# Thyroid Function Tests in the Reference Range and Fracture: Individual Participant Analysis of Prospective Cohorts

Carole E. Aubert,[1] Carmen Floriani,[1] Douglas C. Bauer,[2] Bruno R. da Costa,[3,4] Daniel Segna,[1] Manuel R. Blum,[1] Tinh-Hai Collet,[5] Howard A. Fink,[6,7] Anne R. Cappola,[8] Lamprini Syrogiannouli,[3] Robin P. Peeters,[9] Bjørn O. Åsvold,[10,11] Wendy P. J. den Elzen,[12] Robert N. Luben,[13] Alexandra P. Bremner,[14] Apostolos Gogakos,[15] Richard Eastell,[16] Patricia M. Kearney,[17] Mari Hoff,[10,18] Erin Le Blanc,[19] Graziano Ceresini,[20] Fernando Rivadeneira,[9] André G. Uitterlinden,[9] Kay-Tee Khaw,[14] Arnulf Langhammer,[10] David J. Stott,[21] Rudi G. J. Westendorp,[22] Luigi Ferrucci,[23] Graham R. Williams,[16] Jacobijn Gussekloo,[12,24] John P. Walsh,[25,26] Drahomir Aujesky,[1] and Nicolas Rodondi,[1,3] on behalf of the Thyroid Studies Collaboration

[1]Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland; [2]Departments of Medicine and Epidemiology and Biostatistics, University of California, San Francisco, California 94143; [3]Institute of Primary Health Care (Berner Institut für Hausarztmedizin), University of Bern, 3012 Bern, Switzerland; [4]Department of Cardiology, Swiss Cardiovascular Center Bern, Bern University Hospital, 3010 Bern, Switzerland; [5]Service of Endocrinology, Diabetes, and Metabolism, University Hospital of Lausanne, 1011 Lausanne, Switzerland; [6]Geriatric Research Education and Clinical Center, Veterans Affairs Health Care System, Minneapolis, Minnesota 55417; [7]Department of Medicine, University of Minnesota, Minneapolis, Minnesota 55455; [8]University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104; [9]Departments of Internal Medicine and Epidemiology, Erasmus University Rotterdam, 3062 Rotterdam, The Netherlands; [10]Department of Public Health and Nursing, Norges teknisk-naturvitenskapelige universitet, Norwegian University of Science and Technology, N-7491 Trondheim, Norway; [11]Department of Endocrinology, St. Olav's Hospital, Trondheim University Hospital, N-7006 Trondheim, Norway; [12]Leiden University Medical Center, Department of Clinical Chemistry and Laboratory Medicine, 2300 RC Leiden, The Netherlands; [13]Department of Public Health and Primary Care, University of Cambridge, Cambridge CB2 0SR, United Kingdom; [14]School of Population Health, University of Western Australia, Crawley, Perth, Western Australia 6009, Australia; [15]Molecular Endocrinology Laboratory, Hammersmith Campus, Department of Medicine, Imperial College London, London W12 0NN, United Kingdom; [16]Department of Oncology and Metabolism, University of Sheffield, Sheffield, South Yorkshire S10 2TN, United Kingdom; [17]Department of Epidemiology and Public Health, University College Cork, T12 PX46 Cork, Ireland; [18]Levanger Hospital, Nord-Trøndelag Hospital Trust, 7600 Levanger, Norway; [19]Center for Health Research NW, Kaiser Permanente, Portland, Oregon 97277; [20]Department of Clinical and Experimental Medicine, Geriatric Endocrine Unit, University Hospital of Parma, 43126 Parma, Italy; [21]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow G31 2ER, United Kingdom; [22]Department of Public Health, Center for Health Aging, University of Copenhagen, 1004 Copenhagen, Denmark; [23]National Institute on Aging, National Institutes of Health, Baltimore, Maryland 21224; [24]Department of Gerontology and Geriatrics, Leiden University Medical Center, 2300 RC, Leiden, The Netherlands; [25]Department of Endocrinology and Diabetes, Sir Charles Gairdner Hospital, Nedlands, Western Australia 6009, Australia; and [26]School of Medicine and Pharmacology, University of Western Australia, Crawley, Perth, Western Australia 6009, Australia

**Context:** Hyperthyroidism is associated with increased fracture risk, but it is not clear if lower thyroid-stimulating hormone (TSH) and higher free thyroxine (FT4) in euthyroid individuals are associated with fracture risk.

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 31 January 2017. Accepted 26 April 2017.
First Published Online 5 May 2017

Abbreviations: BMD, bone mineral density; BMI, body mass index; CI, confidence interval; FT4, free thyroxine; HR, hazard ratio; IPD, individual participant data; IQR, interquartile range; SD, standard deviation; TSH, thyroid-stimulating hormone.

**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 to 4.49 mIU/L.

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

**Results:** During 659,059 person-years, 2,565 out of 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 to 1.49) for TSH 0.45 to 0.99 mIU/L, 1.19 (1.01 to 1.41) for TSH 1.00 to 1.49 mIU/L, 1.09 (0.93 to 1.28) for TSH 1.50 to 2.49 mIU/L, and 1.12 (0.94 to 1.33) for TSH 2.50 to 3.49 mIU/L (*P* for trend = 0.004). Hip fracture was also associated with FT4 [HR (95% CI) 1.22 (1.11 to 1.35) per one standard deviation increase in FT4]. FT4 only was associated with any and nonvertebral fractures. 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. **(*J Clin Endocrinol Metab* 102: 2719–2728, 2017)**

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 thyroid-stimulating hormone (TSH) reduces bone resorption and increases bone formation in postmenopausal 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 the TSH reference range was derived did not exclude participants with occult or underlying disease (*e.g.*, those with positive antithyroid antibodies), which might bias the reference range toward 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 postmenopausal 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 free thyroxine (FT4) in euthyroid participants were associated with increased risk of fractures. We therefore aimed to assess the association between TSH within the reference range, 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 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 May 19, 2016) used the 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 antithyroid drugs); (4) studies with only participants with TSH outside the reference range (<0.45 mIU/L or >4.49 mIU/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 later, 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 to 4.49 mIU/L. For clinical relevance, we separated TSH values into five categories: 0.45 to 0.99 mIU/L; 1.00 to 1.49 mIU/L; 1.50 to 2.49 mIU/L; 2.50 to 3.49 mIU/L; and 3.50 to 4.49 mIU/L. The latter was used as a 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 cutoffs. 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, nonvertebral, and clinical vertebral incident fractures were secondary outcomes. We excluded (1) vertebral fractures diagnosed with only radiologic imaging to keep focus on clinical relevance and (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, because these are not related to osteoporosis. "Nonvertebral fractures" were the same as "any fractures" except they excluded vertebral fractures. For any and nonvertebral 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 (CIs). We used Schoenfeld residuals to test the proportional hazards assumption (23). Results were presented as hazard ratios (HRs) compared with 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 the following predefined sensitivity analyses: (1) excluding participants with thyroid medication (thyroxine or antithyroid medication) at baseline; (2) excluding participants with thyroid-altering medication at baseline (thyroid medication, oral corticosteroids, amiodarone, and iodine); (3) excluding participants with antifracture medication at baseline (bisphosphonate, calcitonin, selective estrogen receptor modulator, and parathyroid hormone); (4) including only studies with formal fracture adjudication; (5) including only studies that uniformly defined fractures (except for hip fracture, because it has a common definition and is rarely reported in error); (6) excluding cohorts with loss to follow-up rates >5%; (7) including only participants with TSH remaining within the reference range during follow-up; and (8) 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 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 vs ≥75 years), and duration of follow-up (<5 vs ≥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, TX). All tests were two-sided, at a 0.05 level of significance.

## Results

Our updated literature search identified nine additional reports (Supplemental Fig. 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 (1 mIU/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. We included 13 studies (5, 10, 14, 17, 19, 25, 26, 29–34) from the United States, Europe, and Australia with 61,959 participants, and a median duration of follow-up of 12.1 years [interquartile range (IQR), 8.5 to 12.9], totaling 659,059 person-years. Median age was 64 (range, 18 to 102) with 60.5% women (Table 1). Median (IQR) TSH was 1.60 mIU/L (1.10 to 2.30); 3.1% of participants used thyroid medication at baseline and 5.5% during follow-up; 17.7% had a TSH 0.45 to 0.99 mIU/L, 24.8% 1.00 to 1.49 mIU/L, 37.4% 1.50 to 2.49 mIU/L, 14.2% 2.50 to 3.49 mIU/L, and 5.9% 3.50 to 4.49 mIU/L. Hip fracture occurred in 2565 participants (4.5%; 12 studies), any fracture in 2333 (8.9%; 9 studies), nonvertebral fracture in 1874 (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).

**Table 1. Study Population and Baseline Characteristics of the Participants in the 13 Included Studies (N = 61,959)**

| Study Name, Place | Description of Study Population | Number of Participants | Age, Median (Range)[a] | Women, No. (%) | TSH, Median, mIU/L | Thyroid Medication at Baseline, No. (%)[b,c] | Thyroid Medication During Follow-Up, No. (%)[b,d] | Start of Follow-Up, Year | Duration of Follow-Up, Median (IQR), Years[e] | Person-Years |
|---|---|---|---|---|---|---|---|---|---|---|
| Busselton Health Study, Australia (29) | Adults | 1907 | 51 (18–90) | 919 (48.2) | 1.42 | 10 (0.5) | 15 (0.8) | 1981 | 20.0 (17.6–20.0) | 33,281 |
| CHS, United States (four communities) (25) | Adults with Medicare eligibility | 2853 | 71 (65–100) | 1694 (59.4) | 2.03 | 145 (5.1) | 299 (10.5) | 1989–1990 | 12.9 (7.5–18.9) | 36,466 |
| EPIC-Norfolk Study, England (30) | Adults aged 45 to 79 years | 11,986 | 58 (40–78) | 6365 (53.1) | 1.70 | 275 (2.3) | NA | 1995–1998 | 12.4 (11.7–13.3) | 142,951 |
| Health ABC Study, United States (four communities) (14) | Adults aged 70 to 79 years with Medicare eligibility | 2347 | 74 (69–81) | 1165 (49.6) | 1.99 | 177 (7.5) | 383 (13.9) | 1997 | 12.7 (8.1–13.2) | 24,794 |
| HUNT Study, Norway (31)[f] | Adults | 31,388 | 57 (19–99) | 21,186 (67.5) | 1.60 | 999 (3.2) | NA | 1995–1997 | 12.2 (11.6–12.8) | 345,517 |
| InCHIANTI Study, Italy (17) | Adults aged ≥65 years | 1066 | 71 (21–102) | 590 (55.3) | 1.38 | 17 (1.6) | 28 (2.6) | 1998 | 9.1 (7.8–9.3) | 8562 |
| Leiden 85-Plus Study, The Netherlands (32) | Adults aged 85 years | 456 | 85 (85–85) | 293 (64.3) | 1.66 | 6 (1.3) | 11 (2.4) | 1997–1999 | 4.8 (2.2–8.1) | 2411 |
| MrOS, United States (six clinical centers) (10) | Men aged ≥65 years | 1410 | 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, and United Kingdom (5)[g] | Women aged 20 to 80 years | 1205 | 63 (20–80) | All women | 0.96 | 0 (0.0) | NA | 1999–2001 | 6.0 (5.8–6.3) | 7179 |
| PROSPER, Ireland, Scotland, and The Netherlands (33) | Older adults at high cardiovascular risk | 5124 | 75 (69–83) | 2527 (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 ≥55 years | 1611 | 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 to 85 years | 291 | 63 (50–86) | All women | 2.00 | 2 (0.7) | 9 (3.1) | 1990–1991 | 10.0 (5.5–10.1) | 2301 |
| SOF, United States (four clinical centers) (19)[h] | Women >65 years | 314 | 71 (65–88) | All women | 1.50 | 15 (4.8) | NA | 1986–1998 | 14.3 (9.8–19.8) | 4433 |
| Overall | 13 cohorts | 61,959 | 64 (18–102) | 37,506 (60.5) | 1.60 | 1885 (3.1) | 831 (5.5) | 1981–2002 | 12.1 (8.5–12.9) | 659,059[i] |

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; 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.

[a]We excluded participants younger than 18 years.

[b]Thyroid medication was defined as thyroxine or antithyroid medication.

[c]Data on thyroid medication at baseline was missing for 255 participants in the HUNT Study, 59 participants in the MrOS, 1 participant in the Rotterdam Study, 4 participants in the SOF, and 7 participants in the Health ABC Study.

[d]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.

[e]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].

[f]We included participants excluded from the original article of the HUNT Study (participants <40 years, with previous fracture and/or with previous thyroid disease), which explains the different number of the sample.

[g]We included only the thyroid hormone substudy of the OPUS, which excluded participants on thyroid medication.

[h]We included only a subsample of the SOF (*i.e.*, the participants with TSH measurement at baseline).

[i]It was calculated as time to hip fracture; for the PROSPER, it was calculated as time to any fracture, because data on hip fracture was unavailable.

### Thyroid function and hip fractures

Compared with the reference group (TSH 3.50 to 4.49 mIU/L), pooled age- and sex-adjusted HR (95% CI) for hip fracture was 1.25 (1.05 to 1.49) for TSH 0.45 to 0.99 mIU/L, 1.19 (1.01 to 1.41) for TSH 1.00 to 1.49 mIU/L, 1.09 (0.93 to 1.28) for TSH 1.50 to 2.49 mIU/L, and 1.12 (0.94 to 1.33) for TSH 2.50 to 3.49 mIU/L (*P* for trend 0.004; Fig. 1). After adjusting for BMI, smoking status, and history of diabetes, HR (95% CI) was 1.24 (1.03 to 1.49) for TSH 0.45 to 0.99 mIU/L compared with the reference group, whereas HR (95% CI) for TSH 1.00 to 1.49 mIU/L was somewhat attenuated and no longer statistically significant [1.15 (0.97 to 1.38)]. The risk of hip fracture in participants with TSH 0.45 to 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 to 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 to 2.49 mIU/L, or TSH 2.50 to 3.49 mIU/L. We found no significant interaction with sex, age, or duration of follow-up (Supplemental Fig. 2), although CIs were larger and point estimates smaller for age <75 years and follow-up <5 years. Conversely, there was significant interaction for publication status with an HR (95% CI) of 1.35 (1.13 to 1.61) for the 10 studies that published risk of hip fracture associated with thyroid function tests in a separate manuscript and 0.44 (0.21 to 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 to 1.37) per one SD increase in FT4 (Fig. 2). We found no significant interaction with sex, age, duration of follow-up, or publication status of hip fracture data (Fig. 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 to 0.99 mIU/L compared with the reference group was 1.70 (1.13 to 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, nonvertebral, and vertebral fractures

For all TSH categories when compared with the reference group, we found no significant association for any, nonvertebral, or vertebral fractures (Supplemental Table 5). The HR (95% CI) per one SD increase in FT4 was 1.08 (1.02 to 1.15) for any fracture and 1.10 (1.03 to 1.18) for nonvertebral 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 13 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 nonvertebral fracture.

Although overt and subclinical hyperthyroidism has 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 to 1.6 mIU/L when compared with TSH

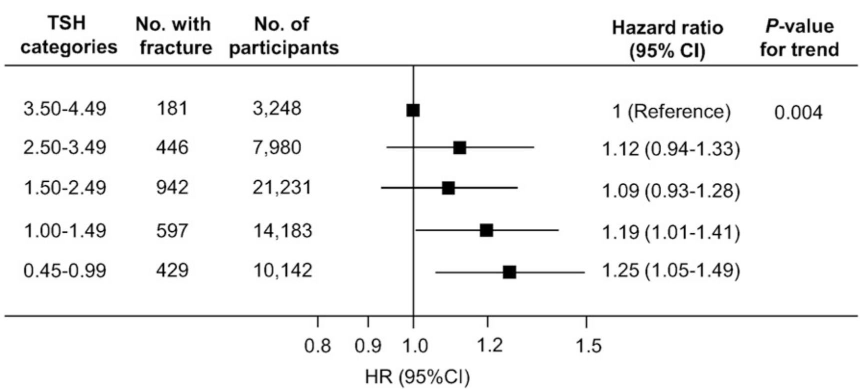


**Figure 1.** Risk of hip fracture according to TSH categories. Data on hip fractures were available for 12 studies (all except PROSPER).

**Table 2. Sensitivity Analyses for the Risk of Hip Fracture According to TSH and FT4**

| | Analysis by TSH Category[a] | | Analysis by SD Increase in FT4[b] | |
|---|---|---|---|---|
| | No. of Events/ Participants | HR (95% CI)[c] | No. of Events/ Participants | HR (95% CI)[d] |
| Main analysis | 610/13,390 | 1.25 (1.05–1.49) | 542/20,633 | 1.24 (1.12–1.37) |
| Medication use | | | | |
| Excluding participants with thyroid medication at baseline[e] | 557/12,728 | 1.28 (1.06–1.53) | 526/20,158 | 1.26 (1.13–1.40) |
| Excluding participants with thyroid-altering medication at baseline[f] | 542/12,396 | 1.28 (1.07–1.55) | 506/19,679 | 1.26 (1.13–1.40) |
| Excluding participants with antifracture medication at baseline[g] | 605/12,739 | 1.27 (1.07–1.52) | 539/20,563 | 1.24 (1.12–1.38) |
| Definition of fracture | | | | |
| Including only studies with formal fracture adjudication[h] | 496/12,048 | 1.31 (1.06–1.60) | 416/17,913 | 1.21 (1.07–1.36) |
| Other | | | | |
| Excluding one study with loss to follow-up $> 5\%$[i] | 606/12,748 | 1.26 (1.05–1.50) | 536/19,463 | 1.24 (1.11–1.37) |
| BMD | | | | |
| Further adjusting for femoral neck BMD at baseline[j] | 94/2020 | 1.68 (1.08–2.61)[k] | 142/4147 | 1.22 (1.01–1.47) |

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

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; 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.

[a]We present a selected analysis for the TSH category 0.45 to 0.99 mIU/L compared with the reference category (TSH 3.50 to 4.99 mIU/L). Numbers are for participants in both of these TSH categories only.

[b]FT4 was measured in all studies but the SOF and the Health ABC Study (FT4 not measured in participants with TSH within reference range).

[c]HRs are for TSH 0.45 to 0.99 mIU/L, compared with the reference group TSH 3.50 to 4.99 mIU/L.

[d]HRs are per one SD increase in FT4.

[e]Thyroid medication was defined as thyroxine or antithyroid medication.

[f]Thyroid-altering medication included oral corticosteroid, amiodarone, iodine, thyroxine, or antithyroid medication.

[g]Antifracture medication was defined as bisphosphonate, calcitonin, selective estrogen receptor modulator, or parathyroid hormone.

[h]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).

[i]OPUS.

[j]Femoral neck BMD at baseline was available in the following studies: CHS, MrOS, Rotterdam Study, Sheffield Study, OPUS, and Health ABC Study.

[k]Participants within the TSH category 3.50 to 4.49 mIU/L had lower femoral neck BMD at baseline than participants within the TSH category 0.45 to 1.50 mIU/L [mean (SD): 0.77 g/cm$^2$ (0.16) vs 0.79 g/cm$^2$ (0.15), respectively, $P = 0.002$], which explains the higher HR after adjusting for femoral neck BMD at baseline.

1.7 to 2.9 mIU/L, but in women only [odds ratio (95% CI), 1.28 (1.03 to 1.59)], whereas 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 <1.5 mIU/L and for FT4 >1.4 ng/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

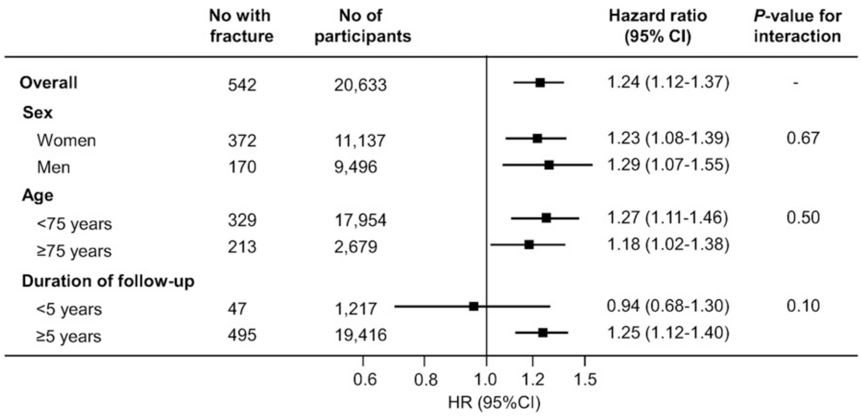


**Figure 2.** Risk of hip fracture per one SD increase in FT4, overall and stratified by sex, age, and duration of follow-up. The analysis stratified for sex was adjusted for age. All other analyses were adjusted for age (as a continuous variable) and sex. FT4 was measured in all studies but the Study of Osteoporotic Fractures (SOF) and the Health, Aging, and Body Composition (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).

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, because 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 nonthyroidal illness, a potential cause of suppressed TSH; however, the prevalence of nonthyroidal illness was likely low, as we only studied community-dwelling adults. Third, fractures were adjudicated in 9 of 13 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 mechanisms, 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, the TSH reference range is still a matter of debate (35). The 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 2.5 mIU/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 nonvertebral fractures, whereas TSH was associated only with hip fractures. TSH and thyroid hormones may act differently on peripheral organs, including bones: TSH may act on osteoblasts and osteoclasts via specific receptors (3), whereas 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 10 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 13 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 nonvertebral fractures. Our findings may help refine the current definition of optimal thyroid function. Meanwhile, clinicians should be aware that lower TSH and

**Table 3. Sensitivity and Stratified Analyses for the Risk of Any, Nonvertebral, and Vertebral Fractures Per One SD Increase in FT4**

| | Any Fracture[a] | | Nonvertebral Fracture[b] | | Vertebral Fracture[c] | |
|---|---|---|---|---|---|---|
| | No. of Events/ Participants | HR (95% CI) | No. of Events/ Participants | HR (95% CI) | No. of Events/ Participants | HR (95% CI) |
| Main analysis | 1629/22,977 | 1.08 (1.02–1.15) | 1273/19,101 | 1.10 (1.03–1.18) | 129/17,711 | 1.06 (0.86–1.30) |
| Sensitivity analyses | | | | | | |
| Medication use | | | | | | |
| Excluding participants with thyroid medication at baseline[d] | 1552/22,440 | 1.09 (1.02–1.16) | 1240/18,697 | 1.14 (1.06–1.23) | 125/17,309 | 1.08 (0.86–1.37) |
| Excluding participants with thyroid-altering medication at baseline[e] | 1537/21,976 | 1.09 (1.02–1.15) | 1200/18,256 | 1.11 (1.03–1.19) | 125/16,868 | 1.07 (0.86–1.32) |
| Excluding participants with antifracture medication at baseline[f] | 1622/22,927 | 1.08 (1.02–1.15) | 1263/19,038 | 1.10 (1.03–1.18) | 127/17,666 | 1.05 (0.85–1.29) |
| Definition of fracture | | | | | | |
| Including only studies with formal fracture adjudication[g] | 1026/15,805 | 1.11 (1.02–1.19) | 1111/17,208 | 1.11 (1.03–1.19) | 113/15,806 | 1.07 (0.86–1.32) |
| Including only studies with most uniform definition of fracture[h] | 1155/19,728 | 1.06 (0.99–1.14) | 685/14,461 | 1.08 (0.98–1.19) | 65/14,462 | 1.10 (0.83–1.47) |
| Other | | | | | | |
| Further adjusting for BMI, smoking status, and diabetes mellitus | 1591/22,536 | 1.21 (1.00–1.46) | 1140/17,562 | 1.09 (1.01–1.18) | 126/17,290 | 1.03 (0.83–1.27) |
| Excluding studies with loss of follow-up rate >5% | NA | NA | 1174/17,981 | 1.13 (1.04–1.22) | NA | NA |
| BMD | | | | | | |
| Further adjusting for lumbar spine BMD at baseline[i] | NA | NA | NA | NA | 39/1399 | 0.96 (0.68–1.36) |
| Further adjusting for whole-body BMD at baseline[j] | 183/1399 | 0.89 (0.75–1.04) | NA | NA | NA | NA |
| Stratified analyses | | | | | | |
| Stratified for sex | | | | | | |
| Women | 1013/11,321 | 1.11 (1.03–1.19) | 827/10,075 | 1.10 (1.01–1.20) | 62/8679 | 1.12 (0.83–1.51) |
| Men | 616/11,656 | 1.05 (0.95–1.15) | 446/9026 | 1.08 (0.96–1.22) | 67/9032 | 1.00 (0.75–1.33) |
| *P* value for interaction | NA | 0.39 | NA | 0.79 | NA | 0.61 |
| Stratified for age | | | | | | |
| <75 years at baseline | 1041/18,367 | 1.10 (1.02–1.19) | 955/17,144 | 1.10 (1.02–1.20) | 87/15,917 | 0.96 (0.74–1.25) |
| ≥75 years at baseline | 588/4610 | 1.06 (0.96–1.16) | 318/1957 | 1.10 (0.97–1.25) | 42/1794 | 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 | 446/5920 | 1.04 (0.93–1.15) | 47/888 | 0.82 (0.59–1.14) | 7/888 | 0.60 (0.26–1.37) |
| ≥5 years | 1183/17,057 | 1.09 (1.02–1.18) | 1226/18,213 | 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 |

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). HRs are per one SD increase in FT4.

Abbreviations: EPIC, European Prospective Investigation of Cancer; Health ABC, Health, Aging, and Body Composition; InCHIANTI, Invecchiare in Chianti Study; 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.

[a]Data on any fractures were available for seven studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, PROSPER, Rotterdam Study, and Busselton Health Study).

[b]Data on nonvertebral fractures were available for seven studies (MrOS, EPIC-Norfolk Study, InCHIANTI Study, Rotterdam Study, Busselton Health Study, Sheffield Study, and OPUS).

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

[d]Thyroid medication was defined as thyroxine or antithyroid medication.

[e]Thyroid-altering medication included oral corticosteroid, amiodarone, iodine, thyroxine, or antithyroid medication.

[f]Antifracture medication was defined as bisphosphonate, calcitonin, selective estrogen receptor modulator, or parathyroid hormone.

[g]EPIC-Norfolk Study, InCHIANTI Study, MrOS, OPUS, Rotterdam Study, and Sheffield Study.

[h]EPIC-Norfolk Study, InCHIANTI Study, Leiden 85-Plus Study, MrOS, and PROSPER.

[i]Lumbar spine BMD was available in MrOS only.

[j]Whole-body BMD was available in MrOS only.

higher FT4, even within the reference range, are associated with an increased risk of hip fracture.

## Acknowledgments

Current Affiliation: G.R. Williams is currently affiliated with Molecular Endocrinology Laboratory, Hammersmith Campus, Department of Medicine, Imperial College London, London W12 0NN, United Kingdom. J. Gussekloo is currently affiliated with Department of Public Health and Primary Care, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

Address all correspondence and requests for reprints to: Nicolas Rodondi, MD, MAS, Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland. E-mail: nicolas.rodondi@insel.ch.

This work was supported by 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; to J.G., P.M.K., N.R., D.J.S., and R.G.J.W.) and research grants from the Swiss National Science Foundation (P3SMP3-155318 and PZ00P3-167826; to T.-H.C.), the Netherlands Organization for Scientific Research (NGI/NOW 911-03-016; to R.G.J.W.), and the Swiss National Science Foundation (SNSF 320030-150025; to N.R.).

Author contributions: C.E.A. and N.R. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. C.E.A., D.C.B., and N.R. conceived and designed the study. C.E.A. and D.S. conducted the literature search and review. B.O.Å., C.E.A., M.R.B., A.P.B., A.R.C., G.C., W.P.J.d.E., J.G., P.M.K., K.-T.K., R.P.P., D.J.S., J.P.W., R.G.J.W., and N.R. conducted data acquisition, analysis, and interpretation. C.E.A. drafted the manuscript. B.O.Å., D.A., D.C.B., M.R.B., A.P.B., A.R.C., G.C., T.-H.C., B.R.d.C., W.P.J.d.E., L.F., H.A.F., C.F., J.G., M.H., P.M.K., K.-T.K., A.L., E.L.B., R.N.L., R.P.P., F.R., N.R., D.S., D.J.S., L.S., A.G.U., J.P.W., R.G.J.W., and G.R.W. performed critical revision of the manuscript for important intellectual content. C.E.A., M.R.B., B.R.d.C., N.R., and L.S. conducted statistical analyses. D.C.B., G.C., W.P.J.d.E., L.F., J.G., K.-T.K., and R.P.P. provided administrative, technical, and material support. D.C.B. and N.R. supervised the study.

Clinical trial registry: PROSPERO registration no. CRD42016039125 (registered 15 May 2016).

Disclosure Summary: R.P.P. 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. R.E. reports grants and personal fees from Amgen; grants from AstraZeneca; grants and personal fees from Alexion; grants, personal fees, and non-financial support from Immunodiagnostic Systems; grants, personal fees, and non-financial support from Roche; personal fees from Lilly; personal fees from D3 Biomedical Science Institutes; and personal fees from GSK Nutrition, all outside of the submitted work.

## References

1. Vestergaard P, Mosekilde L. Hyperthyroidism, bone mineral, and fracture risk—a meta-analysis. *Thyroid*. 2003;**13**(6):585–593.
2. Blum MR, Bauer DC, Collet TH, Fink HA, Cappola AR, da Costa BR, Wirth CD, Peeters RP, Åsvold BO, den Elzen WP, Luben RN, Imaizumi M, Bremner AP, Gogakos A, Eastell R, Kearney PM, Strotmeyer ES, Wallace ER, Hoff M, Ceresini G, Rivadeneira F, Uitterlinden AG, Stott DJ, Westendorp RG, Khaw KT, Langhammer A, Ferrucci L, Gussekloo J, Williams GR, Walsh JP, Jüni P, Aujesky D, Rodondi N; Thyroid Studies Collaboration. Subclinical thyroid dysfunction and fracture risk: a meta-analysis. *JAMA*. 2015;**313**(20):2055–2065.
3. Bassett JH, Williams GR. Role of thyroid hormones in skeletal development and bone maintenance. *Endocr Rev*. 2016;**37**(2): 135–187.
4. Brennan MD, Powell C, Kaufman KR, Sun PC, Bahn RS, Nair KS. The impact of overt and subclinical hyperthyroidism on skeletal muscle. *Thyroid*. 2006;**16**(4):375–380.
5. Murphy E, Glüer CC, Reid DM, Felsenberg D, Roux C, Eastell R, Williams GR. Thyroid function within the upper normal range is associated with reduced bone mineral density and an increased risk of nonvertebral fractures in healthy euthyroid postmenopausal women. *J Clin Endocrinol Metab*. 2010;**95**(7):3173–3181.
6. Mazziotti G, Sorvillo F, Piscopo M, Cioffi M, Pilla P, Biondi B, Iorio S, Giustina A, Amato G, Carella C. Recombinant human TSH modulates in vivo C-telopeptides of type-1 collagen and bone alkaline phosphatase, but not osteoprotegerin production in postmenopausal women monitored for differentiated thyroid carcinoma. *J Bone Miner Res*. 2005;**20**(3):480–486.
7. Wartofsky L, Dickey RA. The evidence for a narrower thyrotropin reference range is compelling. *J Clin Endocrinol Metab*. 2005; **90**(9):5483–5488.
8. Surks MI, Goswami G, Daniels GH. The thyrotropin reference range should remain unchanged. *J Clin Endocrinol Metab*. 2005; **90**(9):5489–5496.
9. Taylor PN, Razvi S, Pearce SH, Dayan CM. Clinical review: a review of the clinical consequences of variation in thyroid function within the reference range. *J Clin Endocrinol Metab*. 2013;**98**(9): 3562–3571.
10. Waring AC, Harrison S, Fink HA, Samuels MH, Cawthon PM, Zmuda JM, Orwoll ES, Bauer DC; Osteoporotic Fractures in Men (MrOS) Study. A prospective study of thyroid function, bone loss, and fractures in older men: the MrOS study. *J Bone Miner Res*. 2013;**28**(3):472–479.
11. Leader A, Ayzenfeld RH, Lishner M, Cohen E, Segev D, Hermoni D. Thyrotropin levels within the lower normal range are associated with an increased risk of hip fractures in euthyroid women, but not men, over the age of 65 years. *J Clin Endocrinol Metab*. 2014;**99**(8): 2665–2673.
12. Mazziotti G, Porcelli T, Patelli I, Vescovi PP, Giustina A. Serum TSH values and risk of vertebral fractures in euthyroid postmenopausal women with low bone mineral density. *Bone*. 2010; **46**(3):747–751.
13. Cappola AR, Arnold AM, Wulczyn K, Carlson M, Robbins J, Psaty BM. Thyroid function in the euthyroid range and adverse outcomes in older adults. *J Clin Endocrinol Metab*. 2015;**100**(3):1088–1096.
14. Rodondi N, den Elzen WP, Bauer DC, Cappola AR, Razvi S, Walsh JP, Asvold BO, Iervasi G, Imaizumi M, Collet TH, Bremner A, Maisonneuve P, Sgarbi JA, Khaw KT, Vanderpump MP, Newman AB, Cornuz J, Franklyn JA, Westendorp RG, Vittinghoff E, Gussekloo J; Thyroid Studies Collaboration. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA*. 2010;**304**(12):1365–1374.
15. Gencer B, Collet TH, Virgini V, Bauer DC, Gussekloo J, Cappola AR, Nanchen D, den Elzen WP, Balmer P, Luben RN, Iacoviello M, Triggiani V, Cornuz J, Newman AB, Khaw KT, Jukema JW, Westendorp RG, Vittinghoff E, Aujesky D, Rodondi N; Thyroid Studies Collaboration. Subclinical thyroid dysfunction and the risk of heart failure events: an individual participant data analysis from 6 prospective cohorts. *Circulation*. 2012;**126**(9):1040–1049.
16. Collet TH, Gussekloo J, Bauer DC, den Elzen WP, Cappola AR, Balmer P, Iervasi G, Åsvold BO, Sgarbi JA, Völzke H, Gencer B, Maciel RM, Molinaro S, Bremner A, Luben RN, Maisonneuve P, Cornuz J, Newman AB, Khaw KT, Westendorp RG, Franklyn JA, Vittinghoff E, Walsh JP, Rodondi N; Thyroid Studies Collaboration. Subclinical hyperthyroidism and the risk of coronary heart disease and mortality. *Arch Intern Med*. 2012;**172**(10):799–809.
17. Ceresini G, Ceda GP, Lauretani F, Maggio M, Usberti E, Marina M, Bandinelli S, Guralnik JM, Valenti G, Ferrucci L. Thyroid status and 6-year mortality in elderly people living in a mildly iodine-deficient area: the aging in the Chianti Area Study. *J Am Geriatr Soc*. 2013;**61**(6):868–874.
18. Wells GA, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 29 May 2016.

19. Bauer DC, Ettinger B, Nevitt MC, Stone KL; Study of Osteoporotic Fractures Research Group. Risk for fracture in women with low serum levels of thyroid-stimulating hormone. *Ann Intern Med*. 2001;**134**(7):561–568.
20. Rodondi N, Newman AB, Vittinghoff E, de Rekeneire N, Satterfield S, Harris TB, Bauer DC. Subclinical hypothyroidism and the risk of heart failure, other cardiovascular events, and death. *Arch Intern Med*. 2005;**165**(21):2460–2466.
21. StataCorp. Stata 14 Base Reference Manual. Available at: http://www.stata.com/manuals14/u.pdf. Accessed 13 May 2017.
22. Siemieniuk RA, Agoritsas T, Manja V, Devji T, Chang Y, Bala MM, Thabane L, Guyatt GH. Transcatheter versus surgical aortic valve replacement in patients with severe aortic stenosis at low and intermediate risk: systematic review and meta-analysis. *BMJ*. 2016; **354**:i5130.
23. Schoenfeld D. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika*. 1980;**67**(1): 145–153.
24. Hofman A, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A, Ikram MA, Klaver CC, Nijsten TE, Peeters RP, Stricker BH, Tiemeier HW, Uitterlinden AG, Vernooij MW. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol*. 2013; **28**(11):889–926.
25. Lee JS, Buzková P, Fink HA, Vu J, Carbone L, Chen Z, Cauley J, Bauer DC, Cappola AR, Robbins J. Subclinical thyroid dysfunction and incident hip fracture in older adults. *Arch Intern Med*. 2010; **170**(21):1876–1883.
26. Finigan J, Greenfield DM, Blumsohn A, Hannon RA, Peel NF, Jiang G, Eastell R. Risk factors for vertebral and nonvertebral fracture over 10 years: a population-based study in women. *J Bone Miner Res*. 2008;**23**(1):75–85. doi:10.1359/jbmr.070814
27. Imaizumi M, Akahoshi M, Ichimaru S, Nakashima E, Hida A, Soda M, Usa T, Ashizawa K, Yokoyama N, Maeda R, Nagataki S, Eguchi K. Risk for ischemic heart disease and all-cause mortality in subclinical hypothyroidism. *J Clin Endocrinol Metab*. 2004;**89**(7): 3365–3370.
28. Goichot B, Sapin R, Schlienger JL. Subclinical hyperthyroidism: considerations in defining the lower limit of the thyrotropin reference interval. *Clin Chem*. 2009;**55**(3):420–424.
29. Walsh JP, Bremner AP, Bulsara MK, O'Leary P, Leedman PJ, Feddema P, Michelangeli V. Subclinical thyroid dysfunction as a risk factor for cardiovascular disease. *Arch Intern Med*. 2005; **165**(21):2467–2472.
30. Boekholdt SM, Titan SM, Wiersinga WM, Chatterjee K, Basart DC, Luben R, Wareham NJ, Khaw KT. Initial thyroid status and cardiovascular risk factors: the EPIC-Norfolk prospective population study. *Clin Endocrinol (Oxf)*. 2010;**72**(3):404–410.
31. Svare A, Nilsen TI, Asvold BO, Forsmo S, Schei B, Bjoro T Langhammer A. Does thyroid function influence fracture risk? Prospective data from the HUNT2 study, Norway. *Eur J Endocrinol*. 2013;**169**(6):845–852.
32. Gussekloo J, van Exel E, de Craen AJ, Meinders AE, Frölich M, Westendorp RG. Thyroid status, disability and cognitive function, and survival in old age. *JAMA*. 2004;**292**(21):2591–2599.
33. Nanchen D, Gussekloo J, Westendorp RG, Stott DJ, Jukema JW, Trompet S, Ford I, Welsh P, Sattar N, Macfarlane PW, Mooijaart SP, Rodondi N, de Craen AJ; PROSPER Group. Subclinical thyroid dysfunction and the risk of heart failure in older persons at high cardiovascular risk. *J Clin Endocrinol Metab*. 2012;**97**(3):852–861.
34. Hofman A, Brusselle GG, Darwish Murad S, van Duijn CM, Franco OH, Goedegebure A, Ikram MA, Klaver CC, Nijsten TE, Peeters RP, Stricker BH, Tiemeier HW, Uitterlinden AG, Vernooij MW. The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol*. 2015;**30**(8):661–708.
35. Biondi B. The normal TSH reference range: what has changed in the last decade? *J Clin Endocrinol Metab*. 2013;**98**(9):3584–3587.
36. Åsvold BO, Vatten LJ, Midthjell K, Bjøro T. Serum TSH within the reference range as a predictor of future hypothyroidism and hyperthyroidism: 11-year follow-up of the HUNT Study in Norway. *J Clin Endocrinol Metab*. 2012;**97**(1):93–99.
37. de Jong FJ, Masaki K, Chen H, Remaley AT, Breteler MM, Petrovitch H, White LR, Launer LJ. Thyroid function, the risk of dementia and neuropathologic changes: the Honolulu-Asia aging study. *Neurobiol Aging*. 2009;**30**(4):600–606.
38. Yeap BB, Alfonso H, Chubb SA, Puri G, Hankey GJ, Flicker L, Almeida OP. Higher free thyroxine levels predict increased incidence of dementia in older men: the Health in Men Study. *J Clin Endocrinol Metab*. 2012;**97**(12):E2230–E2237.
39. Klein I, Ojamaa K. Thyroid hormone and the cardiovascular system. *N Engl J Med*. 2001;**344**(7):501–509.
40. Bassett JH, Boyde A, Howell PG, Bassett RH, Galliford TM, Archanco M, Evans H, Lawson MA, Croucher P, St Germain DL, Galton VA, Williams GR. Optimal bone strength and mineralization requires the type 2 iodothyronine deiodinase in osteoblasts. *Proc Natl Acad Sci USA*. 2010;**107**(16):7604–7609.