Weight-based levothyroxine dosing and long-term cardiovascular, bone, and mortality outcomes in older adults: an emulated target trial using UK primary care data

Mia Holley, Scott Wilkes, Ellen Dorothea Moss, Ian Maxwell, Rosie Dew, Salman Razvi

Summary

Background Levothyroxine is widely prescribed in older adults with hypothyroidism, yet evidence guiding safe initial weight-based dosing is scarce. Both under-replacement and over-replacement of levothyroxine might adversely affect cardiovascular health, bone integrity, and survival. Given the impracticality of conducting a long-term randomised trial in this population, we aimed to emulate a target trial to assess associations between initial levothyroxine dose per bodyweight and long-term clinical outcomes in older adults using routinely collected primary care data.

Methods This emulated target trial was based on observational data from The Health Improvement Network, a UK primary care database, from Jan 1, 2006, to Dec 31, 2021. Eligible participants were adults older than 50 years diagnosed with hypothyroidism with at least one recorded thyroid-stimulating hormone measurement exceeding 4·0 mU/L who initiated levothyroxine during this period. The target trial compared initiation of levothyroxine at greater than 1·0 µg/kg per day (high dose) versus 1·0 µg/kg per day or less (low dose) using total bodyweight and lean bodyweight definitions. Participants were followed up for up to 10 years from levothyroxine initiation until outcome occurrence, death, or study end. Inverse probability of treatment weighting adjusted for baseline confounders. Associations with cardiovascular events, bone outcomes, and all-cause mortality were estimated as primary outcomes using weighted Cox proportional hazards models, and 10-year absolute risk differences were calculated in the pseudo-weighted population. Outcomes were analysed according to baseline prescribed dose group. Each of the primary outcomes was analysed separately, using bespoke cohorts.

Findings Overall, 19 027 records were eligible for inclusion in this study (14 296 [75%] were for female individuals and 4731 [25%] were for male individuals). 15 463 adults were included in the primary cardiovascular analysis, 11 762 (76%) of whom initiated levothyroxine at a dose of 1·0 µg/kg per day or less and 3701 (24%) at a dose of more than 1·0 µg/kg per day. 14 947 adults were included in the bone health analysis, 11 988 (80%) of whom initiated levothyroxine at a dose of 1·0 µg/kg per day or less and 2959 (20%) at a dose of more than 1·0 µg/kg per day. All 19 027 individuals were included in the all-cause mortality analysis, 14 503 (76%) of whom initiated levothyroxine at a dose of 1·0 µg/kg per day or less and 4524 (24%) at a dose of more than 1·0 µg/kg per day. Unweighted event rates were higher in the high-dose group than the low-dose group for cardiovascular events (728 [20%] of 3701 *vs* 2079 [18%] of 11 762), bone outcomes (387 [13%] of 2959 *vs* 1089 [9%] of 11 988), and all-cause mortality (836 [18%] of 4524 *vs* 2532 [17%] of 14 503). High levothyroxine dose was associated with increased risk of cardiovascular events (adjusted hazard ratio 1·18, 95% CI 1·08–1·29), bone outcomes (1·29, 1·16–1·43), and all-cause mortality (1·17, 1·09–1·25). The weighted 10-year absolute risk difference between the high-dose group and the low-dose group was 1·3% (95% CI 0·3–3·0) for cardiovascular events, 3·0% (1·6–4·5) for bone outcomes, and 3·4% (1·7–5·1) for all-cause mortality for total bodyweight dosing. Similar associations were observed across outcomes and analyses using lean bodyweight-based dosing.

Interpretation Our findings suggest that high initial weight-based levothyroxine dosing in older adults is associated with less favourable long-term outcomes, supporting cautious initiation and titration in this population. Our study is based on observational data and findings are associations, so should be interpreted with caution and validated with prospective studies.

Funding National Institute for Health and Care Research Applied Research Collaboration North East and North Cumbria.

Lancet Prim Care 2026; 2: 100194

Published Online September 1, 2026
https://doi.org/10.1016/j.lanprc.2026.100194

School of Medicine, University of Sunderland, Sunderland, UK (M Holley PhD, Prof S Wilkes PhD, I Maxwell MSc, R Dew PhD); **Population Health Sciences Institute** (E D Moss PhD) **and Translational and Clinical Research Institute** (S Razvi MD), **Newcastle University, Newcastle upon Tyne, UK**

Correspondence to: Dr Mia Holley, School of Medicine, University of Sunderland, Sunderland SR1 3SD, UK
mia.holley@sunderland.ac.uk

Introduction

Hypothyroidism is a common chronic endocrine disorder characterised by insufficient production of thyroid hormone, particularly thyroxine.[1] Its prevalence increases with age and is notably higher in women, affecting up to 21% of those older than 74 years.[2] This higher prevalence in

Research in context

Evidence before this study

We searched PubMed from Jan 1, 2000, to Oct 1, 2025, using combinations of terms for levothyroxine ("levothyroxine", "thyroxine", and "LT4"), weight-based dosing ("dose per body weight", "weight-based dose", "mcg/kg", and "µg/kg"), hypothyroidism ("hypothyroidism"), ageing ("older adults", "elderly", and "aged"), and outcomes ("cardiovascular", "fracture", "osteoporosis", "bone", and "mortality"). No language restrictions were applied. No studies were identified that examined the associations between initial levothyroxine dose per bodyweight and cardiovascular, bone, or all-cause mortality outcomes in adults aged 50 years or older. Existing evidence has focused primarily on the risks of overtreatment, defined by suppressed thyroid-stimulating hormone levels, rather than prescribed dose per bodyweight. Current guidance from the UK National Institute for Health and Care Excellence recommends fixed low starting doses of levothyroxine (25–50 µg/day) for adults aged 65 years and older, irrespective of bodyweight. By contrast, younger adults are initiated on weight-based doses. The Baltimore Longitudinal Study of Ageing reported a mean requirement of approximately 1·09 µg/kg per day to maintain euthyroidism in adults aged 65 years and older, suggesting that lower weight-adjusted levothyroxine doses might be appropriate; however, no large-scale study has directly examined their association with long-term outcomes.

Added value of this study

Using primary care data from more than 15 000 older adults (age >50 years) with hypothyroidism in the UK, this observational study emulated a target trial to assess the association between initial levothyroxine dose per bodyweight and cardiovascular, bone, and mortality outcomes. After adjustment for baseline confounding, initiating levothyroxine at doses greater than 1·0 µg/kg per day, whether based on total or lean bodyweight, was associated with higher risks of adverse cardiovascular events, bone health events, and all-cause mortality. This is the first study to evaluate the association between initial weight-based levothyroxine dosing and long-term cardiovascular, bone, and mortality outcomes in older adults with hypothyroidism.

Implications of all the available evidence

UK standard dosing might overestimate levothyroxine requirements in adults older than 50 years. Weight-adjusted, age-appropriate dosing, targeting approximately 1·0 µg/kg per day, might be associated with fewer adverse outcomes. These findings support initiation of discussions to revise clinical guidance to incorporate weight-based levothyroxine prescribing for people aged 65 years and older.

older adults likely reflects age-related increases in thyroid-stimulating hormone (TSH) concentrations,[3] a greater susceptibility to autoimmune thyroid disease, and previous treatments that impair thyroid function, such as thyroid surgery and radioactive iodine therapy.[4]

Levothyroxine, a synthetic form of thyroxine, remains the standard treatment and is among the most widely prescribed medications globally.[5] In the UK, prescribing rates are highest in people older than 80 years, with more than 10% of this group receiving levothyroxine.[6] The treatment goal is to restore and maintain euthyroidism, typically assessed by TSH normalisation.[7]

Older adults, however, are particularly susceptible to the effects of suboptimal levothyroxine dosing. Both under-replacement and over-replacement, indicated by persistently abnormal TSH concentrations outside the population reference range, have been associated with adverse cardiovascular and bone outcomes.[8,9] Standard weight-based dosing (1·6 µg/kg per day) is derived from studies in younger adults and might overestimate levothyroxine needs in older individuals, who generally have lower lean body mass and heightened peripheral sensitivity to thyroid hormones.[10] Reflecting this, the UK National Institute for Health and Care Excellence (NICE) recommends weight-based dosing (1·6 µg/kg per day) in adults younger than 65 years without cardiovascular disease, but substantially lower fixed starting doses (20–50 µg per day) with titration in those aged 65 years and older or with cardiovascular disease.[11] The American Thyroid Association similarly advises individualised dosing based on both age and weight.[12] Additionally, a cautious approach to treatment in older adults is often advocated, particularly where biochemical abnormalities are mild.[13,14] Despite these differences, there remains little evidence directly examining the relationship between initial levothyroxine dose per bodyweight and long-term clinical outcomes in ageing populations.

Long-term randomised trials comparing initial levothyroxine dosing strategies in older adults are unlikely to be feasible because of the sample size and follow-up requirements needed to evaluate long-term safety. Observational data structured within a target trial emulation framework therefore provide an opportunity to address this evidence gap using real-world clinical practice data. By explicitly defining eligibility criteria, treatment strategies, follow-up, and outcomes, target trial emulation can strengthen causal inference from observational data and reduce biases commonly encountered in observational studies.[15]

Using a target trial emulation with UK primary care data, we aimed to examine the association between initial levothyroxine dose per bodyweight and major clinical outcomes in adults with primary hypothyroidism. The causal question of interest was whether initiating levothyroxine at high versus low weight-based doses is associated with

differences in long-term cardiovascular and skeletal outcomes and mortality in routine clinical practice, with the objective of estimating comparative benefits and harms of these dosing strategies in older adults with hypothyroidism.

Methods

Study design and data sources

Data for this observational study were obtained from The Health Improvement Network (THIN), a UK primary care database containing anonymised electronic health records from more than 850 general practices, representing approximately 6% of the UK population. THIN includes longitudinal information on demographics, diagnoses, prescriptions, and laboratory test results recorded in routine primary care and is broadly representative of the UK population in terms of age, sex distribution, major disease prevalence, and mortality rates.[16] Data from Jan 1, 2006, to Dec 31, 2021, were available for analysis. THIN captures data recorded by general practices and does not include linked secondary care or hospital admission data. The database has been validated and widely used in epidemiological research examining levothyroxine therapy, thyroid function, cardiovascular disease, fractures, and mortality.[9,17]

We emulated a target trial to estimate the association between initial levothyroxine dose and long-term cardiovascular, skeletal, and mortality outcomes in adults with hypothyroidism (appendix pp 4–5). The design followed the target trial framework proposed by Hernán and Robins,[18] which uses observational data to reproduce key features of a hypothetical randomised controlled trial, including explicit specification of eligibility criteria, treatment strategies, assignment procedures, and follow-up. This study follows both the STROBE guidelines for observational studies reporting propensity score analysis and the TARGET checklist (appendix pp 6–12).[15,19] The primary causal contrast of interest compared participants according to the initial prescribed levothyroxine dose at treatment initiation; we compared high versus low dosing strategies using a threshold of 1·0 µg/kg per day, informed by data from the Baltimore Longitudinal Study of Aging, which reported a mean levothyroxine requirement of 1·09 µg/kg per day in adults aged 65 years and older.[12] Participants were classified according to their initial prescribed dose at treatment initiation, irrespective of subsequent dose modifications.

People older than 50 years with at least one recorded TSH measurement exceeding 4·0 mU/L between Jan 1, 2006, and Jan 1, 2022, were extracted from THIN, reflecting evidence from the TEARS study demonstrating age-related increases in TSH in older adults.[3] Eligibility was assessed at the time of levothyroxine initiation (time zero). Participants were eligible if they met all of the following: aged older than 50 years at the time of levothyroxine initiation; recorded diagnosis of primary hypothyroidism; initiated levothyroxine therapy between Jan 1, 2006, and Dec 31, 2021; at least one serum TSH measurement recorded within 3 months before or after levothyroxine initiation; at least one bodyweight measurement within 3 months before or after levothyroxine initiation; at least one height measurement (lean body mass analyses only); and at least 6 months of follow-up data available after levothyroxine initiation. 6 months of follow-up data refers to the availability of routine primary care records for at least 6 months after levothyroxine initiation, ensuring sufficient observation time for outcome events.

Participants were excluded if they met any of the following: history of thyroid cancer, hyperthyroidism, or pituitary disease; previous thyroid surgery or radioiodine therapy; use of antithyroid drugs, liothyronine, or medications known to substantially affect thyroid function (including lithium and amiodarone) during the study period; or pre-existing cardiovascular or bone disease relevant to the outcome under assessment.

Diagnostic, prescription, and procedure codes used to define inclusion and exclusion criteria were derived from ICD-10, the National Health Service Dictionary of Medicines and Devices, and the Office of Population Censuses and Surveys Classification of Interventions and Procedures, 4th revision (OPCS-4; appendix pp 13–23). Data cleaning included removal of duplicate records, verification of measurement units, checks for chronological consistency, and assessment of plausibility for key variables. TSH of less than 0·1 mU/L or more than 20 mU/L and bodyweight of less than 35 kg or greater than 200 kg were excluded as extreme values. Records with missing baseline information or with inconsistent follow-up data were excluded.

See Online for appendix

This study received ethical approval from the University of Sunderland Research Ethics Group (reference: 011081). Scientific approval for the use of data from THIN was granted by the THIN Scientific Review Committee (protocol number 22-003). Individual patient consent was waived by the University of Sunderland Research Ethics Group because anonymised, routinely collected electronic health record data were used.

The study protocol was prospectively developed and is available online. The protocol was amended during analysis to refine eligibility criteria, outcome definitions, and the confounder adjustment set, and to include additional sensitivity and supplementary analyses. All protocol changes are summarised in the appendix (p 24).

For the **study protocol** see https://sure.sunderland.ac.uk/id/eprint/19274

Procedures

THIN data are collected during routine primary care consultations by participating UK general practitioners and affiliated health-care professionals using the Vision electronic health record system. Demographic data, diagnoses, prescriptions, laboratory test results, and consultation records entered during routine clinical care were available for analysis. Data on race and ethnicity were not consistently available within THIN and therefore were not analysed in this study.

Baseline demographic variables, including age and sex assigned at birth, were obtained from routinely recorded primary care electronic health records within THIN. Clinical covariates, including hypertension, Charlson Comorbidity Index, smoking status, medication use, serum TSH measurements, and bodyweight, were identified using diagnostic records, prescribing data, and laboratory test entries recorded during routine clinical care. Prescription data were recorded by general practitioners at the point of prescribing and laboratory results were electronically transferred into the medical record from affiliated laboratories (appendix pp 25–170). Cardiovascular, bone, and mortality outcomes were identified using ICD-10 and OPCS-4 codes recorded in THIN (appendix pp 171–84).

Participants were stratified into two age groups: 51–64 years, for whom standard weight-based dosing (1·6 μg/kg per day) is generally recommended, and 65 years or older, for whom NICE advocates lower fixed starting doses (25–50 μg per day).[11] To evaluate the impact of initial levothyroxine dose according to body composition, exposure was defined using two weight-based approaches based on the initial prescribed dose at levothyroxine initiation: (1) total bodyweight dose (μg/kg per day): initial levothyroxine dose (μg per day) divided by total bodyweight (kg); and (2) lean bodyweight dose (μg/kg per day): initial levothyroxine dose (μg per day) divided by estimated lean bodyweight (kg).

Lean bodyweight was calculated using the Boer formula,[20] which incorporates sex assigned at birth, height, and weight. For male individuals: lean bodyweight (kg) = (0·407 × weight [kg]) + (0·267 × height [cm]) – 19·2. For female individuals: lean bodyweight (kg) = (0·252 × weight [kg]) + (0·473 × height [cm]) – 48·3.

For each dosing metric, individuals were categorised into two groups: a high-dose group (>1·0 μg/kg per day) and a low-dose group (≤1·0 μg/kg per day). These classifications were applied separately for total bodyweight and lean bodyweight, with analyses conducted independently for each approach for all outcomes.

TSH and bodyweight were defined as the closest measurements recorded within 3 months before or after levothyroxine initiation. This time period was selected to maximise data completeness and reflects routine UK practice.[11] When multiple measurements were available, the value closest to levothyroxine initiation was used. TSH was used as the primary biochemical marker because it is the most consistently recorded thyroid function test in routine UK clinical practice and is the primary parameter used to guide levothyroxine dose titration. Free thyroxine measurements were inconsistently available and therefore were not included in all analyses. Reported free thyroxine values are based on complete-case data only.

Potential confounders, defined as variables plausibly associated with both levothyroxine dose and study outcomes, were identified a priori using a directed acyclic graph specified in the study protocol and informed by clinical reasoning and previous evidence. The final adjustment set comprised baseline age; sex assigned at birth; Charlson Comorbidity Index; hypertension; serum TSH; smoking status; and use of proton pump inhibitors, oestrogen, or testosterone (appendix p 185).[21,22]

Participants were followed up from the date of levothyroxine initiation until the earliest of: occurrence of the relevant outcome event (cardiovascular, bone, or mortality); death (for non-mortality outcomes); or end of the study period or follow-up period. Follow-up continued for up to 10 years after levothyroxine initiation, with eligible participants contributing follow-up time between Jan 1, 2006, and Dec 31, 2021.

Several measures were undertaken to reduce potential sources of bias. Eligibility criteria, treatment strategies, and follow-up definitions were specified in advance according to a target trial framework to minimise selection and immortal time biases. Time zero was consistently defined as the date of levothyroxine initiation. Extreme, biologically implausible values for TSH and bodyweight were excluded.

Outcomes

Primary outcomes were cardiovascular events, bone health outcomes, and all-cause mortality. Cardiovascular events were defined as the first occurrence of a composite outcome including angina, myocardial infarction, stroke, peripheral vascular disease, heart failure, atrial fibrillation, or any revascularisation procedures recorded during follow-up. Bone outcomes were defined as the first recorded diagnosis of osteoporosis or a fragility fracture during follow-up. All-cause mortality was defined as death recorded during follow-up. Data on cause-specific mortality were unavailable.

Secondary outcomes were health-care utilisation and thyroid function control. Health-care utilisation was defined as the total number of primary care encounters (including face-to-face, telephone, or online consultations) per patient during follow-up. Thyroid function control was defined as the proportion of participants for whom the latest thyroid function test during follow-up showed an abnormal TSH level (<0·4 μIU/mL or >4·0 μIU/mL).

Statistical analysis

The primary causal contrast of interest was comparing high versus low levothyroxine dose groups based on dose at treatment initiation. Because treatment allocation was not randomly assigned and longitudinal adherence data were incomplete, all analyses estimated the observational analogue of the effect of treatment initiation according to baseline prescribed dose group.

Formal sample size calculations were not performed because the study aimed to include the full eligible population defined by the target trial protocol to provide observational estimates rather than test a single prespecified hypothesis.[18] Each of the three primary outcomes (cardiovascular disease, bone health, and all-cause mortality) were analysed separately using bespoke patient cohorts. The cardiovascular disease cohort included individuals without a history of cardiovascular disease at baseline. The bone

health cohort included individuals without a previous fracture or relevant bone outcome at baseline. The mortality cohort included all eligible participants who met the target trial criteria. Secondary outcomes were assessed within the cardiovascular outcome-specific cohort.

To emulate randomisation and adjust for measured confounding, inverse probability of treatment weighting (IPTW) was applied to reduce confounding by indication and other baseline imbalances based on measured covariates. Propensity scores estimating the probability of receiving a high levothyroxine dose were derived from logistic regression models. Unstabilised weights were calculated as the inverse of the propensity score for participants in the high-dose group and the inverse of one minus the propensity score for those in the low-dose group, with extreme weights truncated at the first and 99th percentiles. Propensity score distributions were examined to assess overlap between treatment groups, ensuring adequate positivity (appendix p 186). This created a weighted pseudo-population with balanced baseline covariates between treatment groups. E-values were calculated to assess the robustness of observed associations to potential unmeasured confounding.

Baseline characteristics of participants in the high-dose and low-dose levothyroxine groups were summarised using descriptive statistics. Continuous variables are reported as mean (SD) or median (IQR), and categorical variables as counts and percentages. Covariate balance before and after weighting was assessed using standardised mean differences (SMDs). Baseline data were complete for included participants; therefore, no multiple imputation procedures were required.

Weighted Cox proportional hazards models estimated hazard ratios (HRs) and 95% CIs for each outcome. Proportional hazards assumptions were assessed using Schoenfeld residuals. Weighted Kaplan–Meier curves were used to illustrate event-free survival over time by levothyroxine dose group.[23] Absolute risk differences were estimated at 10 years.

To assess potential non-linear dose–response relationships, IPTW-adjusted Cox models incorporating natural cubic splines with three degrees of freedom were fitted. Health-care utilisation was analysed using IPTW-weighted negative binomial regression, reporting incidence rate ratios (IRRs) with 95% CIs.

Four post-hoc sensitivity analyses were conducted. First, statin use was additionally included in the original IPTW model. Second, calendar year was incorporated into the outcome model. Third, analyses were restricted to participants with TSH measurements recorded before levothyroxine initiation. Fourth, time zero was redefined as the initiation of the second levothyroxine prescription among participants receiving a repeat prescription. Prespecified subgroup analyses were conducted by sex assigned at birth (female and male), age group (51–64 years and ≥65 years), and TSH levels (≤10·0 μIU/mL and >10·0 μIU/mL). Kaplan–Meier analyses were additionally conducted by sex assigned at birth. All analyses were conducted in R version 4.5.1, using the MatchIt package.[24]

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

282 036 records were received from THIN; 263 009 were excluded (figure 1). Overall, 19 027 records were eligible for inclusion in this study (14 296 (75%) individuals were female and 4731 (25%) were male). Mean age was 68 years (SD 10).

Of the 19 027 adults included, 15 463 were included in the primary cardiovascular analysis after excluding 3564 individuals with pre-existing cardiovascular disease at baseline. Based on total bodyweight, 11 762 (76%) participants initiated levothyroxine at a dose of 1·0 μg/kg per day or less and 3701 (24%) at a dose of more than 1·0 μg/kg per day (figure 1). After applying IPTW, the pseudo-population represented 15 430 (51%) and 14 899 (49%) participants in the low-dose and high-dose groups, respectively. Participants had a median follow-up of 10·0 years (IQR 8·6–10·0). Baseline characteristics of the weighted pseudo-population are shown in the table.

After weighting, all covariates in the pseudo-population were well balanced between dose groups in the cardiovascular cohort, with SMDs of less than 0·1 (appendix p 187). The weighted mean age was 67 years (SD 10) in both treatment groups, with similar distributions of sex assigned at birth, smoking status, comorbidity burden, and concomitant medication use (table; appendix pp 188–90). Median TSH was lower in the high-dose group (5·7 μIU/mL [IQR 2·6–10·2]) than in the low-dose group (6·8 μIU/mL [5·2–9·8]) of the cardiovascular cohort, and this difference persisted at the end of follow-up (2·3 μIU/mL [1·2–4·6] *vs* 2·8 μIU/mL [1·6–4·2]). Weighted mean free thyroxine concentrations were higher in the high-dose group both at baseline (15·2 pmol/L [SD 3·6] *vs* 13·1 pmol/L [2·8]) and at the end of follow-up (17·3 pmol/L [SD 3·8] *vs* 15·7 pmol/L [3·1]).

In analyses based on total bodyweight, initiating levothyroxine at a high dose was associated with a statistically significant increased risk of cardiovascular events compared with low-dose therapy (adjusted HR [aHR] 1·18, 95% CI 1·08–1·29; figure 2A). Consistent with these adjusted estimates, unweighted event rates were higher in the high-dose group (728 [20%] of 3701 *vs* 2079 [18%] of 11 762; figure 2A). Similar findings were observed when dosing was defined by lean bodyweight (aHR 1·15, 95% CI 1·06–1·24; figure 2A). Subgroup analyses generally had similar values, although not all estimates were statistically significant (figure 2A). E-values for the cardiovascular outcomes ranged from 1·34 to 2·13, indicating that a moderately strong unmeasured confounder would be required to fully explain the observed associations (appendix p 191). In

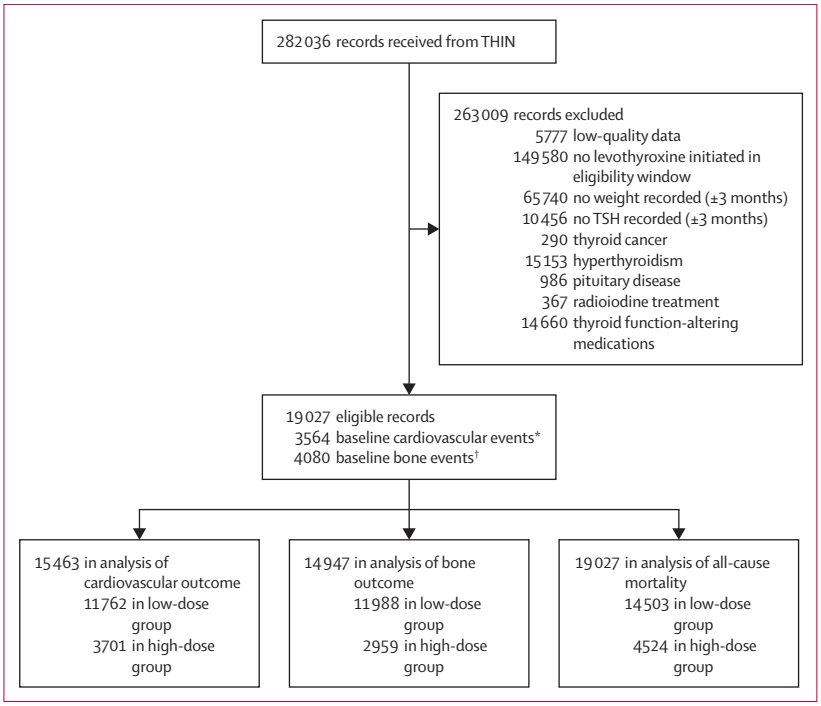


Figure 1: **Flow diagram**
The dose groups shown are based on total bodyweight. THIN=The Health Improvement Network. TSH=thyroid-stimulating hormone. *Individuals with cardiovascular events before time zero were excluded from the primary cardiovascular analysis. †Individuals with bone events before time zero were excluded from the primary bone health analysis.

sensitivity analyses, the association with cardiovascular events was attenuated and no longer statistically significant in analyses including calendar year and with time zero defined at the initiation of the second levothyroxine prescription for all individuals. The remaining sensitivity analyses yielded findings consistent with the primary analysis (appendix pp 192–205).

Kaplan–Meier curves demonstrated lower cardiovascular event-free survival among participants initiated on high-dose levothyroxine (>1·0 µg/kg per day, total bodyweight) compared with low-dose therapy (≤1·0 µg/kg per day, total bodyweight; figure 3A). Similar patterns were observed when dosing was defined using lean bodyweight and split by sex assigned at birth (appendix pp 206–07). The weighted 10-year absolute risk difference between high and low levothyroxine dose groups was 1·3% (95% CI 0·3 to 3·0) for the total bodyweight-based dosing groups and 0·7% (–0·7 to 2·0) for the lean bodyweight-based dosing groups. Spline analyses indicated increasing cardiovascular risk with total bodyweight-based levothyroxine dosing, whereas for lean bodyweight-based dosing, risk appeared to increase modestly at doses between approximately 1·3 µg/kg per day and 3·0 µg/kg per day (appendix pp 208–09).

14 947 individuals were included in the primary bone health analysis. Based on total bodyweight, 11 988 (80%) participants initiated levothyroxine at a dose of 1·0 µg/kg per day or less and 2959 (20%) at a dose of more than 1·0 µg/kg per day. After applying IPTW, the pseudo-population represented 14 939 (51%) and 14 609 (49%) participants in the low-dose and high-dose groups, respectively. Participants had a median follow-up of 10·0 years (IQR 9·1–10·0). Baseline characteristics of the weighted pseudo-population are shown in the table. After weighting, all covariates in the weighted pseudo-population were well balanced between dose groups in the bone health cohort, with SMDs of less than 0·1. The weighted mean age was 68 years (SD 10) in both treatment groups, with similar distributions of sex assigned at birth, smoking status, comorbidity burden, and concomitant medication use. Median TSH was lower in the high-dose group (5·4 µIU/mL [IQR 5·1–10·4]) than in the low-dose group (6·9 µIU/mL [5·2–10·4]). Weighted mean free thyroxine concentrations were higher in the high-dose group than in the low-dose group at baseline (15·1 pmol/L [SD 3·6] *vs* 13·1 pmol/L [2·7]; table).

In analyses based on total bodyweight, initiating levothyroxine at a high dose (>1·0 µg/kg per day) was associated with a statistically significant increased risk of bone outcomes compared with low dose (≤1·0 µg/kg per day; aHR 1·29, 95% CI 1·16–1·43), with unweighted event rates also higher in the high-dose group (387 [13%] of 2959 *vs* 1089 [9%] of 11 988; figure 2B). A similar finding was observed for lean bodyweight-based dosing (aHR 1·16, 95% CI 1·06–1·28; figure 2B). Subgroup analyses had broadly similar values, although not all estimates were statistically significant (figure 2B). E-values for the bone outcomes ranged from 1·49 to 2·79, indicating that a moderately strong unmeasured confounder would be required to fully explain the observed associations (appendix p 191). Sensitivity analyses produced similar results (appendix pp 192–205).

Kaplan–Meier curves demonstrated lower bone event-free survival among participants initiated on high-dose levothyroxine (>1·0 µg/kg per day, total bodyweight) compared with low-dose therapy, although the CIs partly overlapped (figure 3B). Similar patterns were observed for lean bodyweight-based dosing and when split by sex assigned at birth (appendix pp 210–11). The weighted 10-year absolute risk difference between high and low levothyroxine dose groups was 3·0% (95% CI 1·6–4·5) for total bodyweight-based dosing groups and 1·7% (0·5–2·8) for lean bodyweight-based dosing groups. Spline analyses suggested that bone risk was increased at above approximately 1·2 µg/kg per day for total bodyweight-based dosing and above 1·8 µg/kg per day for lean bodyweight-based dosing (appendix pp 212–13).

All 19 027 individuals were included in the primary all-cause mortality analysis. Based on total bodyweight, 14 503 (76%) participants initiated levothyroxine at a dose of 1·0 µg/kg per day or less and 4524 (24%) at a dose of more than 1·0 µg/kg per day. After applying IPTW, the pseudo-population represented 18 990 (51%) and 18 329 (49%) participants in the low-dose and high-dose groups, respectively. Participants had a median follow-up of

	Cardiovascular cohort		Bone health cohort		All-cause mortality cohort	
	Low dose	High dose	Low dose	High dose	Low dose	High dose
N	15 430	14 899	14 939	14 609	18 990	18 329
Sex						
Male	3420 (22%)	3177 (21%)	3955 (26%)	3733 (26%)	4731 (25%)	4390 (24%)
Female	12 010 (78%)	11 722 (79%)	10 984 (74%)	10 876 (74%)	14 259 (75%)	13 939 (76%)
Smoking status						
Smoker	6687 (43%)	6483 (44%)	6756 (45%)	6613 (45%)	8575 (45%)	8307 (45%)
Non-smoker	8743 (57%)	8416 (56%)	8183 (55%)	7996 (55%)	10 415 (55%)	10 022 (55%)
Proton pump inhibitor use						
Yes	5262 (34%)	5053 (34%)	5420 (36%)	5228 (36%)	7173 (38%)	6876 (38%)
No	10 168 (66%)	9846 (66%)	9519 (64%)	9381 (64%)	11 817 (62%)	11 453 (62%)
Hormone therapy						
Yes	213 (1%)	205 (1%)	198 (1%)	191 (1%)	259 (1%)	247 (1%)
No	15 217 (99%)	14 694 (99%)	14 741 (99%)	14 418 (99%)	18 731 (99%)	18 082 (99%)
Statin use						
Yes	6244 (40%)	6190 (42%)	7099 (48%)	7061 (48%)	9188 (48%)	9035 (49%)
No	9186 (60%)	8709 (58%)	7840 (52%)	7548 (52%)	9802 (52%)	9294 (51%)
Chronic pulmonary disease						
Yes	2266 (15%)	2153 (14%)	2333 (16%)	2270 (16%)	3154 (17%)	2981 (16%)
No	13 164 (85%)	12 746 (86%)	12 606 (84%)	12 339 (84%)	15 836 (83%)	15 348 (84%)
Hypertension						
Yes	8431 (55%)	8155 (55%)	8908 (60%)	8695 (60%)	11 619 (61%)	11 225 (61%)
No	6999 (45%)	6744 (45%)	6031 (40%)	5914 (40%)	7371 (39%)	7104 (39%)
Asthma						
Yes	1548 (10%)	1463 (10%)	1522 (10%)	1489 (10%)	2051 (11%)	1964 (11%)
No	13 882 (90%)	13 436 (90%)	13 417 (90%)	13 120 (90%)	16 939 (89%)	16 365 (89%)
Dyslipidaemia						
Yes	1864 (12%)	1801 (12%)	2240 (15%)	2211 (15%)	2942 (15%)	2791 (15%)
No	13 566 (88%)	13 098 (88%)	12 699 (85%)	12 398 (85%)	16 048 (85%)	15 538 (85%)
Diabetes						
Yes	342 (2%)	417 (3%)	383 (3%)	501 (3%)	521 (3%)	606 (3%)
No	15 088 (98%)	14 482 (97%)	14 556 (97%)	14 108 (97%)	18 469 (97%)	17 723 (97%)
Cancer						
Yes	776 (5%)	667 (4%)	827 (6%)	683 (5%)	1044 (5%)	913 (5%)
No	14 654 (95%)	14 232 (96%)	14 112 (94%)	13 926 (95%)	17 946 (95%)	17 416 (95%)
Continuous variables						
Age, years	67 (10)	67 (10)	68 (10)	68 (10)	68 (10)	68 (10)
Weight, kg	79 (19)	75 (16)	80 (19)	75 (16)	79 (19)	75 (16)
BMI, kg/m^2	30 (6)	28 (6)	29·6 (6·3)	28·2 (5·5)	29·6 (6·4)	28·1 (5·6)
TSH, µIU/mL	6·9 (5·2–9·8)	5·7 (2·6–10·2)	6·9 (5·2–10·4)	5·4 (5·1–10·4)	6·9 (5·2–10·5)	4·5 (2·0–8·0)
Free thyroxine, pmol/L	13·1 (2·8)	15·2 (3·6)	13·1 (2·7)	15·1 (3·6)	13·1 (2·8)	15·3 (3·6)
Levothyroxine dose per weight, µg/kg per day	0·5 (0·3–0·7)	1·4 (1·2–1·6)	0·5 (0·3–0·7)	1·4 (1·2–1·6)	0·5 (0·3–0·7)	1·4 (1·2–1·6)

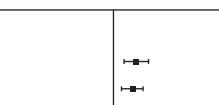
Data are n (%), mean (SD), or median (IQR), unless indicated otherwise. The table shows a weighted pseudo-population with balanced baseline covariates between treatment groups. Low dose was defined as levothyroxine dose <1·0 µg/kg per day; high dose was defined as levothyroxine dose ≥1·0 µg/kg per day. Free thyroxine and TSH were defined as the closest available measurements within 3 months before or after treatment initiation, selecting the value nearest to time zero. TSH=thyroid-stimulating hormone.

Table: **Baseline characteristics of the weighted pseudo-population by levothyroxine dose group, based on total bodyweight**

10·0 years (IQR 9·3–10·0). After weighting, all covariates in the weighted pseudo-population were well balanced between dose groups in the all-cause mortality cohort, with SMDs of less than 0·1. The weighted mean age was 68 years (SD 10) in both treatment groups, with similar distributions of sex assigned at birth, smoking status, comorbidity burden, and concomitant medication use. Median TSH was lower in the high-dose group (4·5 µIU/mL [IQR 2·0–8·0]) than in the low-dose group (6·9 µIU/mL [5·2–10·5]). Weighted mean free thyroxine concentrations were higher

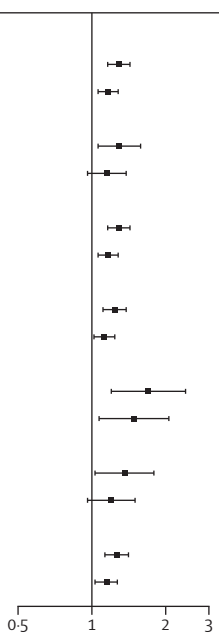
A Cardiovascular events

		Low dose Events/patients	High dose Events/patients	Adjusted HR (95% CI)	p value
Overall					
	Total bodyweight	2079/11762	728/3701	1·18 (1·08–1·29)	<0·0001
	Lean bodyweight	1458/8025	1291/7129	1·15 (1·06–1·24)	<0·0001
Age 51–64 years					
	Total bodyweight	562/5421	180/1674	1·10 (0·92–1·30)	0·30
	Lean bodyweight	395/3669	336/3299	1·15 (0·99–1·34)	0·073
Age ≥65 years					
	Total bodyweight	1517/6341	548/2027	1·18 (1·08–1·29)	<0·0001
	Lean bodyweight	1063/4356	955/3830	1·15 (1·06–1·24)	<0·0001
Female					
	Total bodyweight	1417/8889	564/3154	1·13 (1·02–1·25)	0·015
	Lean bodyweight	873/5456	1064/6357	1·07 (0·98–1·17)	0·14
Male					
	Total bodyweight	662/2873	164/547	1·33 (1·12–1·59)	0·0011
	Lean bodyweight	585/2569	227/772	1·39 (1·19–1·62)	<0·0001
High TSH (>10·0 µIU/mL)					
	Total bodyweight	450/3071	100/558	1·32 (1·05–1·65)	0·016
	Lean bodyweight	317/2077	216/1456	1·10 (0·92–1·31)	0·32
Low TSH (≤10·0 µIU/mL)					
	Total bodyweight	1629/8691	628/3143	1·14 (1·04–1·25)	0·0072
	Lean bodyweight	1141/5948	1075/5673	1·15 (1·06–1·26)	0·0013



B Bone health events

		Low dose Events/patients	High dose Events/patients	Adjusted HR (95% CI)	p value
Overall					
	Total bodyweight	1089/11988	387/2959	1·29 (1·16–1·43)	<0·0001
	Lean bodyweight	730/8606	724/6047	1·16 (1·06–1·28)	0·0018
Age 51–64 years					
	Total bodyweight	320/5153	112/1257	1·29 (1·06–1·58)	0·010
	Lean bodyweight	207/3574	222/2719	1·15 (0·96–1·38)	0·13
Age ≥65 years					
	Total bodyweight	769/6835	275/1702	1·29 (1·16–1·43)	<0·0001
	Lean bodyweight	523/5032	502/3328	1·16 (1·06–1·28)	0·0018
Female					
	Total bodyweight	972/8547	361/2448	1·24 (1·11–1·38)	<0·0001
	Lean bodyweight	622/5476	692/5314	1·12 (1·02–1·24)	0·021
Male					
	Total bodyweight	117/3441	26/511	1·70 (1·20–2·41)	0·0030
	Lean bodyweight	108/3130	32/733	1·48 (1·07–2·06)	0·018
High TSH (>10·0 µIU/mL)					
	Total bodyweight	258/3356	58/552	1·36 (1·03–1·79)	0·032
	Lean bodyweight	163/2355	146/1457	1·20 (0·96–1·50)	0·10
Low TSH (≤10·0 µIU/mL)					
	Total bodyweight	831/8632	329/2407	1·26 (1·13–1·41)	<0·0001
	Lean bodyweight	567/6251	578/4590	1·15 (1·03–1·27)	0·012



(Figure 2 continues on next page)

in the high-dose group than in the low-dose group at baseline (15·3 pmol/L [SD 3·6] *vs* 13·1 pmol/L [SD 2·8]; table).

In analyses based on total bodyweight, initiating levothyroxine at a high dose was associated with a statistically significant increased risk of all-cause mortality compared with low dose (aHR 1·17, 95% CI 1·09–1·25), with a similar association observed using lean bodyweight-based dosing (1·12, 1·05–1·19; figure 2C). Unweighted event rates were higher in the high-dose group (836 [18%] of 4524 *vs* 2532 [17%] of 14 503; figure 2C). Subgroup analyses generally had similar values, although not all estimates were

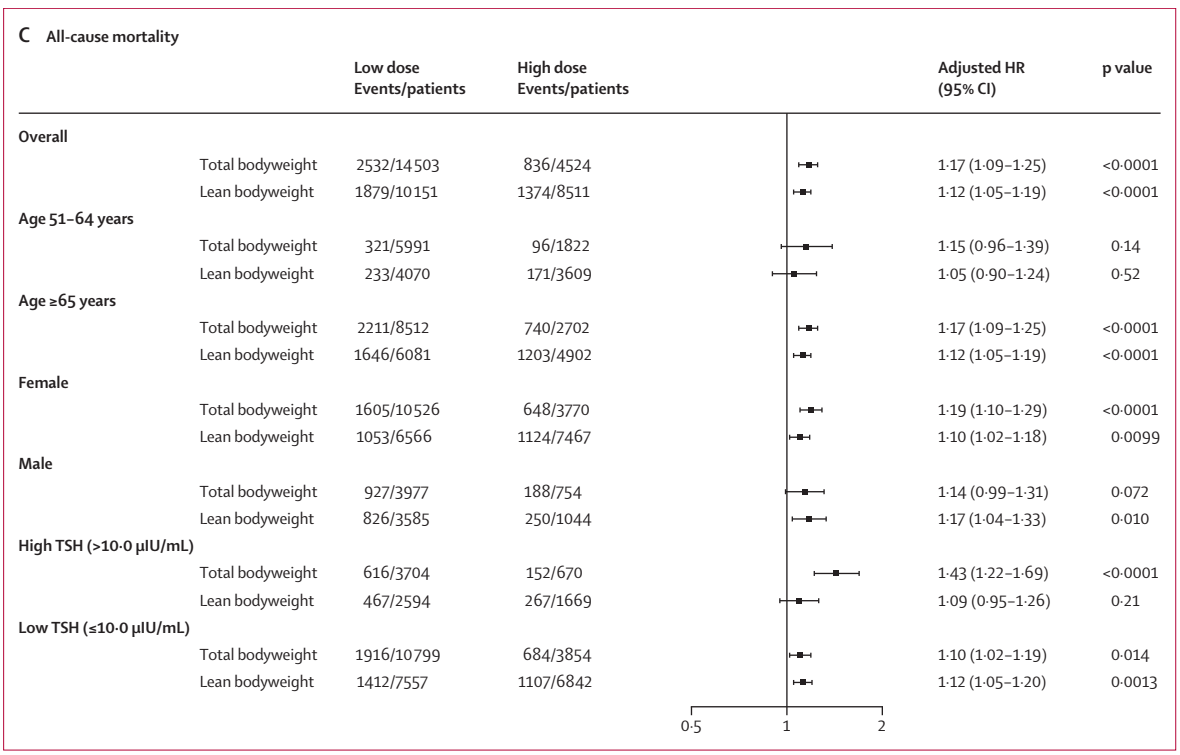


Figure 2: **Cardiovascular events, bone health events, and all-cause mortality by levothyroxine dose based on total bodyweight and lean bodyweight, overall and by subgroup**

Forest plots show unweighted numbers of events and patients, and adjusted HRs with 95% CIs for cardiovascular events (A), bone health outcomes (B), and all-cause mortality (C). HR=hazard ratio. TSH=thyroid-stimulating hormone.

statistically significant (figure 2C). E-values for mortality outcomes ranged from 1·28 to 2·21, indicating that unmeasured confounding would need to be moderately strong to fully account for the observed associations (appendix p 191). Sensitivity analyses produced mostly similar results (appendix pp 192–93, 196–200); however, sensitivity analyses including calendar year and with time zero defined at the initiation of the second levothyroxine prescription among participants receiving a repeat prescription were not significant for individuals with lean bodyweight (appendix pp 194–95, 201–05).

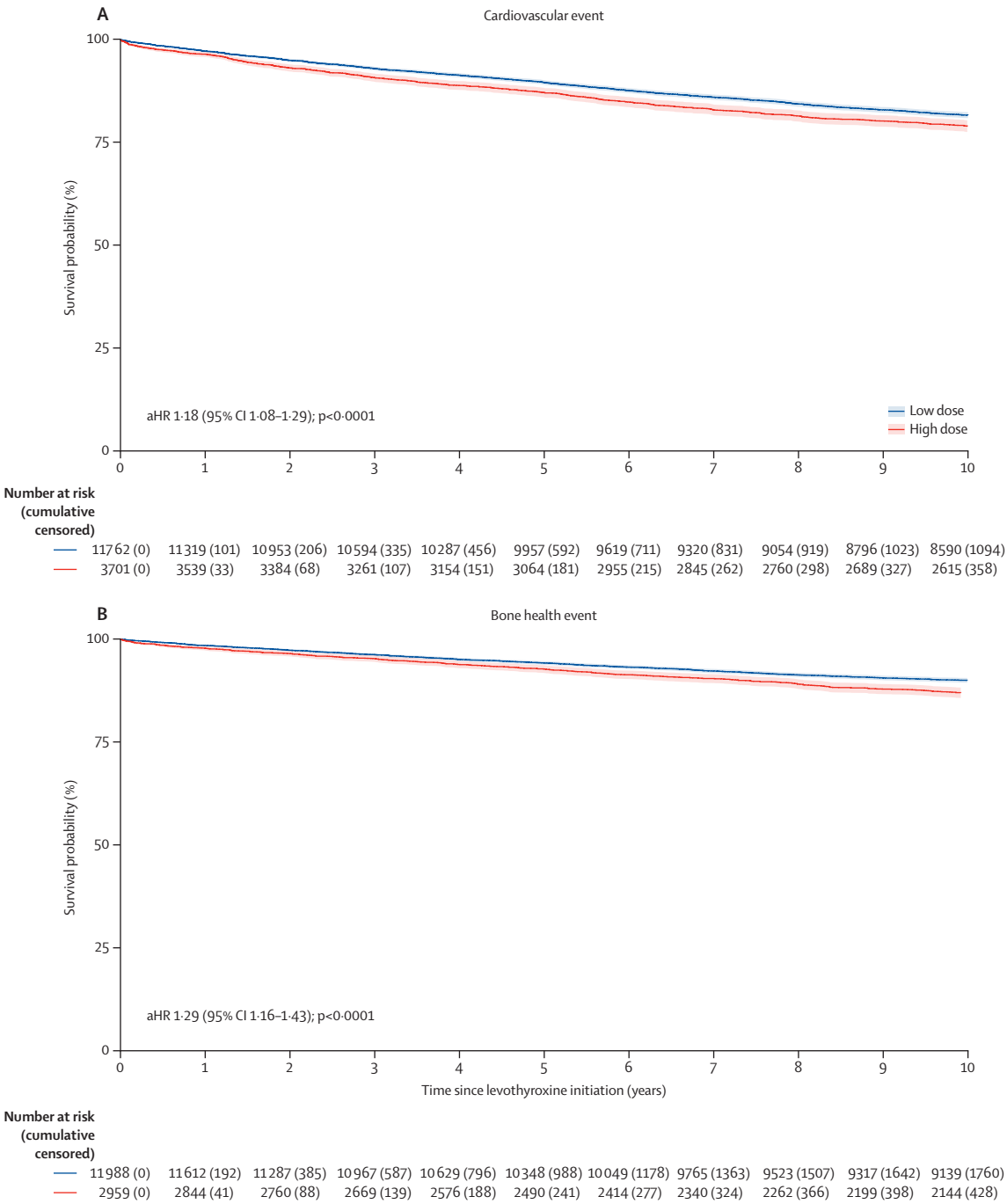
Kaplan–Meier curves indicated no difference in event-free survival for all-cause mortality among participants initiated on high levothyroxine doses compared with low doses, with overlapping 95% CIs (figure 3C). Similar patterns were observed for lean bodyweight-based dosing and when split by sex assigned at birth (appendix pp 214–15). The weighted 10-year absolute risk difference was 3·4% (95% CI 1·7–5·1) for total bodyweight-based dosing groups and 2·2% (0·8–3·5) for lean bodyweight-based dosing groups. Spline analysis demonstrated a gradual increase in mortality risk, with total bodyweight-based levothyroxine dosing above approximately 1·0 μg/kg per day (appendix p 216). For lean bodyweight-based dosing, the relationship was less clear, with risk appearing to increase modestly at doses between approximately 1·4 μg/kg per day and 3·0 μg/kg per day (appendix p 217).

In the weighted cardiovascular cohort, adults with hypothyroidism who were initiated on high levothyroxine dose (>1·0 μg/kg per day, total bodyweight) had greater healthcare utilisation compared with those started on low dose (≤1·0 μg/kg per day, total bodyweight). The high-dose group had a statistically significantly higher number of primary care consultations compared with the low-dose group (640 405 *vs* 1 850 731 consultations; IRR 1·09, 95% CI 1·06–1·13). The median number of consultations was 246 (IQR 152–378) in the low-dose group and 272 (164–422) in the high-dose group.

At study end, abnormal TSH levels (<0·4 μIU/mL or >4·0 μIU/mL) were recorded in 4166 (27%) of 15 430 participants in the weighted low-dose cardiovascular cohort and 4470 (30%) of 14 899 in the weighted high-dose cardiovascular cohort. Within the low-dose group, 917 (22%) of 4166 had low TSH values and 3249 (78%) had elevated TSH values, compared with 1935 (43%) of 4470 and 2535 (57%), respectively, in the high-dose group.

Discussion

In this observational study emulating a target trial of adults older than 50 years with primary hypothyroidism, initiation of levothyroxine at doses greater than 1·0 μg/kg per day, whether based on total bodyweight or lean bodyweight, was associated with higher risks of cardiovascular events, bone outcomes, and all-cause mortality than initiation of



(Figure 3 continues on next page)

levothyroxine at doses of 1·0 µg/kg per day or less. Stratified analyses suggested that the potential harm of high initial levothyroxine dose was more pronounced in adults aged 65 years and older, whereas outcomes in those aged 50–64 years were largely non-significant, highlighting age-related differences in sensitivity to levothyroxine dosing. Participants initiated on high levothyroxine dose also had a greater number of primary care consultations, suggesting an association between high initial dosing and increased health-care utilisation. Together, these findings suggest that high initial weight-based levothyroxine dosing in older adults is associated with less favourable long-term outcomes.

For lean bodyweight-based dosing, binary analyses comparing doses above versus below 1·0 µg/kg per day suggested higher risks of cardiovascular events and all-cause

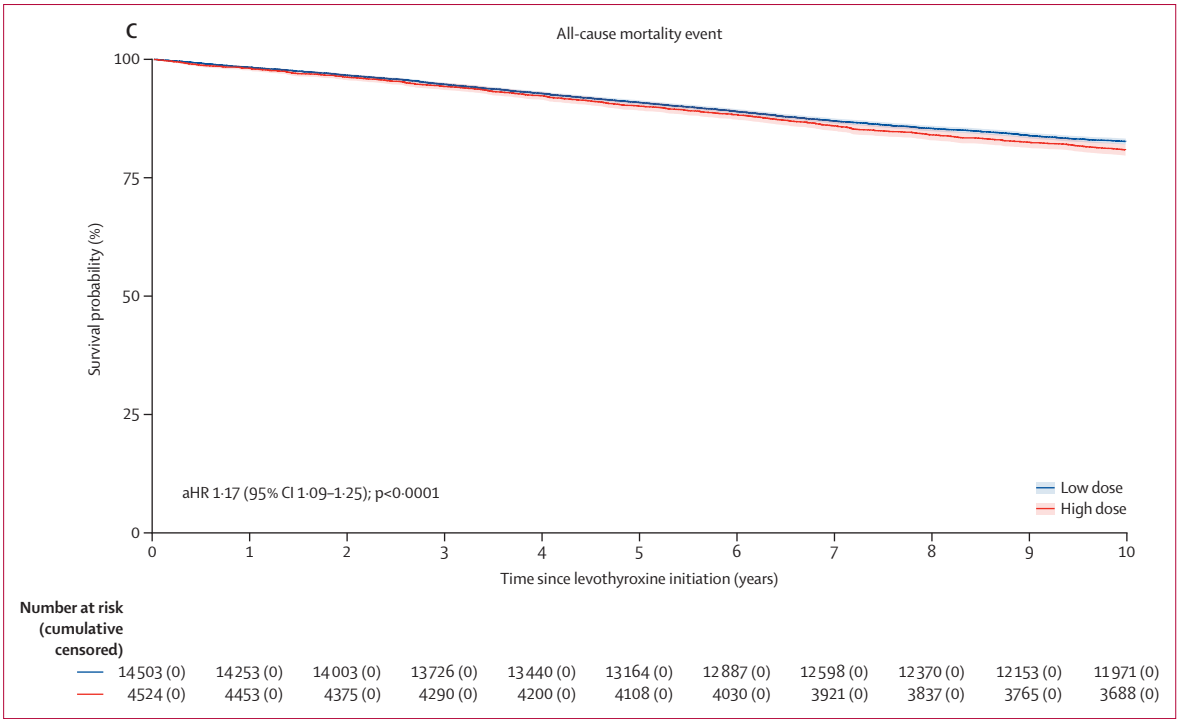


Figure 3: **Kaplan–Meier survival curves for cardiovascular events, bone health events, and all-cause mortality according to levothyroxine dose based on total bodyweight**

(A) Cardiovascular events. (B) Bone health events. (C) All-cause mortality. Red lines represent the high-dose group (>1·0 µg/kg per day, total bodyweight) and black lines represent the low-dose group (≤1·0 µg/kg per day, total bodyweight). Shaded areas represent 95% CIs. aHR=adjusted hazard ratio.

mortality. By contrast, spline analyses treating lean bodyweight as a continuous variable showed HRs mostly near 1 at lower doses of levothyroxine, with only a modest increase in risk at doses between approximately 1·3 µg/kg per day and 3·0 µg/kg per day. This pattern might reflect a threshold effect around 1·0 µg/kg per day, which the binary analysis captures, or the small number of participants at very low or very high lean bodyweight doses, which can flatten the spline curve. Total bodyweight-based dosing consistently showed increasing risk across the relevant dose range in both binary and spline analyses.

Our findings are consistent with previous observational studies linking levothyroxine overtreatment to increased cardiovascular morbidity and mortality in older adults.[8] For example, higher rates of atrial fibrillation have been reported among individuals receiving levothyroxine doses exceeding 75 µg/day.[25] Similarly, the increased risk of osteoporosis and fractures observed in this study aligns with earlier findings of adverse skeletal outcomes among older adults treated with high levothyroxine dose.[26–28] Previous literature also suggests that multiple levothyroxine dose adjustments could be associated with increased health-care utilisation, including more frequent monitoring, medication adjustments, and management of complications.[29]

Together, this literature supports that levothyroxine requirements derived from studies in younger adults, typically approximately 1·6 µg/kg per day, might not be appropriate for older populations, who have lower lean body mass and increased hormone sensitivity. However, the present study examines the initial prescribed dose rather than achieved biochemical control and, therefore, complements rather than replaces evidence based on TSH-defined overtreatment.

Target trial emulation seeks to approximate the causal question that would be addressed in a hypothetical randomised controlled trial. This is achieved by strictly predefining eligibility, treatment, and follow-up so that it matches as closely as possible to a target randomised controlled trial, which minimises the biases usually found in observational studies.

This study's strengths include its use of a large, population-based UK primary care cohort with long-term follow-up and the application of a target trial emulation framework to reduce bias through clearly defined eligibility, treatment strategies, and follow-up.[18] To emulate a randomised trial and improve comparability between treatment groups, IPTW was applied. This method adjusts the contribution of each participant so that baseline characteristics are balanced across a pseudo-population, the size of which reflects the reweighted contributions of participants. Additionally, the evaluation of dosing based on both total bodyweight and lean bodyweight provides a novel insight into practical alternatives for safer levothyroxine prescribing in older adults.

Several limitations should be considered. As with all observational analyses, residual confounding cannot be excluded, particularly confounding by indication, whereby

patients prescribed high initial dose might differ systematically in ways not fully captured by measured covariates. Although IPTW was used to balance observed baseline characteristics, unmeasured factors such as disease severity, frailty, malabsorption, adherence, or clinician prescribing preferences could have influenced both dose selection and outcomes. Race and ethnicity data were not available within THIN and, therefore, could not be accounted for in our analyses, despite their potential association with health-care access, treatment patterns, and health outcomes. Additionally, the study population was highly selected, comprising approximately 5% of patients identified in the database who met eligibility criteria; this reflects the application of strict inclusion and exclusion criteria required for the emulated target trial design, but might limit generalisability to all levothyroxine users in routine clinical practice.

Longitudinal data on dose adjustments, adherence, and serial TSH measurements were incomplete, precluding per-protocol analyses. Analysis therefore focused on initial levothyroxine dose as a proxy for intended treatment strategy, based on exposure at treatment initiation, recognising that subsequent titration commonly occurs in clinical practice. For cardiovascular outcomes, the association observed in the primary analysis was attenuated and no longer statistically significant when time zero was redefined at the second levothyroxine prescription. This finding suggests that part of the observed cardiovascular association could reflect early treatment changes, treatment persistence, or residual confounding. By contrast, associations with bone health outcomes and all-cause mortality remained broadly consistent across sensitivity analyses.

Estimated lean bodyweight might be inaccurate in frail older adults, potentially causing misclassification.[30] Excluding patients without recorded weights could introduce selection bias, because weights are more often measured in those with multiple comorbidities or in individuals who are underweight or overweight. Categorising doses using a threshold of 1·0 μg/kg per day could group widely different doses together while separating similar doses across the cutoff, potentially obscuring dose–response relationships. Sensitivity analyses using natural cubic splines to model dose continuously produced similar associations, supporting the robustness of the findings.

Some baseline measurements, including TSH, might have been recorded shortly after levothyroxine initiation, which could affect the observed associations. TSH and bodyweight were defined as the closest measurements within 3 months before or after treatment, reflecting real-world UK primary care practice where some values might be entered electronically after initiation. Consequently, the lower TSH and higher free thyroxine observed in the high-dose group might partly reflect early treatment effects rather than true pretreatment thyroid function. Free thyroxine and free triiodothyronine were measured inconsistently and therefore not included in weighting. However, in sensitivity analyses restricted to participants with TSH measurements recorded before levothyroxine initiation, effect estimates were materially unchanged, suggesting that inclusion of post-initiation measurements did not substantially bias the findings.

Diagnosis was based on a coded record of primary hypothyroidism, and TSH was included as a baseline variable without systematic free thyroxine or free triiodothyronine measurements, because these were not routinely recorded. This approach captures both overt and subclinical disease. Nevertheless, initial levothyroxine dosing is generally recommended uniformly for elevated TSH, which aligns with the study's focus on initial dosing strategies. Patients with post-surgical or post-ablative hypothyroidism were excluded to maintain a clinically homogeneous cohort, as these individuals typically require immediate full replacement dosing and follow different treatment pathways.

THIN captures only UK primary care data and does not reliably include information on ethnicity, family history, consultation type, or secondary care. Cause of death data were not available, precluding assessment of cause-specific mortality and limiting interpretation of mortality findings. Death represents a competing risk for cardiovascular and bone health outcomes, and this limitation should be considered when interpreting the results. Health-care utilisation was included as a secondary outcome and could also act as a mediator, potentially affecting dose titration and earlier detection of complications. The analyses conducted involved multiple outcomes, exposure definitions, and subgroup comparisons, and some estimates—particularly in subgroups—might reflect chance rather than true associations. Residual confounding from unmeasured factors, as well as from variation between practices and care pathways, cannot be excluded. Prescribing practices for levothyroxine might have changed over the study period, and evolving clinical guidance and clinician behaviour could have introduced residual confounding. Although we adjusted for a wide range of measured baseline characteristics, we did not include a negative control outcome to further assess the potential impact of residual confounding. Diagnostic coding inaccuracies and the restriction to primary care might limit generalisability to other settings or health-care systems.

These findings highlight potential risks associated with high initial levothyroxine dosing in adults older than 50 years, particularly in those aged 65 years and older, and emphasise the need for careful dose selection and monitoring in this population. The results describe associations observed in routine clinical practice and should not be interpreted as prescriptive dosing recommendations. Clinical decisions should continue to balance symptom control, biochemical targets, comorbidities, and patient preferences.

Although weight-based levothyroxine dosing might provide a more tailored approach, its implementation in routine primary care requires consideration of feasibility. Accurate weight measurement is essential, and calculation

tools could be integrated into electronic prescribing systems or clinical decision support prompts to help facilitate safe and practical dosing in everyday practice.

Future research should aim to evaluate alternative levothyroxine initiation strategies in older adults, ideally through pragmatic randomised trials or prospective cohort studies incorporating adherence monitoring, serial thyroid function testing, cardiovascular and skeletal endpoints, and patient-reported outcomes. Further work is also needed to clarify optimal TSH targets and the effect of frailty, multimorbidity, and polypharmacy on levothyroxine requirements in ageing populations.

Contributors

SR developed the study concept. SW, SR, MH, RD, IM, and EDM designed the study. MH conducted all analyses, with guidance from EDM, SR, RD, IM, and SW. MH drafted the manuscript with feedback from SW, SR, RD, IM, and EDM. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication. MH and IM directly accessed and verified the underlying data reported in the manuscript. All authors read and approved the final manuscript and agree to be accountable for all aspects of the work.

Declaration of interests

SR has received speaker fees from Merck KGaA and Berlin Chemie, manufacturers of levothyroxine. All other authors declare no competing interests.

Data sharing

Data used in this study are available from THIN, a wholly owned subsidiary of Cegedim, which owns the proprietary rights to THIN data. Restrictions apply to the availability of these data, which were used under licence for the current study and are not publicly available. Researchers may apply for access directly through THIN or Cegedim, subject to approval and completion of the relevant data access and licensing agreements.

Acknowledgments

This study was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration North East and North Cumbria (NIHR200173). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

References

1 Cooper DS, Biondi B. Subclinical thyroid disease. *Lancet* 2012; **379:** 1142–54.
2 Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med* 2000; **160:** 526–34.
3 Vadiveloo T, Donnan PT, Murphy MJ, Leese GP. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: the Thyroid Epidemiology, Audit, and Research Study (TEARS). *J Clin Endocrinol Metab* 2013; **98:** 1147–53.
4 Surks MI, Hollowell JG. Age-specific distribution of serum thyrotropin and antithyroid antibodies in the US population: implications for the prevalence of subclinical hypothyroidism. *J Clin Endocrinol Metab* 2007; **92:** 4575–82.
5 Fournier J, Barret L, Khouri C, Naudet F, Boussageon R, Roustit M. The evidence base of the 10 most prescribed drugs in England, France, and the United States: a scoping review. *J Clin Epidemiol* 2024; **174:** 111478.
6 Leng O, Razvi S. Hypothyroidism in the older population. *Thyroid Res* 2019; **12:** 2.
7 Centanni M, Duntas L, Feldt-Rasmussen U, et al. ETA guidelines for the use of levothyroxine sodium preparations in monotherapy to optimize the treatment of hypothyroidism. *Eur Thyroid J* 2025; **14:** e250123.
8 Feldt-Rasmussen U, Effraimidis G, Bliddal S, Klose M. Risks of suboptimal and excessive thyroid hormone replacement across ages. *J Endocrinol Invest* 2024; **47:** 1083–90.
9 Thayakaran R, Adderley NJ, Sainsbury C, et al. Thyroid replacement therapy, thyroid stimulating hormone concentrations, and long term health outcomes in patients with hypothyroidism: longitudinal study. *BMJ* 2019; **366:** l4892.
10 Santini F, Pinchera A, Marsili A, et al. Lean body mass is a major determinant of levothyroxine dosage in the treatment of thyroid diseases. *J Clin Endocrinol Metab* 2005; **90:** 124–27.
11 National Institute for Health and Care Excellence. Thyroid disease: assessment and management (NG145). Nov 20, 2019. https://www.nice.org.uk/guidance/ng145 (accessed June 25, 2026).
12 Jonklaas J, Bianco AC, Bauer AJ, 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.
13 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. *Syst Rev* 2024; **13:** 123.
14 Holley M, Razvi S, Maxwell I, Dew R, Wilkes S. Assessing the cardiovascular effects of levothyroxine use in an ageing UK population with subclinical hypothyroidism: emulated target trial (ACEL-UK-ETT). *Thyroid* 2026; **36:** 242–50.
15 Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA* 2022; **328:** 2446–47.
16 Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Inform Prim Care* 2011; **19:** 251–55.
17 Holley M, Razvi S, Maxwell I, Dew R, Wilkes S. Assessing the cardiovascular effects of levothyroxine use in an ageing United Kingdom population (ACEL-UK): cohort study. *J Clin Endocrinol Metab* 2025; **110:** e4137–43.
18 Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016; **183:** 758–64.
19 Yao XI, Wang X, Speicher PJ, et al. Reporting and guidelines in propensity score analysis: a systematic review of cancer and cancer surgical studies. *J Natl Cancer Inst* 2017; **109:** djw323.
20 Boer P. Estimated lean body mass as an index for normalization of body fluid volumes in humans. *Am J Physiol* 1984; **247:** F632–36.
21 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.
22 Vita R, Saraceno G, Trimarchi F, Benvenga S. Switching levothyroxine from the tablet to the oral solution formulation corrects the impaired absorption of levothyroxine induced by proton-pump inhibitors. *J Clin Endocrinol Metab* 2014; **99:** 4481–86.
23 Denz R, Klaaßen-Mielke R, Timmesfeld N. A comparison of different methods to adjust survival curves for confounders. *Stat Med* 2023; **42:** 1461–79.
24 R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, 2024.
25 Gong IY, Atzema CL, Lega IC, et al. Levothyroxine dose and risk of atrial fibrillation: a nested case-control study. *Am Heart J* 2021; **232:** 47–56.
26 Ko Y-J, Kim JY, Lee J, et al. Levothyroxine dose and fracture risk according to the osteoporosis status in elderly women. *J Prev Med Public Health* 2014; **47:** 36–46.
27 Leese GP, Flynn RV. Levothyroxine dose and fractures in older adults. *BMJ* 2011; **342:** d2250.
28 Turner MR, Camacho X, Fischer HD, et al. Levothyroxine dose and risk of fractures in older adults: nested case-control study. *BMJ* 2011; **342:** d2238.
29 Ernst FR, Barr P, Elmor R, et al. The economic impact of levothyroxine dose adjustments: the CONTROL HE Study. *Clin Drug Investig* 2017; **37:** 71–83.
30 Mitchell SJ, Kirkpatrick CMJ, Le Couteur DG, et al. Estimation of lean body weight in older community-dwelling men. *Br J Clin Pharmacol* 2010; **69:** 118–27.