted of 19,236 patients, of whom 70·7% were treated and 29·3% were untreated. All three groups had a median (interquartile range) follow-up time of 10·0 (8·9-10·0) years for bone health outcomes.
A total of 254,773 patients were excluded based on the trial's exclusion criteria for all-cause mortality outcomes, leaving 27,263 patients eligible for weighting. Of these, 60·9% were assigned to LT4, and 39·0% were not. In Cohort 1a, out of 6,371 patients, 28·1% were prescribed LT4, and 71·9% were not. In Cohort 1b, out of 20,892 patients, 71% were prescribed LT4, and 29% were not. All three groups had a median (interquartile range) follow-up time of 10·0 (10·0-10·0) years for all-cause mortality outcomes.

Cardiovascular outcomes: Across all three cohorts, the treatment group had lower cardiovascular event rates than the control group (10·1% vs 13.0% in Cohort 1, 13·1% vs 15.0% in Cohort 1a, and 9·8% vs 11·6% in Cohort 1b). Cohort 1 exhibited significantly reduced long-term cardiovascular risk when prescribed LT4 treatment (Table 4). Cohort 1a showed no association between LT4 treatment and cardiovascular outcomes. Treatment with LT4 in cohort 1b demonstrated a beneficial effect.

Bone health outcomes: Across all three groups, there was a similar proportion of adverse bone health outcomes, regardless of LT4 treatment. Cohort 1 revealed an IPTW-adjusted HR of 1·04 (95% CI 0·93-1·17; p = 0·45), indicating no increased risk of osteoporosis or fragility fractures when prescribed LT4. Cohort 1a and Cohort 1b displayed similar findings.

All-cause mortality: Cohort 1 and Cohort 1b had lower mortality in the treatment group compared to the control group (23·8% vs. 30.0% in Cohort 1 and 18·3% vs. 25.0% in Cohort 1b); Cohort 1a showed similar proportions (37·4% vs 36·5%) across both the LT4-treated and control groups. Cohort 1 indicated a significantly reduced all-cause mortality risk associated with using LT4, presented by IPTW-adjusted HR's. Cohorts 1a and 1b presented a similar association.

IPTW-adjusted HRs were also calculated for various subgroups (Supplementary Data, Table 4). These generally showed no significant associations across outcomes, likely due to limited sample sizes. An exception was observed for all-cause mortality, where LT4 use was consistently associated with a favourable outcome, other than Cohort 1a. Additionally, sensitivity analyses using varying TSH cutoffs (4·5 and 5·0 mU/L) yielded results consistent with the primary HRs (Supplementary Data, Table 5). A third sensitivity analysis, excluding cholesterol from the IPTW model, also produced consistent findings. E-values corresponding to the primary HRs (up to 1·94) suggest that substantial unmeasured confounding would be needed to explain the observed associations (Supplementary Data, Table 5).

The Kaplan-Meier plot for Cohort 1 suggests early and late effects of LT4 treatment, with non-significant cardiovascular outcomes until approximately 2,000 days (Figure 1). Beyond this, the treatment group showed a significantly higher survival rate. The Kaplan-Meier plot for Cohort 1b displayed similar findings. Conversely, the Kaplan-Meier plot for Cohort 1a illustrated no difference in the survival probability, regardless of LT4 status (Supplementary Data, Figure 2).

*Post-Hoc Analyses*

Visual inspection of Kaplan-Meier cardiovascular event-free survival curves (Figure 1) suggested differences between LT4-treated and untreated groups, varying over time, with a trend toward improved cardiovascular outcomes in the LT4 group after approximately five years. To assess this time-varying effect, we conducted piecewise Cox regression, landmark, and RMST analyses. Piecewise Cox regression showed a non-significant adverse effect of LT4 within the first five years (HR 1·08; 95% CI 0·94–1·25; p = 0·25), but a significant protective effect from five to ten years (HR 0·54; 95% CI 0·44–0·66; p < 0·001), consistent across Cohorts 1a and 1b (Supplementary Material, Table 7). Landmark analysis, restricted to patients event-free at five years, confirmed reduced cardiovascular risk beyond five years of LT4 treatment (Cohort 1, HR 0·58; 95% CI 0·50–0·68; p < 0·001, Cohort 1a, HR 0·59; 95% CI 0·43–0·80; p < 0·001, Cohort 1b, HR 0·58; 95% CI 0·49–0·69; p < 0·001). RMST analysis further supported these findings, withLT4 patients experiencing ~21 fewer event-free days in the first five years but gained ~24 event-free days between five and ten years. Together, these post-hoc analyses suggest a delayed but significant cardiovascular benefit of LT4 treatment.

## Discussion

*Summary of the Main Findings*

In this emulated target trial using UK primary care data, LT4 in older adults with SCH was associated with long-term reduced cardiovascular events and all-cause mortality. No cardiovascular benefit was observed among those with age-specific TSH values, suggesting that modest TSH elevations in older adults may reflect physiological ageing rather than clinically relevant thyroid dysfunction. Importantly, we also observed no adverse effects on bone health associated with LT4 in the ageing SCH population. Post-hoc analyses showed cardiovascular benefits of LT4 were not significant in the first five years of treatment but became substantial between five and ten years of follow-up, specifically in individuals treated with an age-specific elevated TSH level.

*Comparison with Other Literature*

This ETT confirms cardiovascular benefits of LT4 in older adults with SCH, reinforcing findings from a previous cohort study involving over 53,000 patients.[17] Notably, it demonstrates no adverse effects on bone health and a significant reduction in all-cause mortality. These differences may be attributed to the superior methodological design of the ETT, which effectively emulates a hypothetical RCT within a large dataset. In contrast, the TRUST trial, a RCT involving just 737 participants, lacked statistical power to detect meaningful cardiovascular outcomes and failed to demonstrate significant benefits. This underscores the limitations inherent in clinical trials, specifically on ageing SCH patients. A recently published systematic review further emphasises this study's relevance. The review revealed inconsistent evidence regarding long-term cardiovascular and bone health effects of LT4 due to variations in sample sizes and follow-up periods.[18] Through a large sample size, decade-long follow-up, and rigorous methods, this ETT provides the most comprehensive real-world evidence to date regarding LT4 outcomes in an ageing SCH population.

*Strengths and Weaknesses*

This study has notable strengths. To our knowledge, it is the most extensive and methodologically robust study to evaluate the long-term effects of LT4 in ageing adults with SCH. With over 22,000 patients included and a 10-year follow-up, the study's size and duration considerably enhance the validity and reliability of its findings.[21] The ETT design, combined with IPTW and adjustment for multiple clinically relevant confounders,

enhances confidence in causal relationships.[26] Importantly, multiple sensitivity analyses confirmed the primary results and provided plausible explanations for the observed effects. While both this ETT and our prior cohort study demonstrated cardiovascular benefits of LT4, only the ETT found no adverse bone effects and a reduction in mortality risk.[17] These differences likely reflect the methodological superiority of the ETT, emulating a hypothetical gold-standard RCT. Fundamentally, by utilising existing big data, this trial addresses the cost, recruitment, and retention challenges that have limited prior RCTs, as highlighted in the recent systematic review.[18]

There are several limitations to consider. Cause-of-death data were unavailable, which limited the interpretation of mortality outcomes and potentially introduced selection bias. Although IPTW was employed, residual confounding from unmeasured variables remains possible. Additionally, many patients had only a single recorded TSH measurement and one LT4 prescription, restricting our ability to evaluate treatment adherence, LT4 dose, and thyroid function control; however, this reflects the intention-to-treat design of the target trial emulation. Defining SCH based on a single baseline TSH may have introduced misclassification bias, as some individuals may have had transient TSH elevations that would normalise on repeat testing. The absence of ethnic data further limited subgroup analyses. Lastly, while the THIN database is representative of the UK population, the findings may not be generalisable to other healthcare settings or populations with different characteristics.

*Implications for Practice and Research*

To our knowledge, this is the first long-term follow-up study describing the causal effects of LT4 in older adults with SCH. As an observational emulated target trial, these findings may be affected by residual confounding and health-care–seeking behaviour and should be interpreted with caution. RCTs in this population are unlikely, as demonstrated by recruitment challenges in prior studies, including the underpowered TRUST trial, which enrolled only 737 patients.[15] In contrast, this study represents the largest observational emulated target trial to date, involving over 22,000 patients. Given practical and financial barriers to conducting a trial of this scale, high-quality emulated trials currently provide the most reliable available evidence. Our findings suggest clinicians and guideline authors should consider the routine LT4 use in older adults with age-specific TSH levels and a selective, risk-based treatment approach. Using standard TSH reference ranges risks over-diagnosis and unnecessary intervention in adults over 50 years. Future research will likely rely on emulated target trials with large, generalisable population datasets.

**Author Contribution**

Mia Holley: writing – original draft (lead); methodology (supporting); data curation (lead), formal analysis (lead); writing – review and editing (equal). Salman Razvi: Conceptualisation (supporting); methodology (supporting); formal analysis (supporting); writing – review and editing (equal). Ian Maxwell: methodology (supporting); formal analysis (supporting); writing – review and editing (equal). Rosie Dew: methodology (supporting); formal analysis (lead); writing – review and editing (equal). Scott Wilkes: Conceptualisation (lead); funding acquisition (lead); methodology (lead); formal analysis (lead); writing – review and editing (equal).

**Author Disclosure Statement**

Salman Razvi has received speaker fees from Meck KGaA and Berlin Chemie AG, makers of Levothyroxine. All other authors declare that there were no conflicts of interest.

**Funding Statement**

This study was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration (ARC) North East and North Cumbria (NENC) (NIHR200173). Mia Holley was funded by National Institute for Health and Care Research Applied Research Collaboration North East and North Cumbria (NIHR200173). Salman Razvi, Ian Maxwell, Rosie Dew, and Scott Wilkes received no specific funding for this work.

**List of Abbreviations**

| | |
|---|---|
| ACEL-UK-ETT | Assessing the Cardiovascular Effects of Levothyroxine Use in an Ageing United Kingdom Population: an Emulated Target Trial |
| aHR | Adjusted hazard ratio |
| ARC | Applied Research Collaboration |
| BMI | Body mass index |
| CCI | Charlson comorbidity index |
| CI | Confidence interval |
| fT4 | Free thyroxine |
| HR | Hazard ratio |
| ICD-10 | International Classification of Diseases, 10th Edition |
| IPTW | Inverse probability of treatment weighting |
| LT4 | Levothyroxine |
| NIHR | National Institute for Health and Care Research |
| RCT | Randomised controlled trial |
| RMST | Restricted Mean Survival Time |
| SCH | Subclinical hypothyroidism |
| TEARS | Thyroid Epidemiology, Audit, and Research Study |
| THIN | The Health Improvement Network |
| TRUST | Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism |
| TSH | Thyrotropin |
| UK | United Kingdom |

**References**


1. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). J Clin Endocrinol Metab 2002;87(2):489–499; doi: 10.1210/jcem.87.2.8182.

2. Parle JV, Franklyn JA, Cross KW, et al. Prevalence and follow-up of abnormal thyrotrophin (TSH) concentrations in the elderly in the United Kingdom. Clin Endocrinol (Oxf) 1991;34(1):77–83; doi: 10.1111/j.1365-2265.1991.tb01739.x.

3. Vadiveloo T, Donnan PT, Murphy MJ, et al. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: the Thyroid Epidemiology, Audit, and Research Study (TEARS). J Clin Endocrinol Metab 2013;98(3):1147–1153; doi: 10.1210/jc.2012-3191.

4. Waring AC, Arnold AM, Newman AB, et al. Longitudinal changes in thyroid function in the oldest old and survival: the cardiovascular health study all-stars study. J Clin Endocrinol Metab 2012;97(11):3944–3950; doi: 10.1210/jc.2012-2481.

5. Bremner AP, Feddema P, Leedman PJ, et al. Age-related changes in thyroid function: a longitudinal study of a community-based cohort. J Clin Endocrinol Metab 2012;97(5):1554–1562; doi: 10.1210/jc.2011-3020.

6. Chaker L, Korevaar TIM, Medici M, et al. Thyroid Function Characteristics and Determinants: The Rotterdam Study. Thyroid® 2016;26(9):1195–1204; doi: 10.1089/thy.2016.0133.

7. Gencer B, Collet T-H, Virgini V, et al. Subclinical thyroid dysfunction and cardiovascular outcomes among prospective cohort studies. Endocr Metab Immune Disord Drug Targets 2013;13(1):4–12; doi: 10.2174/1871530311313010003.

8. Simonsick EM, Newman AB, Ferrucci L, et al. Subclinical Hypothyroidism and Functional Mobility in Older Adults. Arch Intern Med 2009;169(21):2011–2017; doi: 10.1001/archinternmed.2009.392.

9. Rodondi N, Newman AB, Vittinghoff E, et al. Subclinical hypothyroidism and the risk of heart failure, other cardiovascular events, and death. Arch Intern Med 2005;165(21):2460–2466; doi: 10.1001/archinte.165.21.2460.

10. Biondi B, Wartofsky L. Treatment with thyroid hormone. Endocr Rev 2014;35(3):433–512; doi: 10.1210/er.2013-1083.

11. Biondi B, Cooper DS. The Clinical Significance of Subclinical Thyroid Dysfunction. Endocr Rev 2008;29(1):76–131; doi: 10.1210/er.2006-0043.

12. Flynn RW, Bonellie SR, Jung RT, et al. 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(1):186–193; doi: 10.1210/jc.2009-1625.

13. Taylor PN, Iqbal A, Minassian C, et al. Falling threshold for treatment of borderline elevated thyrotropin levels-balancing benefits and risks: evidence from a large community-based study. JAMA Intern Med 2014;174(1):32–39; doi: 10.1001/jamainternmed.2013.11312.

14. Pearce SHS, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. Eur Thyroid J 2013;2(4):215–228; doi: 10.1159/000356507.

15. Stott D, Rodondi N, Kearney P, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. N Engl J Med 2017;376(26):2534–2544; doi: 10.1056/NEJMoa1603825.

16. Xu C, Zhou H, Lu X, et al. Efficacy and safety of levothyroxine monotherapy in lowering the risk of cardiovascular disease in older adults with subclinical hypothyroidism: research protocols of a multicenter, open-label, randomized controlled trial. Trials 2025;26(1):161; doi: 10.1186/s13063-025-08857-z.

17. Holley M, Razvi S, Maxwell I, et al. Assessing the Cardiovascular Effects of Levothyroxine Use in an Ageing United Kingdom Population (ACEL-UK): Cohort Study. J Clin Endocrinol Metab 2025;dgaf208; doi: 10.1210/clinem/dgaf208.

18. Holley M, Razvi S, Farooq MS, et al. Cardiovascular and bone health outcomes in older people with subclinical hypothyroidism treated with levothyroxine: a systematic review and meta-analysis. Syst Rev 2024;13(1):123; doi: 10.1186/s13643-024-02548-7.

19. Bauer BS, Azcoaga-Lorenzo A, Agrawal U, et al. The impact of the management strategies for patients with subclinical hypothyroidism on long-term clinical outcomes: An umbrella review. PloS One 2022;17(5):e0268070; doi: 10.1371/journal.pone.0268070.

20. Holley M, Razvi S, Dew R, et al. Assessing the cardiovascular effects of levothyroxine use in an ageing United Kingdom population (ACEL-UK) protocol: a cohort and target trial emulation study. Thyroid Res 2023;16:43; doi: 10.1186/s13044-023-00186-0.

21. Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. JAMA 2022;328(24):2446–2447; doi: 10.1001/jama.2022.21383.

22. Lewis JD, Schinnar R, Bilker WB, et al. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiol Drug Saf 2007;16(4):393–401; doi: 10.1002/pds.1335.

23. Razvi S, Jabbar A, Addison C, et al. Variation in the reference range limits of thyroid function tests and association with the prevalence of levothyroxine treatment. Eur J Endocrinol 2023;lvad016; doi: 10.1093/ejendo/lvad016.

24. Hernán MA, Sterne JAC, Higgins JPT, et al. A Structural Description of Biases That Generate Immortal Time. Epidemiology 2025;36(1):107–114; doi: 10.1097/EDE.0000000000001808.

25. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016;183(8):758–764; doi: 10.1093/aje/kwv254.

26. Hernan MA, Robins JM. Causal Inference: What If. n.d.

27. Quan H, Sundararajan V, Halfon P, et al. Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data. Med Care 2005;43(11):1130; doi: 10.1097/01.mlr.0000182534.19832.83.

28. Chung WT, Chung KC. The use of the E-value for sensitivity analysis. J Clin Epidemiol 2023;163:92–94; doi: 10.1016/j.jclinepi.2023.09.014.

29. Bland M. An Introduction to Medical Statistics. 4th edition. Oup Oxford; 2015.

30. R Core Team. R: A Language and Environment for Statistical Computing. 2024.

**Table 1 Overview of the target trial and trial emulation study protocols to estimate the effect of levothyroxine on cardiovascular, bone health and mortality outcomes.**[20]

| Protocol Components | Target Trial (ideal randomised controlled trial) | Trial Emulation using observational data |
|---|---|---|
| Eligibility criteria | • Adults aged >50 years and under 101 years on January 1, 2006·<br>• Patients with a baseline thyrotropin level between 4·01 mU/L and 10·0 mU/L.<br>• Patients with a baseline free thyroxine level between 10·0pmol/L and 24·0pmol/L.<br>• No previous thyroid cancer, pituitary disease, radioiodine treatment, thyroid surgery or hyperthyroidism history.<br>• No use of drugs that affect thyroid function (liothyronine, amiodarone, or lithium).<br>• No angina, myocardial infarction, peripheral vascular disease, stent, or stroke history (*cardiovascular outcome only*).<br>• No osteoporosis or fragility fracture history (*bone health outcome only*). | Same as the target trial. |
| Treatment strategies | 1. Patients who initiated levothyroxine.<br>2. Patients who initiated the placebo. | 1. Patients who initiated levothyroxine.<br>2. Patients who did not initiate levothyroxine. |
| Assignment procedures | Participants were randomly assigned to either strategy at baseline and were unaware of the strategy to which they were assigned. | Randomisation was emulated via inverse probability of treatment weighting (IPTW), using baseline confounders. |
| Follow-up period | Started at randomisation and ended at death, outcome event, administrative censoring, or ten years after follow-up began. | Similar: Started at assignment to treatment strategy and ended at death, outcome event, administrative censoring, or ten years after trial inception. |
| Outcome | Cardiovascular, bone health, or mortality event within ten years of baseline. | Same as the target trial. |
| Causal contrasts of interest | Intention-to-treat effect. | Observational analogue of the intention-to-treat principle. |
| Analysis plan | Cox proportional hazard models (one for each outcome event type: cardiovascular, bone health, and all-cause mortality). | Same as the target trial, except that Cox proportional hazard models were weighted through IPTW. |

**Table 2 Categorisation of patient cohorts in the study, split by age and thyrotropin (TSH) levels.**

| Age, y | Cohort 1 TSH level (mU/L) | Cohort 1a TSH level (mU/L) | Cohort 1b TSH level (mU/L) |
|---|---|---|---|
| 51-60 | [4·01–10·00] | [4·01-4·40] | [4·41-10·00] |
| 61-70 | [4·01–10·00] | [4·01-4·60] | [4·61-10·00] |
| 71-80 | [4·01–10·00] | [4·01-5·00] | [5·01-10·00] |
| 81-90 | [4·01–10·00] | [4·01-5·50] | [5·51-10·00] |
| 91-100 | [4·01–10·00] | [4·01-5·90] | [5·91-10·00] |

493 **Table 3 Baseline characteristics of participants.**

| Characteristic, N (%) | Cohort 1, Treatment (n = 14,064) | Cohort 1, Control (n = 8,557) | Cohort 1a, Treatment (n = 1,398) | Cohort 1a, Control (n = 3,553) | Cohort 1b, Treatment (n = 12,666) | Cohort 1b, Control (n = 5,004) |
|---|---|---|---|---|---|---|
| **Sex Assigned at Birth** | | | | | | |
| Female | 10793 (76·7) | 6086 (71·1) | 1106 (79·1) | 2537 (71·4) | 9687 (76·5) | 3549 (70·9) |
| Male | 3271 (23·3) | 2471 (28·9) | 292 (20·9) | 1016 (28·6) | 2979 (23·5) | 1455 (29·1) |
| **Age (years)** | | | | | | |
| 51-60 | 4085 (29) | 2710 (31·7) | 193 (13·8) | 683 (19·2) | 3892 (30·7) | 2027 (40·5) |
| 61-70 | 4696 (33·4) | 2797 (32·7) | 335 (24) | 1059 (29·8) | 4361 (34·4) | 1738 (34·7) |
| 71-80 | 3387 (24·1) | 2303 (26·9) | 435 (31·1) | 1286 (36·2) | 2952 (23·3) | 1017 (20·3) |
| 81-90 | 1654 (11·8) | 689 (8·1) | 369 (26·4) | 482 (13·6) | 1285 (10·1) | 207 (4·1) |
| 91+ | 242 (1·7) | 58 (0·7) | 66 (4·7) | 43 (1·2) | 176 (1·4) | 15 (0·3) |
| Median [lower quartile, upper quartile] | 66 [59, 75] | 66 [58, 74] | 74 [65, 82] | 71 [62, 77] | 66 [59, 74] | 63 [57, 70] |
| **Ethnicity** | | | | | | |
| Asian | 434 (3·1) | 240 (2·8) | 36 (2·6) | 99 (2·8) | 398 (3·1) | 141 (2·8) |
| Black | 68 (0·5) | 39 (0·5) | 7 (0·5) | 12 (0·3) | 61 (0·5) | 27 (0·5) |
| Mixed | 4572 (32·5) | 2825 (33) | 432 (30·9) | 1130 (31·8) | 4140 (32·7) | 1695 (33·9) |
| Other | 132 (0·9) | 77 (0·9) | 21 (1·5) | 34 (1) | 111 (0·9) | 43 (0·9) |
| White | 8858 (63) | 5376 (62·8) | 902 (64·5) | 2278 (64·1) | 7956 (62·8) | 3098 (61·9) |
| **Location** | | | | | | |
| London | 1000 (7·1) | 571 (6·7) | 83 (5·9) | 209 (5·9) | 917 (7·2) | 362 (7·2) |
| Midlands and East | 2364 (16·8) | 1256 (14·7) | 167 (11·9) | 520 (14·6) | 2197 (17·3) | 736 (14·7) |
| North | 2369 (16·8) | 1943 (22·7) | 195 (13·9) | 782 (22) | 2174 (17·2) | 1161 (23·2) |
| Northern Ireland | 957 (6·8) | 381 (4·5) | 110 (7·9) | 185 (5·2) | 847 (6·7) | 196 (3·9) |
| Scotland | 1609 (11·4) | 805 (9·4) | 200 (14·3) | 362 (10·2) | 1409 (11·1) | 443 (8·9) |
| South | 3518 (25) | 1879 (22) | 358 (25·6) | 688 (19·4) | 3160 (24·9) | 1191 (23·8) |
| Wales | 2247 (16) | 1722 (20·1) | 285 (20·4) | 807 (22·7) | 1962 (15·5) | 915 (18·3) |
| **Smoker status** | | | | | | |
| Smoker | 1065 (7·6) | 601 (7) | 86 (6·2) | 228 (6·4) | 979 (7·7) | 373 (7·5) |
| Past smoker | 4491 (31·9) | 2654 (31) | 439 (31·4) | 1095 (30·8) | 4052 (32) | 1559 (31·2) |
| Non-smoker | 8508 (60·5) | 5302 (62) | 873 (62·4) | 2230 (62·8) | 7635 (60·3) | 3072 (61·4) |
| **Comorbidities** | | | | | | |
| Asthma | 1328 (9·4) | 647 (7·6) | 112 (8) | 242 (6·8) | 1216 (9·6) | 405 (8·1) |
| BMI, median [lower quartile, upper quartile] | 28·1 [24·9, 32·2] | 28·1 [24·8, 32·1] | 27·7 [24·6, 32·0] | 27·9 [24·7, 31·8] | 28·1 [24·9, 32·2] | 28·1 [24·8, 32·3] |
| Chronic kidney disease | 1621 (11·5) | 36 (0·4) | 235 (16·8) | 15 (0·4) | 1386 (10·9) | 21 (0·4) |
| Chronic obstructive pulmonary disease | 1978 (14·1) | 1020 (11·9) | 170 (12·2) | 403 (11·3) | 1808 (14·3) | 617 (12·3) |
| Dementia | 160 (1·1) | 17 (0·2) | 37 (2·6) | 14 (0·4) | 123 (1) | 3 (0·1) |
| Depression | 2032 (14·4) | 846 (9·9) | 162 (11·6) | 326 (9·2) | 1870 (14·8) | 520 (10·4) |
| Diabetes | 168 (1·2) | 178 (2·1) | 19 (1·4) | 75 (2·1) | 149 (1·2) | 103 (2·1) |
| Dyslipidaemia | 1589 (11·3) | 681 (8) | 175 (12·5) | 283 (8) | 1414 (11·2) | 398 (8) |
| Hormones | 201 (1·4) | 104 (1·2) | 13 (0·9) | 44 (1·2) | 188 (1·5) | 60 (1·2) |
| Hypertension | 7071 (50·3) | 4299 (50·2) | 796 (56·9) | 1941 (54·6) | 6275 (49·5) | 2358 (47·1) |
| Rheumatoid arthritis | 445 (3·2) | 205 (2·4) | 57 (4·1) | 96 (2·7) | 388 (3·1) | 109 (2·2) |
| **Thyroid Hormone Levels** | | | | | | |
| Low-normal free thyroxine | 8081 (57·5) | 2900 (33·9) | 709 (50·7) | 1143 (32·2) | 7372 (58·2) | 1757 (35·1) |
| High-normal free thyroxine | 5983 (42·5) | 5657 (66·1) | 689 (49·3) | 2410 (67·8) | 5294 (41·8) | 3247 (64·9) |

| Characteristic, N (%) | Cohort 1, Treatment (n = 14,064) | Cohort 1, Control (n = 8,557) | Cohort 1a, Treatment (n = 1,398) | Cohort 1a, Control (n = 3,553) | Cohort 1b, Treatment (n = 12,666) | Cohort 1b, Control (n = 5,004) |
|---|---|---|---|---|---|---|
| TSH, median [lower quartile, upper quartile] | 6·5 [5·5, 7·9] | 4·8 [4·4, 5·6] | 4·5 [4·3, 4·8] | 4·3 [4·2, 4·5] | 6·8 [5·8, 8·1] | 5·4 [4·9, 6·1] |

*Cohort 1: Patients aged over 50 years with a thyrotropin (TSH) level between 4·1 mU/L and 10·0 mU/L and a normal free thyroxine (fT4) level.*
*Cohort 1a: Patients aged over 50 years with a thyrotropin level between 4·1 mU/L and the age-specific upper limit and a normal fT4 level.*
*Cohort 1b: Patients aged over 50 years with a thyrotropin level between the age-specific upper limit and 10·0 mU/L and a normal fT4 level.*
*Low-normal and high-normal fT4 levels are defined as above or below the median baseline fT4 level.*

**Table 4 Outcomes of the 10-year follow-up cohort study represented by raw numbers and IPTW-adjusted hazard ratios (aHR), 95% confidence intervals (CI) and p-values.**

| Outcome | Cohort 1 | | | Cohort 1a | | | Cohort 1b | | |
|---|---|---|---|---|---|---|---|---|---|
| | N (%), Treatment | N (%), Control | aHR (95% CI), p-value | N (%), Treatment | N (%), Control | aHR (95% CI), p-value | N (%), Treatment | N (%), Control | aHR (95% CI), p-value |
| **Cardiovascular** | 1426 (10·1) | 1114 (13.0) | 0·82 (0·74, 0·91), p < 0·0001 | 183 (13·1) | 532 (15.0) | 0·93 (0·78, 1·11), p = 0·42 | 1243 (9·8) | 582 (11·6) | 0·85 (0·75, 0·96), p = 0·011 |
| **Bone Health** | 1286 (8·5) | 841 (8·6) | 1·04 (0·93, 1·17), p = 0·45 | 185 (11·5) | 419 (10.0) | 1·14 (0·95, 1·37), p = 0·15 | 1101 (8·1) | 422 (7·5) | 1·03 (0·89, 1·18), p = 0·71 |
| **All-Cause Mortality** | 3951 (23·8) | 3189 (30.0) | 0·71 (0·67, 0·75), p < 0·0001 | 669 (37·4) | 1671 (36·5) | 0·8 (0·72, 0·89), p < 0·0001 | 981 (18·3) | 1343 (25.0) | 0·68 (0·63, 0·73), p < 0·0001 |

*Cohort 1: Patients aged over 50 years with a thyrotropin (TSH) level between 4·1 mU/L and 10·0 mU/L and a normal free thyroxine (fT4) level.*
*Cohort 1a: Patients aged over 50 years with a TSH level between 4·1 mU/L and the age-specific upper limit and a normal fT4 level.*
*Cohort 1b: Patients aged over 50 years with a TSH level between the age-specific upper limit and 10·0 mU/L and a fT4 level.*
Hazard ratios were adjusted through inverse probability of treatment weighting (IPTW) for age, sex assigned at birth, body mass index, Charlson comorbidity index, total cholesterol, hypertension, thyrotropin, hormonal medications, and smoking.

**Figure 1 Kaplan-Meier plots illustrating survival probabilities over time for cardiovascular outcomes.**
*Cohort 1: Patients aged over 50 years with a thyrotropin level between 4·1 mU/L and 10·0 mU/L and a normal free thyroxine level.*