Thyroid function tests in the reference range and fracture: individual participant analysis of prospective cohorts

Aubert, Carole E.; Floriani, Carmen; Bauer, Douglas C.; da Costa, Bruno R.; Segna, Daniel; Blum, Manuel R.; Collet, Tinh-Hai; Fink, Howard A.; Cappola, Anne R.; Syrogianouli, Lamprini; Peeters, Robin P.; Åsvold, Bjørn O.; den Elzen, Wendy P. J.; Luben, Robert N.; Bremmer, Alexandra P.; Gogakos, Apostolos; Eastell, Richard; Kearney, Patricia M.; Hoff, Mari; Le Blanc, Erin; Ceresini, Graziano; Rivadeneira, Fernando; Uitterlinden, André G.; Khaw, Kay-Tee; Langhammer, Arnulf; Stott, David J.; Westendorp, Rudi G. J.; Ferrucci, Luigi; Williams, Graham R.; Gussekloo, Jacobijn; Walsh, John P.; Aujesky, Drahomir; Rodondi, Nicolas

2017-05-05

Aubert, C. E. et al. (2017) 'Thyroid function tests in the reference range and fracture: individual participant analysis of prospective cohorts', Journal of Clinical Endocrinology and Metabolism, 102(8), pp. 2719-2728. doi:10.1210/jc.2017-00294

Article (peer-reviewed)

http://dx.doi.org/10.1210/jc.2017-00294

Access to the full text of the published version may require a subscription.

© 2017, Endocrine Society. Published by Oxford University Press. All rights reserved. This is a pre-copyedited, author-produced version of an article accepted for publication in Journal of Clinical Endocrinology and Metabolism following peer review. The version of record [Aubert, C. E. et al. (2017) 'Thyroid function tests in the reference range and fracture: individual participant analysis of prospective cohorts', Journal of Clinical Endocrinology and Metabolism, 102(8), pp. 2719-2728. doi:10.1210/jc.2017-00294] is available online at: https://doi.org/10.1210/jc.2017-00294

http://hdl.handle.net/10468/6369

Downloaded on 2018-12-11T02:20:34Z
Thyroid function tests in the reference range and fracture: individual participant analysis of prospective cohorts

Carole E Aubert, MD, Carmen Floriani, MD, Douglas C Bauer, MD, Bruno R da Costa, PhD, Daniel Segna, MD, Manuel R Blum, MD, Tinh-Hai Collet, MD, Howard A Fink, MD, MPH, Anne R Cappola, MD, ScM, Lamprini Syrogianouli, Robin P Peeters, MD, PhD, Bjørn O Åsvold, MD, PhD, Wendy PJ den Elzen, PhD, Robert N Luben, PhD, Alexandra P Bremner, PhD, Apostolos Gogakos, MD, PhD, Richard Eastell, MD, Patricia M Kearney, MD, PhD, MPH, Mari Hoff, MD, PhD, Erín Le Blanc, MD, MPH, Graziano Ceresini, MD, PhD, Fernando Rivadeneira, MD, PhD, André G Uitterlinden, PhD, Kay-Tee Khaw, MD, Arnulf Langhammer, MD, PhD, David J Stott, MD, PhD, Rudi GJ Westendorp, MD, PhD, Luigi Ferrucci, MD, PhD, Graham R Williams, MBBS, PhD, Jacobijn Gussekloo, MD, PhD, John P Walsh, MBBS, PhD, Drahomir Aujesky, MD, MSc, Nicolas Rodondi, MD, MAS, for the Thyroid Studies Collaboration.

The Journal of Clinical Endocrinology & Metabolism
Endocrine Society

Submitted: January 31, 2017
Accepted: April 26, 2017
First Online: May 05, 2017

Advance Articles are PDF versions of manuscripts that have been peer reviewed and accepted but not yet copyedited. The manuscripts are published online as soon as possible after acceptance and before the copyedited, typeset articles are published. They are posted "as is" (i.e., as submitted by the authors at the modification stage), and do not reflect editorial changes. No corrections/changes to the PDF manuscripts are accepted. Accordingly, there likely will be differences between the Advance Article manuscripts and the final, typeset articles. The manuscripts remain listed on the Advance Article page until the final, typeset articles are posted. At that point, the manuscripts are removed from the Advance Article page.

DISCLAIMER: These manuscripts are provided "as is" without warranty of any kind, either express or particular purpose, or non-infringement. Changes will be made to these manuscripts before publication. Review and/or use or reliance on these materials is at the discretion and risk of the reader/user. In no event shall the Endocrine Society be liable for damages of any kind arising references to, products or publications do not imply endorsement of that product or publication.
Normal thyroid function tests and fracture risk

**Thyroid function tests in the reference range and fracture: individual participant analysis of prospective cohorts**

Carole E Aubert, MD, Carmen Floriani, MD, Douglas C Bauer, MD, Bruno R da Costa, PhD, Daniel Segna, MD, Manuel R Blum, MD, Tinh-Hai Collet, MD, Howard A Fink, MD, MPH, Anne R Cappola, MD, ScM, Lamprini Syrogiannouli, Robin P Peeters, MD, PhD, Bjørn O Asvold, MD, PhD, Wendy PJ den Elzen, PhD, Robert N Luben, PhD, Alexandra P Bremner, PhD, Apostolos Gogakos, MD, PhD, Richard Eastell, MD, Patricia M Kearney, MD, PhD, MPH, Mari Hoff, MD, PhD, Erin Le Blanc, MD, MPH, Graziano Ceresini, MD, PhD, Fernando Rivadeneira, MD, PhD, André G Uitterlinden, PhD, Kay-Tee Khaw, MD, Arnulf Langhammer, MD, PhD, David J Stott, MD, PhD, Rudi GJ Westendorp, MD, PhD, Luigi Ferrucci, MD, PhD, Graham R Williams, MBBS, PhD, Jacobijn Gussekloo, MD, PhD, John P Walsh, MBBS, PhD, Drahomir Aujesky, MD, MSc, Nicolas Rodondi, MD, MAS, for the Thyroid Studies Collaboration.

1 Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland; 2 Departments of Medicine and Epidemiology & Biostatistics, University of California, San Francisco, United States; 3 Institute of Primary Health Care (BIHAM), University of Bern, Switzerland; 4 Department of Cardiology, Swiss Cardiovascular Center Bern, Bern University Hospital, Bern, Switzerland; 5 Service of Endocrinology, Diabetes and Metabolism, University Hospital of Lausanne, Lausanne, Switzerland; 6 Geriatric Research Education & Clinical Center, Veterans Affairs Health Care System, Minneapolis, United States; 7 Department of Medicine, University of Minnesota, Minneapolis, United States; 8 University of Pennsylvania School of Medicine, Philadelphia, PA, United States; 9 Department of Internal Medicine & Department of Epidemiology, Erasmus University Rotterdam, Rotterdam, The Netherlands; 10 Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway; 11 Department of Endocrinology, St. Olav’s Hospital, Trondheim University Hospital, Trondheim, Norway; 12 Leiden University Medical Center, Department of Clinical Chemistry and Laboratory Medicine, Leiden, The Netherlands; 13 Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom; 14 School of Population Health, University of Western Australia, Crawley, WA, Australia; 15 Molecular Endocrinology Laboratory, Department of Medicine, Imperial College London, London, United Kingdom; 16 Department of Human Metabolism, University of Sheffield, Sheffield, United Kingdom; 17 Department of Epidemiology and Public Health, University College Cork, Cork, Ireland; 18 Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway; 19 Department of Public Health and General Practice, Norwegian University of Science and Technology; 20 Center for Health Research NW, Kaiser Permanente, Portland, OR, USA; 21 Department of Clinical and Experimental Medicine, Geriatric Endocrine Unit, University Hospital of Parma, Parma, Italy; 22 Department of Clinical and Experimental Medicine, Geriatric Endocrine Unit, University Hospital of Parma, Parma, Italy; 23 Institute of Cardiovascular and Medical Sciences, University of Glasgow, United Kingdom; 24 Department of Public Health, University of Copenhagen, Copenhagen, Denmark; 25 National Institute of Health, Bethesda, United States; 26 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands; 27 Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Nedlands, WA, United States; 28 School of Medicine and Pharmacology, University of Western Australia, Crawley, WA, Australia.

Received 31 January 2017. Accepted 26 April 2017.

**Context**

Hyperthyroidism is associated with increased fracture risk, but it is not clear if lower TSH and higher free thyroxine (FT4) in euthyroid individuals are associated with fracture risk.

**Objective**

To evaluate the association of TSH and FT4 with incident fractures in euthyroid individuals.

**Design**
Individual participant data analysis.

**Setting**
Thirteen prospective cohort studies with baseline examinations between 1981 and 2002.

**Participants**
Adults with baseline TSH 0.45-4.49 mIU/L.

**Main Outcome Measures**
Primary outcome was incident hip fracture. Secondary outcomes were any, non-vertebral, and vertebral fractures. Results were presented as hazard ratios (HR) with 95% confidence interval (CI) adjusted for age and sex. For clinical relevance, we studied TSH according to five categories: 0.45-0.99mIU/L; 1.00-1.49mIU/L; 1.50-2.49mIU/L; 2.50-3.49mIU/L; 3.50-4.49mIU/L (reference). FT4 was assessed as study-specific standard deviation increase, because assays varied between cohorts.

**Results**
During 659,059 person-years, 2,565/56,835 participants had hip fracture (4.5%; 12 studies with data on hip fracture). The pooled adjusted HR (95% CI) for hip fracture was 1.25 (1.05-1.49) for TSH 0.45-0.99mIU/L, 1.19 (1.01-1.41) for TSH 1.00-1.49mIU/L, 1.09 (0.93-1.28) for TSH 1.50-2.49mIU/L, and 1.12 (0.94-1.33) for TSH 2.50-3.49mIU/L (P for trend = 0.004). Hip fracture was also associated with FT4 (HR [95%CI] 1.22 [1.11-1.35] per one standard deviation increase in FT4). FT4 only was associated with any and non-vertebral fracture. Results remained similar in sensitivity analyses.

**Conclusions**
Among euthyroid adults, lower TSH and higher FT4 are associated with an increased risk of hip fracture. These findings may help refine the definition of optimal ranges of thyroid function tests.

In an individual participant data analysis of 13 prospective cohorts, lower thyroid-stimulating hormone within reference range and higher FT4 were associated with an increased hip fractures risk.

**INTRODUCTION**
Overt hyperthyroidism is a well-known risk factor for fracture and is associated with decreased bone mineral density (BMD) (1). We recently showed that subclinical hyperthyroidism was also associated with increased fracture incidence (2). Thyroid hormones stimulate bone turnover acting directly and indirectly on osteoclasts and osteoblasts (3). Anabolic action is net during growth, but in adults, catabolic action leads to greater bone loss and higher fracture risk (3). Thyroid hormones might also decrease muscular strength and coordination, and increase the risk of fall (4, 5). Administering TSH reduces bone resorption and increases bone formation in post-menopausal women monitored for thyroid cancer (6). Conversely, high TSH levels can degrade bone quality by increasing cortical, rather than trabecular bone.

The reference range for thyroid function tests – “euthyroidism” – is defined by the 2.5 to 97.5 percentiles in an apparently healthy population. However, the studies from which TSH reference range was derived did not exclude participants with occult or underlying disease, e.g. those with positive anti-thyroid antibodies, which might bias the reference range towards higher TSH values (7, 8). In medicine, reference ranges can be derived from normative data, as for thyroid function tests, or preferably determining levels associated with important risks or outcomes, as for lipids, or blood pressure.

TSH within the lower reference range has been associated with osteoporosis and fracture mostly in cross-sectional studies of healthy post-menopausal women, but prospective data are limited and conflicting (5, 9-13). If we can better understand the association between TSH and
health outcomes, we could make more accurate estimates of fracture risk, which would help refine thyroxine treatment targets. We hypothesized that lower TSH and higher FT4 in euthyroid participants were associated with increased risk of fractures. We therefore aimed to assess the association between TSH within the reference range, free thyroxine (FT4), and fracture risk by analyzing individual participant data (IPD) of population-based prospective cohort studies participating to the international Thyroid Studies Collaboration (2, 14).

METHODS

Data source, searches and study selection
The study protocol was registered on PROSPERO prior to study conduct (available on http://www.crd.york.ac.uk/PROSPERO; registration number: CRD42016039125).

We updated our previous systematic literature search, which had identified in Ovid (MEDLINE) and EMBASE from inception to March 2015 all prospective cohorts of adults with baseline TSH and FT4 measurement and follow-up evaluation for incident fracture (2). Additionally, we searched for studies with participants with only euthyroidism, which may have been omitted in our initial search. Our Ovid (MEDLINE) and EMBASE search (until 05/19/2016) used following medical search terms: euthyroid, euthyroidism or normal TSH and fractures or osteoporosis. After retrieving studies according to titles and abstracts, two authors (C.E.A. and D.S.) independently reviewed full-texts to confirm study eligibility. Disagreements were resolved by consensus with a third author (N.R.). We also requested unpublished fracture data from all cohorts of the Thyroid Studies Collaboration (2, 14-17). Exclusion criteria were: 1) cohorts using first-generation TSH assays because these assays were not sensitive enough; 2) studies with only participants aged <18 years; 3) studies with only participants with thyroid medication (thyroxine or anti-thyroid drugs); 4) studies with only participants with TSH outside the reference range (<0.45mIU/L or >4.49mIU/L); and, 5) studies exclusively on participants after thyroid surgery. Agreement between reviewers was 100% (K =1.00). For the IPD analysis, we included all participants aged ≥18 years at enrollment with measured TSH at baseline evaluation, and fracture assessment, as defined below, at follow-up.

Data extraction and quality assessment
If the cohorts identified met our eligibility criteria, they were invited to provide IPD. Each study was approved by its local ethics committee. All participants gave informed consent for the original studies. We collected information on demographics, anthropometrics, medications, other risk factors for fracture, history of thyroid disorders, BMD and incident fracture.

Risk of bias and study quality were independently assessed by C.E.A and D.S., using the following Newcastle-Ottawa Quality Assessment Scale items (18): 1) cohorts selection; 2) cohorts representativeness; 3) ascertainment of exposure; 4) availability of relevant confounding factors for adjustment; 5) outcome assessment based on objective fracture assessment, with adjudication procedure for fractures other than hip; 6) length of follow-up; 7) adequacy of follow-up; 8) researchers/participants/physicians blinding to thyroid values; and, 9) publication status. In sensitivity analyses, we excluded cohorts that did not meet one or more item(s).

Data synthesis and analysis

Definition of thyroid function
All included studies used a third-generation TSH radioimmunoassay. Details on the assays used for TSH and FT4 measurement are described in Supplemental Table 1. To maximize comparability, we used uniform TSH thresholds based on previously established thresholds, as
done in previous reports of the Thyroid Studies Collaboration (2, 14). We defined euthyroidism as TSH 0.45-4.49mIU/L. For clinical relevance, we separated TSH values into five categories: 0.45-0.99mIU/L; 1.00-1.49mIU/L; 1.50-2.49mIU/L; 2.50-3.49mIU/L; 3.50-4.49mIU/L. The later was used as reference category because we hypothesized, based on our previous publication (2), that lower TSH might be associated with higher fracture risk. Because of different FT4 reference ranges across studies, we used standard deviation (SD) rather than specific cut-offs. FT4 was available for all but two studies in the euthyroid range (19, 20).

**Definition of outcomes**

Our primary outcome was incident hip fracture, including femoral neck, pertrochanteric, and subtrochanteric fractures, as previously defined (2). Briefly, we excluded pathologic (i.e. associated with metastasis or rare bone disease) and periprosthetic fractures. Any, non-vertebral and clinical vertebral incident fractures were secondary outcomes. We excluded 1) vertebral fractures diagnosed with only radiologic imaging to keep focus on clinical relevance; 2) cervical and sacral vertebral fractures because fractures at these locations are usually associated with trauma rather than osteoporosis. “Any fractures” included fractures at any location, except for skull, face, ankle, finger, or toe, since these are not related to osteoporosis. “Non-vertebral fractures” was the same as “any fractures” except it excluded vertebral fractures. For any and non-vertebral fractures, we excluded cohorts that collected fracture data on only part of the skeleton. **Supplemental Table 2** describes fracture definitions by study.

**Statistical analyses**

We used a shared frailty Cox regression model with random-effects at study level to conduct an IPD meta-analysis, which used data from all included cohorts to assess the relationship of incident fractures with TSH categories and FT4, respectively (21, 22). The random-effects accounted for the between-study variation caused by different definitions of TSH reference range across the studies, incorporating the extra uncertainty in the confidence intervals. We used Schoenfeld residuals to test the proportional-hazards assumption (23). Results were presented as hazard ratios (HR) compared to the reference category. Time-to-event was defined for each outcome from baseline TSH measurement to first fracture event. We adjusted primary analyses for age and sex, and then for other risk factors for fracture (body mass index [BMI], smoking, and history of diabetes), because they might mediate the association between thyroid function tests and fractures. We conducted following predefined sensitivity analyses: 1) excluding participants with thyroid medication (thyroxine or anti-thyroid medication) at baseline; 2) excluding participants with thyroid-altering medication at baseline (thyroid medication, oral corticosteroids, amiodarone, iodine); 3) excluding participants with anti-fracture medication at baseline (bisphosphonate, calcitonin, selective estrogen receptor modulator, parathyroid hormone); 4) including only studies with formal fracture adjudication; 5) including only studies that uniformly defined fractures (except for hip fracture, since it has a common definition and is rarely reported in error); 6) excluding cohorts with loss to follow-up rates >5%; 7) excluding participants who developed overt or subclinical thyroid dysfunction over time; 8) including only participants with TSH remaining within the reference range during follow-up; and 9) further adjusting for BMD, which reflects bone loss and may be a potential mediator between TSH or FT4 and incident fractures. In this last analysis, we used BMD as a continuous variable, and included only studies that used dual energy X-ray absorptiometry (DXA) devices with femoral neck BMD for hip fractures (available for six studies) (5, 10, 14, 24-26), lumbar spine BMD for vertebral fractures (available for one study) (10), and whole body BMD for any fractures.
(available for one study) (10). We conducted predefined stratified analyses by sex, age (<75 versus ≥75 years), and duration of follow-up (<5 versus ≥5 years).

For the FT4 analysis, we used the whole range of FT4 values including only participants with TSH within the reference range. FT4 values were converted to ng/mL (12.87 pmol/L = 1 ng/mL). We used study-specific SD to assess fracture risk per one SD increase in FT4 because FT4 assays varied between cohorts (14). We performed the same sensitivity and stratified analyses as for TSH.

We used STATA release 13.1 for all analyses (StataCorp LP, College Station, Texas). All tests were two-sided, at a 0.05 level of significance.

RESULTS

Our updated literature search identified nine additional reports (Supplemental Figure 1) (2). Eight of them concerned studies already identified in our previous search (2). The newly identified study (Study of Osteoporotic Fractures) (19) agreed to participate. We excluded the Nagasaki Adult Health Study (27), because it used first-generation TSH assays, which have a low functional sensitivity (1mIU/L) (28). For the same reason, this study had been included in our previous work (2) only in the analysis on subclinical hypothyroidism, but not on subclinical hyperthyroidism (16). We included thirteen studies (5, 10, 14, 17, 19, 25, 26, 29-34) from the USA, Europe, and Australia with 61,959 participants, and a median duration of follow-up of 12.1 years (interquartile range [IQR] 8.5-12.9), totaling 659,059 person-years. Median age was 64 (range 18-102) with 60.5% women (Table 1). Median (IQR) TSH was 1.60 mIU/L (1.10-2.30); 3.1% of participants used thyroid medication at baseline and 5.5% during follow-up; 17.7% had a TSH 0.45-0.99 mIU/L, 24.8% 1.00-1.49 mIU/L, 37.4% 1.50-2.49 mIU/L, 14.2% 2.50-3.49 mIU/L and 5.9% 3.50-4.49 mIU/L. Hip fracture occurred in 2,565 participants (4.5%; 12 studies), any fracture in 2,333 (8.9%; 9 studies), non-vertebral fracture in 1,874 (8.5%; 9 studies), and vertebral fracture in 263 (1.3%; 7 studies). Overall quality was good (Supplemental Table 3): one study reported loss to follow-up >5% (5), four did not perform formal fracture adjudication (25, 29, 32, 33), and three had not published fracture data in a separate manuscript (17, 29, 33).

Tests of the proportional-hazards assumption on the basis of Schoenfeld residuals indicated that assumptions were met for all analyses (P > 0.11 for all).

Thyroid function and hip fractures

Compared with the reference group (TSH 3.50-4.49 mIU/L), pooled age- and sex-adjusted HR (95% CI) for hip fracture was 1.25 (1.05-1.49) for TSH 0.45-0.99 mIU/L, 1.19 (1.01-1.41) for TSH 1.00-1.49 mIU/L, 1.09 (0.93-1.28) for TSH 1.50-2.49 mIU/L, and 1.12 (0.94-1.33) for TSH 2.50-3.49 mIU/L (P for trend 0.004, Figure 1). After adjusting for BMI, smoking status, and history of diabetes, HR (95% CI) was 1.24 (1.03-1.49) for TSH 0.45-0.99 mIU/L compared with the reference group, while HR (95% CI) for TSH 1.00-1.49 mIU/L was somewhat attenuated and no longer statistically significant (1.15 [0.97-1.38]). The risk of hip fracture in participants with TSH 0.45-0.99 mIU/L remained significantly higher in all sensitivity analyses, and was even higher after adjusting for femoral neck BMD (Table 2). For TSH 1.00-1.49 mIU/L, the risk of hip fractures remained significantly higher in all sensitivity analyses, except after adjusting for femoral neck BMD, or after excluding participants with thyroid-altering medication at baseline. This association remained not significant for TSH 1.50-2.49 mIU/L, or TSH 2.50-3.49 mIU/L. We found no significant interaction for sex, age, or duration of follow-up (Supplemental Figure 2), although confidence intervals were larger and point estimates smaller for age <75 years and
follow-up <5 years. Conversely, there was significant interaction for publication status with a
HR (95% CI) of 1.35 (1.13-1.61) for the ten studies that published risk of hip fracture associated
with thyroid function tests in a separate manuscript, and 0.44 (0.21-0.90) for the two studies (17,
29) that did not previously publish hip fracture data associated with thyroid function tests in a
separate article (P for interaction 0.0001, Supplemental Table 4).

The HR (95% CI) for hip fracture was 1.24 (1.12-1.37) per one SD increase in FT4 (Figure
2). We found no significant interaction with sex, age, duration of follow-up, or publication status
of hip fracture data (Figure 2, Supplemental Table 4), although point estimate was smaller
when follow-up was <5 years. All sensitivity analyses yielded similar results (Table 2). In the
25,760 participants of the five cohorts with available data on thyroid function tests during
follow-up (25, 29, 31-33), 146 (0.6%) participants developed subclinical hyperthyroidism and 46
(0.2%) overt hyperthyroidism. When we included only endogenous forms of thyroid dysfunction
(i.e., participants without thyroxine use at baseline, N=25,049), 102 (0.4%) and 25 (0.1%)
participants developed subclinical and overt hyperthyroidism, respectively. The HR (95% CI) for
hip fracture for TSH 0.45-0.99mIU/L compared with the reference group was 1.70 (1.13-2.57) in
the sensitivity analysis including only participants with TSH remaining within the reference
range (four cohorts with data on hip fracture and thyroid function tests during follow-up) (25, 29,
31, 32).

Thyroid function and any, non-vertebral, and vertebral fractures
For all TSH categories when compared with the reference group, we found no significant
association for any, non-vertebral, or vertebral fractures (Supplemental Table 5). The HR (95%
CI) per one SD increase in FT4 was 1.08 (1.02-1.15) for any fracture, and 1.10 (1.03-1.18) for
non-vertebral fracture. These associations remained significant in most sensitivity analyses
(Table 3), except when adjusting for BMD. Association between FT4 and vertebral fracture was
not statistically significant, possibly because of the lower number of data (Table 3). We found
no significant interaction in the analyses stratified by sex, age, duration of follow-up, or
publication status for any of these fracture outcomes (Table 3, Supplemental Table 4).

DISCUSSION
In this analysis of 61,959 euthyroid participants of thirteen prospective cohorts with 659,059
person-years of follow-up, lower TSH levels within the reference range were associated with
increased risk of hip fracture, and higher FT4 levels with increased risk of hip, any, and non-
vertebral fracture.

While overt and subclinical hyperthyroidism have been associated with increased fracture
risk (1, 2), previous studies on the relationship between TSH within the reference range and
fracture risk had conflicting results. The Clalit Health Services, a large historical cohort study,
found a borderline increased incidence of hip fracture with TSH 0.35-1.6mIU/L when compared
with TSH 1.7-2.9mIU/L, but in women only (odds ratio [95%CI] 1.28 [1.03-1.59]), while the
association with other osteoporotic fractures was not statistically significant (11). A small cross-
sectional study (N=129) found an association between low TSH and vertebral fracture (12). The
Cardiovascular Health Study found no significant association between TSH within the reference
range or FT4 assessed as continuous variables and hip fracture (13), but, consistent with our
findings, curves bent with an increased fracture risk for TSH <1.5mIU/L and for FT4
>1.4ng/mL. Our thorough IPD analysis across multiple prospective cohorts confirms the
association between low TSH and hip fractures, and an association between high FT4 and all but
vertebral fractures in participants with TSH within the reference range, suggesting that even a
modest increase in thyroid hormone levels among euthyroid adults is associated with higher fracture risk.

Our study was strengthened, first, by an IPD analysis that allowed us to standardize the definitions of predictors and outcomes, adjust for similar potential confounders, and avoid aggregation bias for subgroup analyses. This was the best way to perform time-to-event analysis. Second, our study is the largest to assess fracture risk in prospective cohorts with TSH within the reference range. Third, we included all international prospective cohorts available on this topic, since all the studies we identified agreed to participate.

Our study had several limitations. First, our population consisted mostly of Caucasians and included few young adults. Second, thyroid function tests were performed only at baseline in most cohorts, so we may have included adults who later developed subclinical or overt thyroid dysfunction. However, our sensitivity analysis including only participants with persistent TSH within the reference range yielded an even stronger association between low TSH and hip fracture. In addition, other participants may have had a non-thyroidal illness, a potential cause of suppressed TSH; however, the prevalence of non-thyroidal illness was likely low as we only studied community-dwelling adults. Third, fractures were adjudicated in nine of thirteen cohorts, and we could not uniformly define each fracture type across all cohorts. Nevertheless, sensitivity analyses limited to cohorts with the most uniform fracture definitions or adjudicated fracture yielded similar results. Fourth, we did not know fracture mechanism, but we excluded pathological fractures and fracture locations typically not associated with osteoporosis to reduce bias related to traumatic fractures. Fifth, data on fractures other than hip location were available in a more limited number of studies, reducing the number of outcomes and the related power to identify associations. Sixth, we had no information on other factors that may have influenced bone integrity or accounted for variations in circulating TSH or FT4, such as nutrition or deiodinase activities. Finally, thyroid antibodies were not systematically measured and their potential impact on bone metabolism could not be assessed.

Our findings may have two important clinical implications. First, TSH reference range is still a matter of debate (35). TSH reference range was indeed defined in a population that included persons with occult or underlying thyroid disease (7, 8). TSH between 0.4 and 2.5mIU/L is associated with a lower incidence of thyroid dysfunction (36), but previous studies showed various adverse outcomes associated with subclinical thyroid dysfunction (14-16), and with TSH at both extremities of the reference range (e.g. higher risk of cardiovascular disease with high TSH/low FT4, and higher risk of fractures, osteoporosis, and dementia with low TSH/high FT4) (9, 13). There may be optimal values of thyroid function tests within the reference range. Second, similar to previous studies showing stronger association of adverse outcomes with FT4 than TSH (37, 38), FT4 was associated with hip, any and non-vertebral fracture, while TSH was associated only with hip fracture. TSH and thyroid hormones may act differently on peripheral organs, including bones: TSH may act on osteoblasts and osteoclasts via specific receptors (3), while thyroid hormones may act on target tissues via nuclear receptors controlled locally by deiodinases (3, 39, 40). This may explain why TSH and FT4 are associated with different fracture types. FT4 may therefore help evaluate osteoporosis and fracture risk, which is now usually done with the World Health Organization FRAX score, but future studies should determine if adding FT4 improves clinical accuracy of this score. Of note, FT4 was not significantly associated with vertebral fracture. One explanation may be that FT4 acts differently on vertebral bone. It may however also be due to lack of power, as we could include about ten times fewer vertebral than other fractures (Table 3).
We may have expected a stronger association of fracture risk with TSH and FT4, respectively, after excluding participants with thyroid medication at baseline, but the risk was only slightly increased, probably because of the low number of participants with thyroid medication at baseline (N=1897, 3%).

In conclusion, analyzing individual data of 61,959 adults from thirteen large prospective cohorts, we found that TSH at the lower extremity of the reference range was associated with higher risk of hip fractures, and high FT4 with higher risk of hip, any, and non-vertebral fractures. Our findings may help refine the current definition of optimal thyroid function. Meanwhile, clinicians should be aware that lower TSH and higher FT4, even within the reference range, are associated with an increased risk of hip fracture.

ACKNOWLEDGMENTS
Dr. Gussekloo, Dr. Kearney, Dr. Rodondi, Dr. Stott and Dr. Westendorp received funding for a randomized controlled trial on subclinical hypothyroidism (TRUST trial) from the European Commission FP7-HEALTH-2011, Specific Programme “Cooperation” – Theme “Health” Investigator-driven clinical trials for therapeutic interventions in elderly populations (Proposal No: 278148-2). Dr. Collet was supported by research grants from Swiss National Science Foundation (P3SMP3-155318 and PZ00P3-167826) during the conduct of the study. Dr. Peeters reports lecture and/or advisory board fees from Genzyme B.V., EISAI, IPSEN, and Goodlife Fertility; and grant support from Veracyte, all outside of the submitted work. Dr. Eastell reports grants from Amgen, Department of Health, AstraZeneca, Immunodiagnostic Systems (IDS), Canadian Institutes of Health Research, National Osteoporosis Society, ARUK/MRC Centre of Excellence in Musculoskeletal Ageing Research, National Institute for Health Research, Cancer Research UK, MRC/AZ Mechanisms of Disease Call during the conduct of the study, and personal fees from Novartis, Roche, Alexion, Otsuka, Teijan Pharma Limited, IBMS, Merck, Amgen, Eli Lilly, Foundation of the National Institutes of Health (NIH), Endocrine Society, Johnson & Johnson, SPD Development, Fonterra Brands, Janssen Research, Ono Pharma, Immunodiagnostic Systems, Alere (Unipath), Chronos, GSK Nutrition, Radius Health, European Calcified Tissue Society, Efficacy & Mechanism Evaluation Board of the Medical Research Council, IOF CSA, all outside of the submitted work. Dr. Westendorp is supported by the Netherlands Organization for Scientific Research (NGI/NOW 911-03-016). No other conflicts were reported.

Corresponding author and person to whom reprint requests should be addressed: Nicolas Rodondi, Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland; e-mail: nicolas.rodondi@insel.ch; Phone +41 31 632 41 63; Fax +41 31 632 88 85.

Primary funding source: Swiss National Science Foundation (SNSF 320030-150025).

FINANCIAL SUPPORT
Primary funding source: Swiss National Science Foundation (SNSF 320030-150025 to Prof. Rodondi).

The Busselton Health Study had no financial support to disclose.

The Cardiovascular Health Study (CHS) was supported by contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from
the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The European Prospective Investigation of Cancer (EPIC)-Norfolk Study was supported by research grants from the Medical Research Council UK and Cancer Research UK.

The Health, Aging and Body Composition (Health ABC) Study was supported by National Institute on Aging (NIA) Contracts N01-AG-6-2101; N01-AG-6-2103; N01-AG-6-2106; NIA grant R01-AG028050, and NINR grant R01-NR012459. This research was funded in part by the Intramural Research Program at the National Institute on Aging.

The InCHIANTI Study was supported as a target project ICS 110.1]RS97.71 by the Italian Ministry of Health, and in part by the U.S. National Institute on Aging, contracts 263-MD-9164-13 and 263-MD-821336.

The Nord-Trøndelag Health (HUNT) Study is a collaborative effort of HUNT Research Center (Faculty of Medicine, NTNU, Norwegian University of Science and Technology), the Norwegian Institute of Public Health, Central Norway Health Authority, and the Nord-Trøndelag County Council. Thyroid function testing in the HUNT Study was financially supported by WallacOy (Turku, Finland). Data were provided by HUNT Research Centre and by Nord-Trøndelag Hospital Trust.

The Leiden 85-plus Study was partly funded by the Dutch Ministry of Health, Welfare and Sports.

The Osteoporotic Fractures in Men Study (MrOS) was supported by U.S. National Institutes of Health funding. The following institutes provide support: the National Institute on Aging (NIA), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Advancing Translational Sciences (NCATS), and NIH Roadmap for Medical Research, under the following grant numbers: U01 AG027810; U01 AG042124; U01 AG042139; U01 AG042140; U01 AG042143; U01 AG042145; U01 AG042168; U01 AR066160; and, UL1 TR000128.

The Osteoporosis and Ultrasound Study (OPUS) was supported by Sanofi-Aventis, Eli Lilly, Novartis, Pfizer, Proctor and Gamble Pharmaceuticals, and Roche.

The original PROSPER Study was supported by an unrestricted, investigator-initiated grant from Bristol-Myers Squibb.

The Rotterdam Study was funded by the following: Erasmus MC and Erasmus University, Rotterdam, the Netherlands; the Netherlands Organisation for Scientific Research (NWO); the Netherlands Organisation for the Health Research and Development (ZonMW); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science; the Dutch Ministry for Health, Welfare and Sports; the European Commission (DG XII); and, the Municipality of Rotterdam.

The Sheffield Study was supported by Arthritis Research UK.
The Study of Osteoporotic Fractures (SOF) was supported by National Institutes of Health funding. The National Institute on Aging (NIA) provides support under the following grant numbers: R01 AG005407, R01 AR35582, R01 AR35583, R01 AR35584, R01 AG005394, R01 AG027574, and R01 AG027576.

AUTHORS CONTRIBUTIONS:

Carole Aubert and Prof Rodondi had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Aubert, Bauer, Rodondi.

Literature search and review: Aubert, Segna.

Acquisition, analysis, or interpretation of data: Åsvold, Aubert, Blum, Bremner, Cappola, Ceresini, den Elzen, Gussekloo, Kearney, Khaw, Peeters, Stott, Walsh, Westendorp, Rodondi.

Drafting of the manuscript: Aubert.

Critical revision of the manuscript for important intellectual content: Åsvold, Aujesky, Bauer, Blum, Bremner, Cappola, Ceresini, Collet, da Costa, den Elzen, Ferrucci, Fink, Floriani, Gussekloo, Hoff, Kearney, Khaw, Langhammer, Le Blanc, Luben, Peeters, Rivadeneira, Rodondi, Segna, Stott, Syrogiannouli, Uitterlinden, Walsh, Westendorp, Williams, Rodondi.

Statistical analyses: Aubert, Blum, da Costa, Rodondi, Syrogiannouli.

Administrative, technical, or material support: Bauer, Ceresini, den Elzen, Ferrucci, Gussekloo, Khaw, Peeters.

Study supervision: Bauer, Rodondi.

Registration: The protocol was published on PROSPERO (registration number CRD42016039125).

Conflicts of interest statement: nothing to disclose.

REFERENCES


**Figure 1.** Title: Risk of hip fracture according to thyroid-stimulating hormone categories. Legend: Abbreviations: CI, confidence interval; HR; hazard ratio; No., number; TSH, thyroid-stimulating hormone. Data on hip fractures were available for 12 studies (all except PROSPER).

**Figure 2.** Title: Risk of hip fracture per one standard deviation increase in free thyroxine, overall and stratified by sex, age, and duration of follow-up. Legend: Abbreviations: CI, confidence interval; FT4, free thyroxine; Health ABC, Health, Aging and Body Composition; HR; hazard ratio; No., Number; SOF, Study of Osteoporotic fractures, TSH, thyroid-stimulating hormone. The analysis stratified for sex was adjusted for age. All other analyses were adjusted for age (as a continuous variable) and sex. FT4 was measured in all studies but SOF and Health ABC Study (FT4 not measured in participants with TSH within the reference range). Data on hip fractures were available for 10 studies with measured FT4 (all except PROSPER).

**Table 1.** Study population and baseline characteristics of the participants in the 13 included studies (n=61,959)

<table>
<thead>
<tr>
<th>Study name, place</th>
<th>Description of study population</th>
<th>Number of participants</th>
<th>Age, median (range)</th>
<th>Women, No. (%)</th>
<th>TSH, median, mIU/L</th>
<th>Thyroid medication at baseline, No. (%)</th>
<th>Thyroid medication during follow-up, No. (%)</th>
<th>Start of follow-up, year</th>
<th>Duration of follow-up, median (IQR), years</th>
<th>Person-years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busselton Health Study, Australia (29)</td>
<td>Adults</td>
<td>1,907</td>
<td>51 (18-90)</td>
<td>919 (48.2)</td>
<td>1.42</td>
<td>10 (0.5)</td>
<td>15 (0.8)</td>
<td>1981</td>
<td>20.0 (17.6-20.0)</td>
<td>33,281</td>
</tr>
<tr>
<td>CHS, USA (4 communities) (25)</td>
<td>Adults with Medicare eligibility</td>
<td>2,853</td>
<td>71 (65-100)</td>
<td>1,694 (59.4)</td>
<td>2.03</td>
<td>145 (5.1)</td>
<td>299 (10.5)</td>
<td>1989-1990</td>
<td>12.9 (7.5-18.9)</td>
<td>36,466</td>
</tr>
<tr>
<td>EPIC-Norfolk Study, England (30)</td>
<td>Adults aged 45-79y</td>
<td>11,986</td>
<td>58 (40-78)</td>
<td>6,365 (53.1)</td>
<td>1.70</td>
<td>275 (2.3)</td>
<td>NA</td>
<td>1995-1998</td>
<td>12.4 (11.7-13.3)</td>
<td>142,951</td>
</tr>
<tr>
<td>Health ABC Study, USA (4 communities) (14)</td>
<td>Adults aged 70-79y with Medicare eligibility</td>
<td>2,347</td>
<td>74 (69-81)</td>
<td>1,165 (49.6)</td>
<td>1.99</td>
<td>177 (7.5)</td>
<td>383 (13.9)</td>
<td>1997</td>
<td>12.7 (8.1-13.2)</td>
<td>24,794</td>
</tr>
<tr>
<td>HUNT Study, Norway (31)</td>
<td>Adults</td>
<td>31,388</td>
<td>57 (19-99)</td>
<td>21,186 (67.5)</td>
<td>1.60</td>
<td>999 (3.2)</td>
<td>NA</td>
<td>1995-1997</td>
<td>12.2 (11.6-12.8)</td>
<td>345,517</td>
</tr>
<tr>
<td>InCHIANTI Study, Italy (17)</td>
<td>Adults aged 65y</td>
<td>1,066</td>
<td>71 (21-102)</td>
<td>590 (55.3)</td>
<td>1.38</td>
<td>17 (1.6)</td>
<td>28 (2.6)</td>
<td>1998</td>
<td>9.1 (7.8-9.3)</td>
<td>8,562</td>
</tr>
<tr>
<td>Leiden 85-Plus Study, The Netherlands (32)</td>
<td>Adults aged 85y</td>
<td>456</td>
<td>85 (85-85)</td>
<td>293 (64.3)</td>
<td>1.66</td>
<td>6 (1.3)</td>
<td>11 (2.4)</td>
<td>1997-1999</td>
<td>4.8 (2.2-8.1)</td>
<td>2,411</td>
</tr>
</tbody>
</table>
MrOS, USA (6 clinical centers) (10)
Men aged ≥65y 1,410 73 (65-99) All men 1.97 83 (5.9) 98 (6.9) 2000-2002 11.1 (8.1-11.8) 13,568

OPUS, Germany, France, UK (5)**
Women aged 20-80y 1,205 63 (20-80) All women 0.96 0 (0.0) NA 1999-2001 6.0 (5.8-6.3) 7,179

PROSPER, Ireland, Scotland, The Netherlands (33)
Older adults at high cardiovascular risk 5,124 75 (69-83) 2,527 (49.3) 1.75 135 (2.6) 163 (3.2) 1997-1999 3.2 (3.0-3.5) 15,833

Rotterdam Study, The Netherlands (34)
Adults aged ≥55y 1,611 68 (55-93) 957 (59.4) 1.54 21 (1.3) NA 1989-1992 15.2 (10.4-16.2) 21,130

Sheffield Study, England (26)
Women aged 50-80y 291 63 (50-86) All women 2.00 2 (0.7) 9 (3.1) 1990-1991 10.0 (5.5-10.1) 2,301

SOF, USA (4 clinical centers) (19)††
Women >65y 314 71 (65-88) All women 1.50 15 (4.8) NA 1986-1998 14.3 (9.8-19.8) 4,433

Overall 13 cohorts 61,959 64 (18-102) 37,506 (60.5) 1.60 1,885 (3.1) 831 (5.5) 1981-2002 12.1 (8.5-12.9) 659,059 ‡‡

Abbreviations: CHS, Cardiovascular Health Study; EPIC, European Prospective Investigation of Cancer; Health ABC, Health, Aging and Body Composition; HUNT, Nord-Trøndelag Health Study; InCHIANTI, Invecchiare in Chianti; IQR, interquartile range; MrOS, Osteoporotic Fractures in Men Study; No., number; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone; UK, United Kingdom; USA, United States of America; y, years.

* We excluded participants younger than 18y.
† Thyroid medication was defined as thyroxine or anti-thyroid medication.
‡ Data on thyroid medication at baseline was missing for 255 participants in the HUNT Study, 59 participants in the MrOS, one participant in the Rotterdam Study, four participants in the SOF and seven participants in the Health ABC Study.
§ Data on thyroid mediation at follow-up was missing for 243 participants in the MrOS, 96 participants in the InCHIANTI Study, 45 participants in the Sheffield Study, and all participants in the HUNT Study, EPIC-Norfolk Study, Rotterdam Study, OPUS and SOF.
‖ Duration of follow-up was defined as the maximum duration of follow-up that was available, i.e. the time to the first hip (or any if unavailable) fracture or censor date/death.
¶ We included participants excluded from the original article of the HUNT Study (participants <40y, with previous fracture and/or with previous thyroid disease), which explains the different number of the sample.
** We included only the thyroid hormone sub-study of the OPUS, which excluded participants on thyroid medication.
†† We included only a subsample of the SOF, i.e., the participants with TSH measurement at baseline.
‡‡ It was calculated as time to hip fracture; for the PROSPER, it was calculated as time to any fracture, since data on hip fracture was unavailable.

Table 2: Sensitivity analyses for the risk of hip fracture according to thyroid-stimulating hormone and free thyroxine

<table>
<thead>
<tr>
<th>Analysis by TSH category*</th>
<th>Analysis by SD increase in FT4†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of events/ participants</strong></td>
<td><strong>Hazard ratio (95% CI)</strong> ‡</td>
</tr>
<tr>
<td>Main analysis</td>
<td>610/13,390</td>
</tr>
<tr>
<td>Medication use</td>
<td></td>
</tr>
<tr>
<td>Excluding participants with thyroid medication at baseline‖</td>
<td>557/12,728</td>
</tr>
<tr>
<td>Excluding participants with thyroid-altering medication at baseline§</td>
<td>542/12,396</td>
</tr>
<tr>
<td>Excluding participants with anti-fracture medication at baseline**</td>
<td>605/12,739</td>
</tr>
<tr>
<td>Definition of fracture</td>
<td></td>
</tr>
<tr>
<td>Including only studies with formal fracture adjudication††</td>
<td>496/12,048</td>
</tr>
<tr>
<td>Other</td>
<td>Excluding one study with loss to follow-up &gt;5%‡‡</td>
</tr>
</tbody>
</table>
Further adjusting for femoral neck BMD at baseline§§

| BMD | 942/2,020 | 1.68 (1.08-2.64) || 142/4,147 | 1.22 (1.01-1.47) |

Abbreviations: BMD, bone mineral density; CHS, Cardiovascular Health Study; CI, confidence interval; EPIC, European Prospective Investigation of Cancer; FT4, free thyroxine; Health ABC, Health, Aging and Body Composition; HUNT, Nord-Trøndelag Health Study; InCHIANTI, Invecchiare in Chianti; MrOS, Osteoporotic Fractures in Men Study; No., number; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SD, standard deviation; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone.

All analyses were adjusted for age (as a continuous variable) and sex. Data for hip fractures were available for 12 cohorts (all but PROSPER).

* We present a selected analysis for the TSH category 0.45-0.99mIU/L compared with the reference category (TSH 3.50-4.99mIU/L). No. are for participants in these both TSH categories only.
† FT4 was measured in all studies but the SOF and the Health ABC Study (FT4 not measured in participants with TSH within reference range).
§ Hazard ratios are for TSH 0.45-0.99mIU/L, compared with the reference group 3.50-4.99mIU/L.
¶ Thyroid medication was defined as thyroxine or anti-thyroid medication.
** Anti-fracture medication was defined as bisphosphonate, calcitonin, selective estrogen receptor modulator, or parathyroid hormone.
†† EPIC-Norfolk Study, HUNT Study, InCHIANTI Study, MrOS, OPUS, Rotterdam Study, Sheffield Study, Health ABC Study, and SOF (Health ABC Study and SOF only in the TSH analysis).
§§ Femoral neck BMD at baseline was available in following studies: CHS, MrOS, Rotterdam Study, Sheffield Study, OPUS, Health ABC Study.
|| Participants within the TSH category 3.50-4.49mIU/L had lower femoral neck BMD at baseline than participants within the TSH category 0.45-1.50mIU/L (mean [SD]: 0.77g/cm² [0.16] versus 0.79 g/cm² [0.15], respectively, \( P = 0.002 \)), which explains the higher hazard ratio after adjusting for femoral neck BMD at baseline.

Table 3. Sensitivity and stratified analyses for the risk of any, non-vertebral, and vertebral fractures, per one standard deviation increase in free thyroxine

<table>
<thead>
<tr>
<th>Medication use</th>
<th>Any fracture*</th>
<th>Non-vertebral fracture†</th>
<th>Vertebral fracture‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excluding participants with thyroid medication at baseline§</td>
<td>1,552/22,440</td>
<td>1.09 (1.02-1.16)</td>
<td>1,240/18,697</td>
</tr>
<tr>
<td>Excluding participants with thyroid-altering medication at baseline</td>
<td></td>
<td>1,537/21,976</td>
<td>1.09 (1.02-1.15)</td>
</tr>
<tr>
<td>Excluding participants with anti-fracture medication at baseline¶</td>
<td>1,622/22,927</td>
<td>1.08 (1.02-1.15)</td>
<td>1,263/19,038</td>
</tr>
<tr>
<td>Definition of fracture</td>
<td>Including only studies with formal fracture</td>
<td>1,026/15,805</td>
<td>1.11 (1.02-1.19)</td>
</tr>
<tr>
<td>adjudication*†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Including only studies with most uniform definition of fracture\‡‡</td>
<td>1,155/19,728</td>
<td>1.06 (0.99-1.14)</td>
<td>685/14,461</td>
</tr>
</tbody>
</table>

**Other**

Further adjusting for BMI, smoking status, and diabetes mellitus

| Including only studies with most uniform definition of fracture\‡‡ | 1,591/22,536 | 1.21 (1.00-1.46) | 1,140/17,562 | 1.09 (1.01-1.18) | 126/17,290 | 1.03 (0.83-1.27) |

Excluding studies with loss of follow-up rate >5%

| Including only studies with most uniform definition of fracture\‡‡ | NA | NA | 1,174/17,981 | 1.13 (1.04-1.22) | NA | NA |

**Other**

Further adjusting for BMI, smoking status, and diabetes mellitus

| Including only studies with most uniform definition of fracture\‡‡ | NA | NA | NA | NA | 39/1,399 | 0.96 (0.68-1.36) |

**BMD**

Further adjusting for lumbar spine BMD at baseline\‡‡

| Including only studies with most uniform definition of fracture\‡‡ | NA | NA | NA | NA | 39/1,399 | 0.96 (0.68-1.36) |

Excluding studies with loss of follow-up rate >5%

| Including only studies with most uniform definition of fracture\‡‡ | 183/1,399 | 0.89 (0.75-1.04) | NA | NA | NA | NA |

---

**STRATIFIED ANALYSES**

**Stratified for sex**

<table>
<thead>
<tr>
<th>Women</th>
<th>1,013/11,321</th>
<th>1.11 (1.03-1.19)</th>
<th>827/10,075</th>
<th>1.10 (1.01-1.20)</th>
<th>628,679</th>
<th>1.12 (0.83-1.51)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>616/11,656</td>
<td>1.05 (0.95-1.15)</td>
<td>446/9,026</td>
<td>1.08 (0.96-1.22)</td>
<td>679,032</td>
<td>1.00 (0.75-1.33)</td>
</tr>
</tbody>
</table>

**P-value for interaction**

| NA | 0.39 | NA | 0.79 | NA | 0.61 |

**Stratified for age**

<75 years at baseline

| NA | 1.04 (1.02-1.19) | 955/17,144 | 1.10 (1.01-1.20) | 87/15,917 | 0.96 (0.74-1.25) |

≥75 years at baseline

| NA | 1.06 (0.96-1.16) | 318/1,957 | 1.10 (0.97-1.25) | 42/1,794 | 1.25 (0.88-1.76) |

**P-value for interaction**

| NA | 0.47 | NA | 0.99 | NA | 0.25 |

**Stratified for duration of follow-up**

<5 years

| 446/5,920 | 1.04 (0.93-1.15) | 47/888 | 0.82 (0.59-1.14) | 7/888 | 0.60 (0.26-1.37) |

≥5 years

| 1,183/17,057 | 1.09 (1.02-1.18) | 1,226/18,213 | 1.10 (1.03-1.18) | 122/16,823 | 1.07 (0.87-1.33) |

**P-value for interaction**

| NA | 0.39 | NA | 0.07 | NA | 0.18 |

---

**Abbreviations:** BMD, bone mineral density; BMI, body mass index; CI, confidence interval; EPIC, European Prospective Investigation of Cancer; FT4, free thyroxine; Health ABC, Health, Aging and Body Composition; InCHIANTI, Invecchiare in Chianti Study; MrOS, Osteoporotic Fractures in Men Study; NA, not appropriate; No., number; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SD, standard deviation; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone.

All analyses were adjusted for age (as a continuous variable) and sex; FT4 was measured in all studies but SOF and Health ABC Study (FT4 not measured in participants with TSH within the reference range).

Hazard ratios are per one standard deviation increase in FT4.

\* Data on any fractures were available for 7 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, PROSPER, Rotterdam Study, Busselton Health Study).

\† Data on non-vertebral fractures were available for 7 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, OPUS).

\‡ Data on vertebral fractures were available for 5 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study). Vertebral fracture was defined as a clinical symptomatic dorsal or lumbar fracture.

§ Thyroid medication was defined as thyroxine or anti-thyroid medication.

|| Thyroid-altering medication included oral corticosteroid, amiodarone, iodine, thyroxine, or anti-thyroid drug.
Anti-fracture medication was defined as bisphosphonate, calcitonin, selective estrogen receptor modulator, or parathyroid hormone.

** EPIC-Norfolk Study, InCHIANTI Study, MrOS, OPUS, Rotterdam Study, Sheffield Study.

†† EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, MrOS, PROSPER.

‡‡ Lumbar spine BMD was available in MrOS only.

§§ Whole body BMD was available in MrOS only.
<table>
<thead>
<tr>
<th>TSH categories</th>
<th>No. with fracture</th>
<th>No. of participants</th>
<th>Hazard ratio (95% CI)</th>
<th>P-value for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.50-4.49</td>
<td>181</td>
<td>3,248</td>
<td>1 (Reference)</td>
<td>0.004</td>
</tr>
<tr>
<td>2.50-3.49</td>
<td>446</td>
<td>7,980</td>
<td>1.12 (0.94-1.33)</td>
<td></td>
</tr>
<tr>
<td>1.50-2.49</td>
<td>942</td>
<td>21,231</td>
<td>1.09 (0.93-1.28)</td>
<td></td>
</tr>
<tr>
<td>1.00-1.49</td>
<td>597</td>
<td>14,183</td>
<td>1.19 (1.01-1.41)</td>
<td></td>
</tr>
<tr>
<td>0.45-0.99</td>
<td>429</td>
<td>10,142</td>
<td>1.25 (1.05-1.49)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>No of participants</td>
<td>Hazard ratio (95% CI)</td>
<td>P-value for interaction</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>--------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>495</td>
<td>1.217 (1.12-1.31)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>&lt; 75 years</td>
<td>213</td>
<td>1.27 (1.11-1.46)</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>≥ 75 years</td>
<td>282</td>
<td>1.26 (1.07-1.50)</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>170</td>
<td>1.23 (1.08-1.40)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>329</td>
<td>1.18 (1.02-1.38)</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

**Duration of follow-up**
- < 45 years: 47
- 45-74 years: 329
- ≥ 75 years: 382

**No with fracture**
- Overall: 542
- < 75 years: 213
- ≥ 75 years: 329

**HR (95% CI)**
- Overall: 1.217 (1.12-1.31)
- < 75 years: 1.27 (1.11-1.46)
- ≥ 75 years: 1.26 (1.07-1.50)

**P-value for interaction**
- Overall: 0.04
- < 75 years: 0.09
- ≥ 75 years: 0.07
- Women: 0.02
- Men: 0.03
Thyroid function within the reference range and fracture risk: individual participant analysis of prospective cohorts

Supplemental Data
### Supplemental Table 1. Description of assays used for TSH and FT4 measurement

<table>
<thead>
<tr>
<th>Study</th>
<th>TSH assay</th>
<th>TSH reference range (mIU/L)</th>
<th>FT4 assay</th>
<th>FT4 reference range pmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busselton Health Study</td>
<td>Immulite 2000 chemiluminescent analyzer (Diagnostic Products Corporation, Los Angeles, Calif.), FS 0.02IU/L, CV 7.6%</td>
<td>0.4-4.0</td>
<td>Immulite 2000 chemiluminescent analyzer (Diagnostic Products Corporation, Los Angeles, Calif.), CV 9.6%</td>
<td>9-23</td>
</tr>
<tr>
<td>CHS</td>
<td>Chemiluminescent immunometric assay (LumaTag hTSH; Nichols Institute, San Juan Capistrano, California), FS 0.008mIU/L</td>
<td>0.1-4.5</td>
<td>Direct monoclonal antibody assay (Amerlex-MAB; Amersham International, Amersham, England)</td>
<td>4.9-22</td>
</tr>
<tr>
<td>EPIC-Norfolk Study</td>
<td>AutoDelfia time resolved fluoroimmunoassay kits (Wallac, Finland), FS 0.03mIU/L, CV &lt;3.5%</td>
<td>0.4-4.0</td>
<td>AutoDelfia time resolved fluoroimmunoassay kits (Wallac, Finland), FS 2.0pmol/L, CV &lt;4.5%</td>
<td>9-20</td>
</tr>
<tr>
<td>Health ABC Study</td>
<td>Immunoassay (ACS, ChironDiagnosticsCorp, Emeryville, Calif), CV 4.1% at 18.94mIU/L and 3.6% at 1.26mIU/L</td>
<td>0.1-4.5</td>
<td>Competitive immunoassay (ACS; Chiron Diagnostics Corp)</td>
<td>10.3-23.2</td>
</tr>
<tr>
<td>HUNT Study</td>
<td>DELFIA hTSH Ultra (Wallac Oy, Turku, Finland), FS 0.03 mIU/L, VC 5%</td>
<td>0.5-3.5</td>
<td>DELFIA FT4, CV &lt;7%*</td>
<td>8-20</td>
</tr>
<tr>
<td>InCHIANTI Study</td>
<td>Chemiluminescent Immunoassay (Vitros Reagent, Ortho-Clinical Diagnostics, Johnson &amp; Johnson Medical Section, Milan, Italy), FS 0.003mIU/L, VC &lt;5.4%</td>
<td>0.46-4.68</td>
<td>Chemiluminescent Immunoassay (Vitros Reagent, Ortho-Clinical Diagnostics, Johnson &amp; Johnson Medical Section, Milan, Italy), FS 0.39pmol/L, CV &lt;5.3%</td>
<td>9.9-28.2</td>
</tr>
<tr>
<td>Leiden 85-Plus Study</td>
<td>Elecsys 2010 system (Hitachi, Tokyo, Japan) with an electrochemiluminescence technique (Boehringer, Mannheim, Germany), VC 5-11%</td>
<td>0.3-4.8</td>
<td>Elecsys 2010 system (Hitachi, Tokyo, Japan) with an electrochemiluminescence technique (Boehringer, Mannheim, Germany), VC 5-8%</td>
<td>13-23</td>
</tr>
<tr>
<td>MrOS</td>
<td>ADVIA Centaur (Siemens Diagnostics, Deerfield, IL, USA)</td>
<td>0.55-4.78</td>
<td>Competitive immunoassay (Siemens Diagnostics), CV 4.1%</td>
<td>9.0-23.8</td>
</tr>
<tr>
<td>OPUS</td>
<td>Single automated analyzer using the ARCHITECT System (Abbott ARCHITECT i2000; Abbott Laboratories, Maidenhead, UK), CV&lt;10.4%</td>
<td>0.13-3.48</td>
<td>Single automated analyzer using the ARCHITECT System (Abbott ARCHITECT i2000; Abbott Laboratories, Maidenhead, UK), CV&lt;10.4%</td>
<td>9.15-16.99</td>
</tr>
<tr>
<td>PROSPER</td>
<td>Three different immunoassays, FS 0.05 mIU/L, CV &lt;5%</td>
<td>0.45-4.5</td>
<td>Three different immunoassays, FS 0.05 mIU/L, CV &lt;5%</td>
<td>12-18**</td>
</tr>
<tr>
<td>Rotterdam Study</td>
<td>TSH Lumitest (Henning currently Brahms, Berlin, Germany)</td>
<td>0.4-4.0</td>
<td>Chemiluminescence assay (Vitros, ECI Immunodiagnostic System, Ortho-Clinical Diagnostics, Amersham, UK)</td>
<td>11-25</td>
</tr>
<tr>
<td>Sheffield Study</td>
<td>High-sensitivity immunoradiometric assay with reagents supplied by NETRIA (St. Bartholomew’s Hospital, London, UK), CV 8.0%</td>
<td>0.4-4.0</td>
<td>One step labelled antibody assay (Amerlex MAB, Lifescreen Ltd, Watford, UK), CV 4.8%</td>
<td>11-25</td>
</tr>
<tr>
<td>SOF</td>
<td>Chemiluminescent assay (Endocrine Science, Calabasas, California), FS 0.05mIU/L, CV&lt;6.3%</td>
<td>0.5-5.5</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: CHS, Cardiovascular Health Study; CV, coefficient of variation; EPIC, European Prospective Investigation of Cancer; FS, functional sensitivity; FT4, free thyroxine; Health ABC, Health, Aging and Body Composition; HUNT, Nord-Trøndelag Health; InCHIANTI, Invecchiare in Chianti; MrOS, Osteoporotic Fractures in Men Study; NA, not appropriate; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone.

Functional sensitivity was defined as the concentration at which the interassay CV is ≥20%.

* In the HUNT Study, FT4 was measured only if TSH was >4.0mIU/L or if the participant had known thyroid disease.

** Narrowed range among 3 different assays used for measurement.
## Supplemental Table 2. Definition of fractures in each study

<table>
<thead>
<tr>
<th>Study</th>
<th>Hip fracture</th>
<th>Any fracture</th>
<th>Non-vertebral fracture</th>
<th>Vertebral fracture</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHS</td>
<td>ICD9: 820.0-820.9 for inpatients, plus CPT procedure code on fracture treatment for outpatients</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>EPIC-Norfolk Study</td>
<td>ICD10: S72.0-2</td>
<td>Non-vertebral or vertebral fracture (first event)</td>
<td>Excluding skull, face, ankle, fingers, toes</td>
<td>Clinically diagnosed; thoracic (ICD10: S22), lumbar vertebrae (ICD10: S32), vertebrae of unknown location (ICD10: T08)</td>
</tr>
<tr>
<td>Health ABC Study</td>
<td>Femoral neck, intertrochanteric, proximal femur</td>
<td>Non-vertebral or vertebral fracture (first event)</td>
<td>Excluding ankle, fingers, toes</td>
<td>Clinically diagnosed; thoracic or lumbar vertebrae</td>
</tr>
<tr>
<td>HUNT Study</td>
<td>ICD9: 820.0-820.9, SIFF-95 procedure codes; ICD10: S72.0-2, S72.9, NCSP codes</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>InCHIANTI Study</td>
<td>ICD9: 820.0-820.9 for inpatients, plus CPT procedure code on fracture treatment for outpatients</td>
<td>Non-vertebral or vertebral fracture (first event)</td>
<td>Excluding skull, face, ankle, fingers, toes</td>
<td>Clinically diagnosed; thoracic (ICD9: 805.2-5) or lumbar vertebrae (ICD9: 806.2-5)</td>
</tr>
<tr>
<td>Leiden 85-Plus Study</td>
<td>Any hip fracture</td>
<td>Any fracture</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>MrOS</td>
<td>Femoral neck, intertrochanteric, subtrochanteric</td>
<td>Non-vertebral or vertebral fracture (first event)</td>
<td>Excluding skull, face, ankle, fingers, toes</td>
<td>Clinically diagnosed; thoracic or lumbar vertebrae</td>
</tr>
<tr>
<td>OPUS</td>
<td>Any low-traumatic hip fracture</td>
<td>NA</td>
<td>Any low-traumatic non-vertebral fracture</td>
<td>NA</td>
</tr>
<tr>
<td>PROSPER</td>
<td>NA</td>
<td>Any fracture</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Rotterdam Study</td>
<td>Any hip fracture</td>
<td>Non-vertebral or vertebral fracture (first event)</td>
<td>Excluding skull, ankle/foot, fingers/and/wrist</td>
<td>Any clinically diagnosed vertebral fracture</td>
</tr>
<tr>
<td>Sheffield Study</td>
<td>Any hip fracture</td>
<td>NA</td>
<td>Any non-vertebral fracture</td>
<td>NA</td>
</tr>
<tr>
<td>SOF</td>
<td>Any hip fracture, excluding severe traumatic fracture</td>
<td>Any fracture</td>
<td>Any non-vertebral fracture, excluding severe traumatic fracture</td>
<td>Any clinically diagnosed vertebral fracture, excluding severe traumatic fracture</td>
</tr>
</tbody>
</table>

Abbreviations: CHS, Cardiovascular Health Study; CPT, Current Procedural Terminology; EPIC, European Prospective Investigation of Cancer; GP, general practitioner; Health ABC, Health, Aging and Body Composition; HUNT, Nord-Trøndelag Health; ICD, international classification of disease; InCHIANTI, Invecchiare in Chianti; MrOS, Osteoporotic Fractures in Men Study; NA, not appropriate; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures.
## Supplemental Table 3. Study quality assessment

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Setting</th>
<th>Ascertainment of exposure</th>
<th>Covariates available for adjustment</th>
<th>Assessment of fractures</th>
<th>Adjudication blinded to thyroid function</th>
<th>Median (IQR) length of follow-up</th>
<th>Loss to follow-up</th>
<th>Fractures data published*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busselton Health Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid and anti-osteoporotic medication</td>
<td>ICD9 and ICD10 coded diagnoses in hospital discharge records</td>
<td>No</td>
<td>NA</td>
<td>5%</td>
<td>No</td>
</tr>
<tr>
<td>CHS</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Interview and hospital records reviewed by experts</td>
<td>No</td>
<td>NA</td>
<td>0%</td>
<td>Yes</td>
</tr>
<tr>
<td>EPIC-Norfolk Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid and thyroid-altering medication</td>
<td>Hospital discharge coding by data linkage with NHS central register</td>
<td>Yes</td>
<td>Yes</td>
<td>1.3%</td>
<td>Yes</td>
</tr>
<tr>
<td>Health ABC Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Interview, hospital records and other documents reviewed by clinicians</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;5%</td>
<td>Yes</td>
</tr>
<tr>
<td>HUNT Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid and thyroid-altering medication</td>
<td>Hospital and radiology records reviewed by physicians, health secretaries and nurses</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;5%</td>
<td>Yes</td>
</tr>
<tr>
<td>InCHIANTI Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Hospital records and other documents</td>
<td>Yes</td>
<td>Yes</td>
<td>&lt;5%</td>
<td>No</td>
</tr>
<tr>
<td>Leiden 85-Plus Study</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Annual interview of treating GP or nursing home physician and review of their medical records</td>
<td>No</td>
<td>NA</td>
<td>&lt;4%</td>
<td>Yes</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Setting</td>
<td>Ascertainment of exposure</td>
<td>Covariates available for adjustment</td>
<td>Assessment of fractures</td>
<td>Adjudication blinded to thyroid function</td>
<td>Median (IQR) length of follow-up</td>
<td>Loss to follow-up</td>
<td>Fractures data published*</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
<td>----------------------------------------------</td>
<td>---------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>MrOS</td>
<td>Random sample of a prospective cohort study</td>
<td>Sample of community-dwelling volunteers</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Interviewed reported fractures. Central adjudication by physician through radiology reports or X-rays</td>
<td>Yes</td>
<td>Yes</td>
<td>11.1 (8.1-11.8)</td>
<td>2%</td>
</tr>
<tr>
<td>OPUS</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Interview after 6 years follow-up, validated by medical records and imaging reviewed by radiologist</td>
<td>Yes</td>
<td>Yes</td>
<td>6.0 (5.8-6.3)</td>
<td>40%</td>
</tr>
<tr>
<td>PROSPER</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Fractures documented as adverse events</td>
<td>No</td>
<td>NA</td>
<td>3.3 (3.0-3.5)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid medication</td>
<td>GP and hospital registry records, reviewed by independent medical experts</td>
<td>Yes</td>
<td>Yes</td>
<td>15.2 (10.4-16.2)</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Sheffield</td>
<td>Prospective cohort study</td>
<td>Population-based study</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid medication</td>
<td>GP records and interviews, if confirmed by radiology or orthopedic report</td>
<td>Yes</td>
<td>Yes</td>
<td>10.0 (2.8-10.1)</td>
<td>2%</td>
</tr>
<tr>
<td>SOF</td>
<td>Prospective cohort study</td>
<td>Sample of community-dwelling volunteers</td>
<td>Third generation TSH assay</td>
<td>Age, sex, BMI, smoking status, diabetes; thyroid, thyroid-altering and anti-osteoporotic medication</td>
<td>Mail interview, with confirmation by X-rays or written report review by radiologist</td>
<td>Yes</td>
<td>Yes</td>
<td>14.3 (9.8-19.8)</td>
<td>5%</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; CHS, Cardiovascular Health Study; EPIC, European Prospective Investigation of Cancer; GP, general practitioner; Health ABC, Health, Aging and Body Composition; HUNT, Nord-Trøndelag Health; ICD, international classification of disease; InCHIANTI, Invecchiare in Chianti Study; IQR, interquartile range; MrOS, Osteoporotic Fractures in Men Study; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone.

* Four cohorts had not published their fractures data in a separate manuscript previously.
### Supplemental Table 4. Analysis stratified by publication status of fractures data

#### Analysis by thyroid-stimulating hormone categories

<table>
<thead>
<tr>
<th></th>
<th>Hip fracture</th>
<th>Any fracture</th>
<th>Non-vertebral fracture</th>
<th>Vertebral fracture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of events/ participants</td>
<td>Hazard ratio (95% CI)</td>
<td>No. of events/ participants</td>
<td>Hazard ratio (95% CI)</td>
</tr>
<tr>
<td><strong>Main analysis</strong></td>
<td>610/13,390</td>
<td>1.25 (1.05-1.49)</td>
<td>561/5,587</td>
<td>1.00 (0.83-1.19)</td>
</tr>
<tr>
<td><strong>Studies with published fractures studies</strong></td>
<td>572/12,460</td>
<td>1.35 (1.13-1.61)</td>
<td>387/4,097</td>
<td>1.16 (0.93-1.43)</td>
</tr>
<tr>
<td><strong>Studies with unpublished fractures data</strong></td>
<td>33/865</td>
<td>0.44 (0.21-0.90)</td>
<td>172/2,039</td>
<td>0.70 (0.50-0.96)</td>
</tr>
<tr>
<td><em>P-value for interaction</em></td>
<td>NA</td>
<td>0.0001</td>
<td>NA</td>
<td>0.17</td>
</tr>
</tbody>
</table>

#### Analysis by one standard deviation increase in free thyroxine¶

<table>
<thead>
<tr>
<th></th>
<th>Hip fracture**</th>
<th>Any fracture††</th>
<th>Non-vertebral fracture‡‡</th>
<th>Vertebral fracture§§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of events/ participants</td>
<td>Hazard ratio (95% CI)</td>
<td>No. of events/ participants</td>
<td>Hazard ratio (95% CI)</td>
</tr>
<tr>
<td><strong>Main analysis</strong></td>
<td>542/20,633</td>
<td>1.24 (1.12-1.37)</td>
<td>1,629/22,977</td>
<td>1.08 (1.02-1.15)</td>
</tr>
<tr>
<td><strong>Studies with published fractures studies</strong></td>
<td>453/17,663</td>
<td>1.25 (1.12-1.40)</td>
<td>1,003/15,192</td>
<td>1.09 (1.01-1.18)</td>
</tr>
<tr>
<td><strong>Studies with unpublished fractures data</strong>†</td>
<td>89/2,970</td>
<td>1.16 (0.90-1.89)</td>
<td>626/7,785</td>
<td>1.07 (0.97-1.17)</td>
</tr>
<tr>
<td><em>P-value for interaction</em></td>
<td>NA</td>
<td>0.59</td>
<td>NA</td>
<td>0.70</td>
</tr>
</tbody>
</table>

**Abbreviations:** CHS, Cardiovascular Health Study; CI, confidence interval; EPIC, European Prospective Investigation of Cancer; Health ABC, Health, Aging and Body Composition; InCHIANTI, Invecchiare in Chianti; MrOS, Osteoporotic Fractures in Men Study; OPUS, Osteoporosis and Ultrasound Study; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone; FT4, free thyroxine; No., number.

* All analyses were adjusted for age (as a continuous variable) and sex.
* Data on hip fractures were available for 12 studies (all but PROSPER).
† Data on any fractures were available for 9 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, PROSPER, Rotterdam Study, Busselton Health Study, SOF, Health ABC Study).
‡ Data on non-vertebral fractures were available for 9 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, OPUS, SOF, Health ABC Study).
§ Data on vertebral fracture were available for 7 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, OPUS, SOF, Health ABC Study). Vertebral fracture was defined as a clinical symptomatic dorsal or lumbar fracture.
¶ FT4 was measured in all studies but SOF and Health ABC Study (FT4 not measured in participants with TSH within normal range).
** Data on hip fracture were available for 10 studies with measured FT4 (all but PROSPER).
†† Data on any fracture were available for 7 studies with measured FT4 (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, PROSPER, Rotterdam Study, Busselton Health Study).
‡‡ Data on non-vertebral fracture were available for 7 studies with measured FT4 (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, OPUS).
§§ Data on vertebral fracture were available for 5 studies with measured FT4 (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study). Vertebral fracture was defined as a clinical symptomatic dorsal or lumbar fracture.
## Supplemental Table 5. Risk of any, non-vertebral, and vertebral fractures according to thyroid-stimulating hormone categories

<table>
<thead>
<tr>
<th>TSH level (mIU/L)</th>
<th>Any fracture*</th>
<th>Non-vertebral fracture†</th>
<th>Vertebral fracture‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of events/ participants</td>
<td>Hazard ratio (95% CI)</td>
<td>No. of events/ participants</td>
</tr>
<tr>
<td>3.50-4.49</td>
<td>179/1,769</td>
<td>1 (Reference)</td>
<td>140/1,396</td>
</tr>
<tr>
<td>2.50-3.49</td>
<td>376/4,096</td>
<td>0.98 (0.82-1.17)</td>
<td>280/3,301</td>
</tr>
<tr>
<td>1.50-2.49</td>
<td>841/9,847</td>
<td>1.01 (0.86-1.19)</td>
<td>650/8,033</td>
</tr>
<tr>
<td>1.00-1.49</td>
<td>555/6,105</td>
<td>1.07 (0.90-1.27)</td>
<td>440/5,165</td>
</tr>
<tr>
<td>0.45-0.99</td>
<td>382/4,379</td>
<td>1.00 (0.83-1.19)</td>
<td>364/4,121</td>
</tr>
<tr>
<td>P-value for trend</td>
<td>NA</td>
<td>0.56</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; Health ABC, Health, Aging and Body Composition; No., number; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SOF, Study of Osteoporotic Fractures; TSH, thyroid-stimulating hormone.

All analyses were adjusted for age (as a continuous variable) and for sex.

* Data on any fractures were available for 9 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, PROSPER, Rotterdam Study, Busselton Health Study, SOF, Health ABC Study).

† Data on non-vertebral fractures were available for 9 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, OPUS, SOF, Health ABC Study).

‡ Data on vertebral fractures were available for 7 studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, SOF, Health ABC Study). Vertebral fracture was defined as a clinical symptomatic dorsal or lumbar fracture.
Supplemental Figure 1. Flow diagram of the studies assessed for inclusion

13 studies identified through our previous literature search

254 studies excluded based on title and abstract because not meeting inclusion criteria

2 studies excluded:
- 1 meta-analysis reporting studies already included or not meeting inclusion criteria
- 1 study not prospective

8 studies already identified through our previous literature search

1 study excluded because using 1st generation TSH assay

265 studies identified through new database search and screened

11 full-text articles assessed for eligibility

9 reports meeting inclusion criteria

1 study meeting inclusion criteria and not yet included

13 studies included in the individual participants data (IPD) analysis

Abbreviations: IPD, individual participants data; TSH, thyroid-stimulating hormone.
Supplemental Figure 2. Risk of hip fracture in participants with thyroid-stimulating hormone 0.45-1.49mIU/L, compared to the reference group with thyroid-stimulating hormone 3.50-4.49mIU/L, stratified by sex, age, and duration of follow-up

<table>
<thead>
<tr>
<th></th>
<th>No. with fracture</th>
<th>No. of participants</th>
<th>Hazard ratio (95% CI)</th>
<th>P-value for interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>429</td>
<td>10,142</td>
<td>1.25 (1.05-1.49)</td>
<td>-</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>483</td>
<td>8,629</td>
<td>1.21 (0.99-1.47)</td>
<td>0.42</td>
</tr>
<tr>
<td>Men</td>
<td>127</td>
<td>4,761</td>
<td>1.51 (1.00-2.26)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;75 years</td>
<td>354</td>
<td>11,246</td>
<td>1.16 (0.92-1.46)</td>
<td>0.07</td>
</tr>
<tr>
<td>≥75 years</td>
<td>256</td>
<td>2,144</td>
<td>1.36 (1.04-1.77)</td>
<td></td>
</tr>
<tr>
<td>Duration of follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 years</td>
<td>41</td>
<td>1,029</td>
<td>1.03 (0.52-2.02)</td>
<td>0.58</td>
</tr>
<tr>
<td>≥5 years</td>
<td>569</td>
<td>12,361</td>
<td>1.26 (1.05-1.51)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; HR, hazard ratio; No, number; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; TSH, thyroid-stimulating hormone.

We present a selected analysis for the TSH categories 0.45-0.99mIU/L and 3.50-4.99mIU/L. Hazard ratios are for TSH 0.45-0.99mIU/L, compared with the reference group 3.50-4.99mIU/L. The analysis stratified by sex was adjusted for age. All other analyses were adjusted for age (as a continuous variable) and sex. Data on hip fractures were available for 12 studies (all except PROSPER).