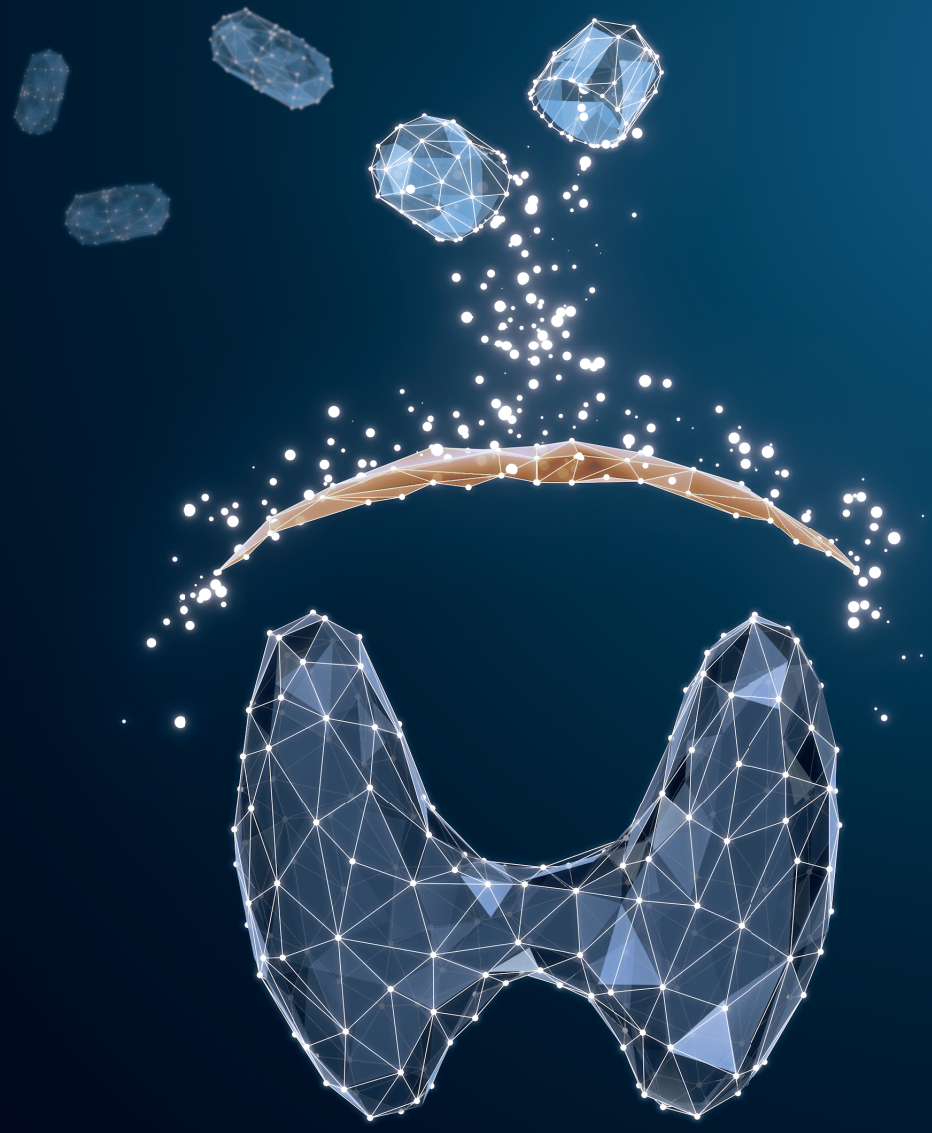
Subclinical hypothyroidism in community-dwelling older people: consequences and treatment outcomes

Du Puy, R.S.

Subclinical Hypothyroidism in Community-Dwelling Older People

Consequences and Treatment Outcomes

Robert Du Puy



Subclinical hypothyroidism in community-dwelling older people

Consequences and treatment outcomes

Robert S. Du Puy

Subclinical hypothyroidism in community-dwelling older people
Consequences and treatment outcomes

The work in this thesis was conducted at the department of Public Health and Primary Care of the Leiden University Medical Center

ISBN: 978-94-6361-574-7

Cover artwork: Optima Grafische Communicatie, Rotterdam, The Netherlands
Layout: Optima Grafische Communicatie, Rotterdam, The Netherlands
Printing: Optima Grafische Communicatie, Rotterdam, The Netherlands

Financial support for printing of this thesis was kindly provided by the SBOH (employer GP trainees in the Netherlands

Subclinical hypothyroidism in community-dwelling older people
Consequences and treatment outcomes

Proefschrift

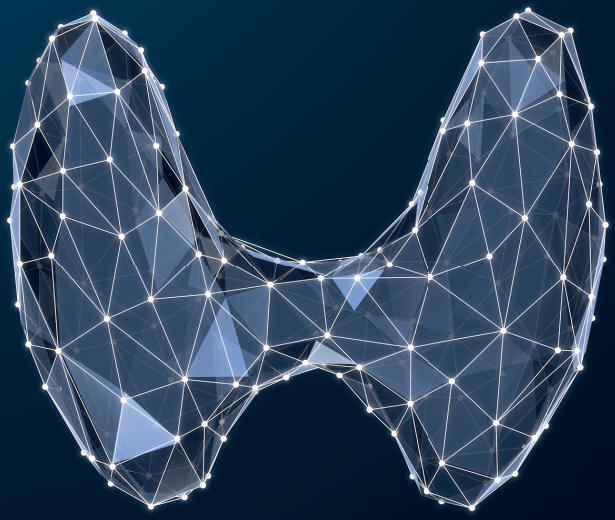
ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 23 september 2021
klokke 12.30 uur

door

Robert Simon Du Puy
Geboren te Voorburg in 1989

Promotor:	Prof. dr. J. Gussekloo
Co-promotores:	Dr. S.P. Mooijaart Dr. R.K.E. Poortvliet
Leden promotiecommissie	Prof. dr. W.P. Achterberg Prof. dr. A.R. Cappola, University of Pennsylvania, United States of America Prof. dr. M.G.M. Olde Rikkert, Radboud Universiteit Prof. dr. T.J.M. Verheij, Universiteit van Utrecht

CONTENTS



Chapter 1

General introduction

The consequences and treatment considerations regarding subclinical hypothyroidism in community-dwelling older people (arbitrarily 65 years and older) are long-debated subjects.[1] In the absence of robust and conclusive scientific evidence, and with decades of arguably trial-and-error experimentation with levothyroxine treatment, opinions about whether monitoring is required and whether it requires levothyroxine treatment (artificial thyroid hormone) or not, are spread out wide and bolstered. Patients, physicians and researchers often adhere to different schools of thought and can be found voicing opinions on both ends of spectra. For instance, some consider subclinical hypothyroidism to be consequential to health and longevity, others don't. Some believe it is a natural evolutionary process that may even be protective, others feel it's a sign of pre-clinical thyroid dysfunction. Finally, some consider treatment with levothyroxine to be necessary, others maintain that the condition is best left alone. Decades of research and experience have to date not been able to bridge these divides. To help address these issues, this thesis set out to investigate whether older, community-dwelling people with subclinical hypothyroidism have a neutral, beneficial or detrimental condition, and whether treatment with levothyroxine is beneficial.

Thyroid function

The thyroid gland, a small organ producing thyroid hormones, is found in all vertebrates, and plays a crucial role throughout an organism's lifecycle by regulating all metabolic processes.[2] It is often underappreciated when functioning normally. However, it becomes a powerful modifier of health and life when under- or overactive. For instance, in animals physiologic (non-diseased) low blood levels of thyroid hormone are implicated as the cause of longevity in long-lived species of squirrels and bats,[3-6] while physiologic high levels signals tadpoles to metamorphose into frogs.[7]

However, pathophysiologic (diseased or abnormal) low blood levels of thyroid hormones unquestionably lead to unhealthy fetal brain development[8] whilst pathophysiologic high levels, especially when all stored hormone is released at once, is not rarely fatal.[9] Differentiating between normal and diseased states of thyroid function is vital to good health and, although the effects of thyroid (dys-)function have been recorded and investigated for at least 4700 years,[10] the exact role in human health remains ever to be elucidated.

Thyroid hormone synthesis starts with cleaving thyroglobulin into up to 120 individual tyrosine molecules.[11] With the help of the thyreoperoxidase (TPO) enzymes, iodine atoms provided through the diet and the now freed tyrosine molecules, are combined to produce the two primary thyroid hormones thyroxine (T4, with 4 iodine atoms) and triiodothyronine (T3, with 3 iodine atoms), ready to enter the circulation.

The production and secretion of thyroid hormones are under strict regulation (figure 1). One particular section of the brain, the hypothalamus, acts much like a thermostat to the central heating of a house. It determines the set-point of thyroid hormone regulation by releasing Thyrotropin-Releasing Hormone (TRH) in specific intervals. The secreted TRH stimulates a second organ in the skull, the pituitary gland, to release its stimulatory Thyroid-Stimulating Hormone (TSH or Thyrotropin). Additionally, both organs respond to changing blood levels of circulating T3 and T4. When secreted TSH binds with the TSH-receptor on the thyroid follicles the thyroid gland releases its T4 and T3 in a 14:1 ratio.[12] However, since 99.98% of T4 and 99.70% of T3 is immediately bound by proteins in the blood only 0.02% of free T4 (fT4) and 0.3% of free T3 (fT3) is active throughout the body.[13] Only the free serum T4 and T3 concentrations determine the biological activity. The increased concentrations of fT3 and in particular fT4 then signal the hypothalamus and pituitary gland to decrease their production of TRH and TSH respectively and this eventually decreases the availability of circulating levels of fT4 and fT3, maintaining balance.

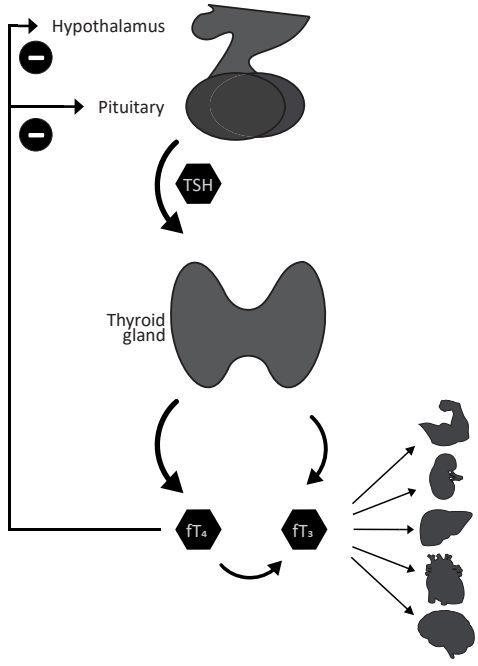


Figure 1. The Hypothalamic-Pituitary-Thyroid Axis. Serum measurements of TSH and fT4 are used by physicians as proxies of thyroid function. Adapted from [15].

In virtually all tissues in the human body binding of T3 increases the speed of protein synthesis and substrate turnover, effectively making the cell, tissue or organ work harder.[14] A majority of the locally available fT3 is not dependent on the secreted fraction from the

thyroid gland however, which is after all very small, but on the conversion of fT4 to fT3 in the liver and kidneys through deiodinase enzymes that allow for organ-specific regulation of extrathyroidal fT3 production. Because of these tissue-specific differences in fT3, clinicians mostly rely on the measurement of TSH and fT4 to assess overall thyroid function in the entire body and to weigh treatment modalities.

Abnormal thyroid function

If, however, the hypothalamic-pituitary-thyroid axis functions abnormally, and an overproduction (hyperthyroidism) or underproduction (hypothyroidism) of thyroid hormone exists, overt thyroid disorders start to emerge. The multifarious effects of thyroid hormones in all organs and tissues explain the pleiotropic clinical signs and symptoms. For instance, in overt hyperthyroidism signs and symptoms may include palpitations, tremors, anxiety, weight loss, heat intolerance, mood instability and shortness of breath.[16] Conversely, overt hypothyroidism may manifest predominantly with fatigue, cold intolerance, weight gain, anaemia, constipation and dry skin.[17] Since thyroid status cannot always be predicted from the often non-specific clinical signs and symptoms, a clinician will, when thyroid disease is suspected, order laboratory analyses of serum samples to investigate whether thyroid dysfunction is present and in what form. Normal thyroid function (euthyroidism) is defined biochemically according to the local laboratory reference ranges (normal reference ranges for TSH commonly between 0.4 and 4.0 mIU/L and for fT4 commonly between 9.0 and 24.0 pmol/L) although inter-individual variances are common with a genetically determined unique TSH and fT4 setpoint.

Overt or clinical thyroid dysfunction is diagnosed when serum analyses deviate from the normal reference ranges, irrespective of the underlying cause. Overt hypothyroidism is defined as a TSH above the laboratory reference range with an fT4 below the laboratory reference range. Overt hyperthyroidism is the exact opposite; defined as a low TSH with fT4 above the reference ranges. In general populations of iodine-sufficient parts of the world the prevalence of overt hypothyroidism ranges from 0.2% to 5.3% and for overt hyperthyroidism ranges from 0.2% to 1.3%.[18] Since the symptoms of overt thyroid disease can often be resolved and serum levels normalised, most international guidelines recommend routine and sometimes lifelong treatment of overt hypothyroidism with levothyroxine administration[19] and treatment of overt hyperthyroidism with either drugs, surgery or radioactive iodine administration.[20]

Subclinical hypothyroidism in old age

In between the two overt thyroid function disorders, but different from the normal euthyroid state, we define two additional subclinical thyroid dysfunctions that are much more common in the community. Subclinical hypothyroidism, defined as TSH levels above the reference

range but with normal circulating fT4 levels is prevalent in 4.0% to 15.4% of the general population[22-25] and patients are generally asymptomatic.[26] Subclinical hyperthyroidism is defined as TSH levels below the reference range but with normal circulating fT4 levels and is prevalent in 0.7% to 12.4% of the community.[25] At the time of writing this thesis, international experts have not reached a consensus on whether or how these subclinical thyroid diseases require medical treatment.

It has long been recognised that thyroid function disorders appear to occur more frequently in older people, in particular subclinical hypothyroidism (figure 2). The prevalence of subclinical hypothyroidism in women aged 60 years and older may be as high as 20% in the general population,[21] higher than that of overt hypothyroidism (0.2-5.3%), subclinical hyperthyroidism (0.6-9.8%) and overt hyperthyroidism (0.8-1.3%) combined. With increasing age more changes in thyroid anatomy, function and outcomes can be detected, yet how much this influences health and longevity is not fully understood.

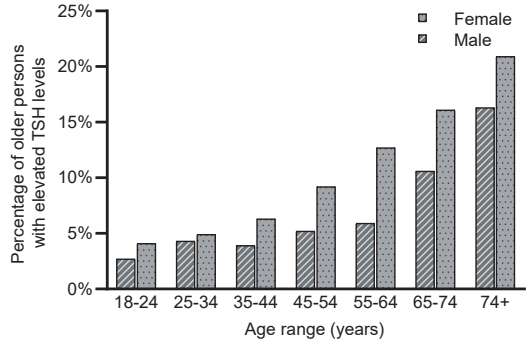


Figure 2. Percentage of community-dwelling older people with elevated TSH levels by age range. Adapted from [21].

Anatomically, over time, the thyroid gland reduces in volume, decreases its iodine uptake and decreases its T4 secretion.[27] Contrastingly, T4 and T3 clearance is decreased compensating for the loss of production. The natural course of thyroid function shows more variation in old age, with subclinical thyroid disorders being found more often, yet often reverting to a euthyroid state without intervention as well.[28] In older people clinical signs and symptoms are generally even more subtle, atypical or absent compared to younger age groups, rendering them unhelpful in physician decision making.[29] If findings are present at all these are not seldomly attributed to old age or confused with other comorbidities. Earlier studies yielded greatly conflicting results, with both increased, neutral and decreased risks or odds, for associations between subclinical hypothyroidism and cardiovascular disease, cognition, mood, lipid metabolism, quality of life, fractures and cause-specific and all-cause mortality. [30-32]

Subclinical hypothyroidism in old age: physicians' predicament

The medical care for community-dwelling older people with subclinical hypothyroidism is not only the responsibility of General Practitioners (GPs) but may be accommodated by different medical practitioners (including but not limited to family physicians, endocrinologists, internists and thyroidologists). In this thesis the combined group of professionals that deal with community-dwelling older people with subclinical hypothyroidism will be referred to as physicians henceforth.

In light of the arguments presented in the aforementioned paragraphs, physicians often struggle with managing subclinical hypothyroidism in older people. While patients with an elevated TSH laboratory finding look to their physician for support, without compelling clinical trial experience, uncertainty about the prognosis and with guidelines offering equivocal directions, physicians demonstrate significant variability in diagnostic and treatment strategies;[33] 'Should I handle this elevated TSH level finding as a disease? Or do I reassure my patient that this is normal and does not explain any potential symptoms? Do I monitor thyroid function over time? Or should I start levothyroxine treatment instead?'.

In an ageing population, it is to be expected that physicians will have to deal with similar dilemmas more and more. As the numbers of older people with subclinical hypothyroidism rise, with better access to their lab results, more influence from sometimes indiscriminate online sources and backing from assertive advocacy groups, lacking evidence-based support as a physician may shift the balance of the narrative and increase tension of the doctor-patient relationship. Simultaneously physicians have a societal responsibility to allocate the limited healthcare resources available (i.e. time, money, referrals and manpower) equally and just, based on the available evidence base and agreements with healthcare insurers and the health ministries. Evidently, future physicians would benefit greatly from scientifically grounded guiding principles and treatment strategies.

Aim

The overall aim of this thesis is to establish the effects of subclinical hypothyroidism on clinically and biologically relevant outcomes, and the effects of levothyroxine treatment on similar endpoints, in community-dwelling older people.

Research objectives

1. To establish whether subclinical hypothyroidism in older patients is a neutral, beneficial or detrimental condition by establishing if subclinical hypothyroidism in older people is associated with:
   a. clinically relevant outcomes
   b. biologically relevant outcomes
2. To investigate if levothyroxine treatment for subclinical hypothyroidism in older people provides long-term benefits in clinically or biologically relevant outcomes

Part 1: Consequences of subclinical hypothyroidism in older people

The first part will focus on observational studies and aims to establish whether subclinical hypothyroidism in older patients is a neutral, beneficial or detrimental condition by establishing if subclinical hypothyroidism in older people is associated with clinically relevant outcomes or associated with biologically relevant outcomes.

Chapter 2 discusses the findings from an extensive Individual Patient Data Meta-Analysis conducted in cohorts of people aged 80 years and older from the Netherlands, the United Kingdom, New Zealand and Japan. The study focuses on discovering associations between thyroid dysfunction as assessed traditionally using TSH and fT4 (i.e. overt hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism and overt hyperthyroidism) and activities of daily living, cognitive function, physical function, depressive symptoms and mortality. Its advantage compared to earlier studies is the harmonization and standardization of determinants, including TSH and fT4, and outcomes.

In **chapter 3** we examine associations between elevated levels of thyreoperoxidase antibodies, a biological marker commonly found in patients with thyroid dysfunction, and mortality, incident thyroid disease, changes in physical function, disability in activities of daily living, cognitive function and depressive symptoms using the data from the Leiden 85-plus Study. In **chapter 4**, an Individual Patient Data Meta-Analysis using data from sixteen international cohorts, we investigate cross-sectional odds of having, and longitudinal hazards of developing, anaemia in different categories of thyroid dysfunction. As well as stratifying the findings for different sexes, ethnicities and ages.

Part 2: Treatment outcomes for subclinical hypothyroidism in older people

The second part of the thesis will shift the focus to experimental studies and aims to discover if levothyroxine treatment for subclinical hypothyroidism in older people provides long-term benefits in clinically or biologically relevant outcomes

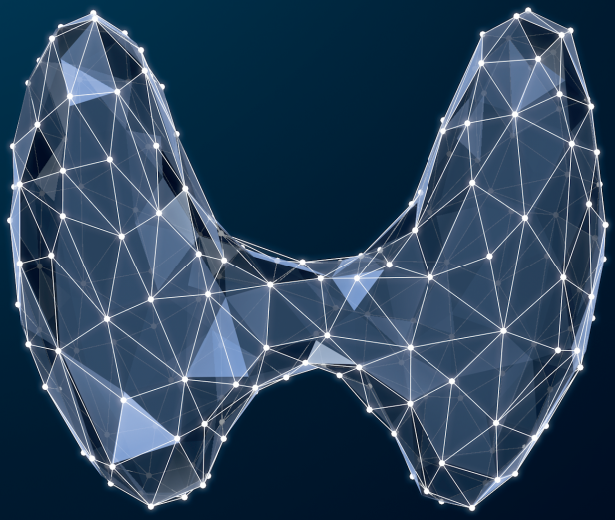
Chapter 5 presents the results from an international randomised controlled trial (TRUST study), focusing on the effects of levothyroxine treatment for subclinical hypothyroidism in >65-year-old people. In **chapter 6** and **chapter 7** we describe the design and results of a complementary international randomised controlled trial aimed specifically for the treatment of people with subclinical hypothyroidism aged 80 years and older in the IEMO 80+ thyroid trial. **Chapter 8** presents the results from levothyroxine treatment for subclinical hypothyroidism on haemoglobin levels in people aged 65 years and older using data from both the TRUST study and IEMO 80+ thyroid trial.

Finally, **chapter 9** presents a summary and discussion of the main findings of this thesis, complete with implications for physician practice and future perspectives for thyroid research.

REFERENCES:

1. Peeters RP, Brito JP. Subclinical hypothyroidism: to treat or not to treat? *European journal of endocrinology.* 2020;183(6):D15-D24.
2. Nilsson M, Fagman H. Development of the thyroid gland. *Development (Cambridge, England).* 2017;144(12):2123-2140.
3. Buffenstein R. The naked mole-rat: a new long-living model for human aging research. *The journals of gerontology Series A, Biological sciences and medical sciences.* 2005;60(11):1369-1377.
4. Buffenstein R, Woodley R, Thomadakis C, Daly TJ, Gray DA. Cold-induced changes in thyroid function in a poikilothermic mammal, the naked mole-rat. *American journal of physiology Regulatory, integrative and comparative physiology.* 2001;280(1):R149-155.
5. Hulbert AJ, Hinds DS, MacMillen RE. Minimal metabolism, summit metabolism and plasma thyroxine in rodents from different environments. *Comparative biochemistry and physiology A, Comparative physiology.* 1985;81(3):687-693.
6. Kwiecinski GG, Damassa DA, Gustafson AW. Control of sex steroid-binding protein (SBP) in the male little brown bat: relationship of plasma thyroxine levels to the induction of plasma SBP in immature males. *The Journal of endocrinology.* 1986;110(2):271-278.
7. Sachs LM, Buchholz DR. Insufficiency of Thyroid Hormone in Frog Metamorphosis and the Role of Glucocorticoids. *Frontiers in endocrinology.* 2019;10:287.
8. Pemberton HN, Franklyn JA, Kilby MD. Thyroid hormones and fetal brain development. *Minerva ginecologica.* 2005;57(4):367-378.
9. Devereaux D, Tewelde SZ. Hyperthyroidism and thyrotoxicosis. *Emergency medicine clinics of North America.* 2014;32(2):277-292.
10. Niazi AK, Kalra S, Irfan A, Islam A. Thyroidology over the ages. *Indian journal of endocrinology and metabolism.* 2011;15(Suppl 2):S121-126.
11. Boron W, Boron WF, Boulpaep EL. *Medical Physiology: A Cellular and Molecular Approach*. W.B. Saunders; 2003.
12. Wiersinga WM, Duntas L, Fadeyev V, Nygaard B, Vanderpump MP. 2012 ETA Guidelines: The Use of L-T4 + L-T3 in the Treatment of Hypothyroidism. *European thyroid journal.* 2012;1(2):55-71.
13. Cooper DS, Sipos JA, Cooper DS, Sipos J. *Medical Management of Thyroid Disease.* Boca Raton: CRC Press; 2018.
14. Brent GA. Mechanisms of thyroid hormone action. *The Journal of clinical investigation.* 2012;122(9):3035-3043.
15. Berkow R. *The Merck Manual of Medical Information*. Vol 2: Elsevier Health Services; 1997.
16. Goichot B, Caron P, Landron F, Bouee S. Clinical presentation of hyperthyroidism in a large representative sample of outpatients in France: relationships with age, aetiology and hormonal parameters. *Clinical endocrinology.* 2016;84(3):445-451.
17. Chaker L, Bianco AC, Jonklaas J, Peeters RP. Hypothyroidism. *Lancet (London, England).* 2017;390(10101):1550-1562.
18. Taylor PN, Albrecht D, Scholz A, et al. Global epidemiology of hyperthyroidism and hypothyroidism. *Nature reviews Endocrinology.* 2018;14(5):301-316.
19. Jonklaas J, Bianco AC, Bauer AJ, et al. Guidelines for the treatment of hypothyroidism: prepared by the american thyroid association task force on thyroid hormone replacement. *Thyroid : official journal of the American Thyroid Association.* 2014;24(12):1670-1751.

20. Ross DS, Burch HB, Cooper DS, et al. 2016 American Thyroid Association Guidelines for Diagnosis and Management of Hyperthyroidism and Other Causes of Thyrotoxicosis. *Thyroid : official journal of the American Thyroid Association.* 2016;26(10):1343-1421.
21. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Archives of internal medicine.* 2000;160(4):526-534.
22. Bemben DA, Hamm RM, Morgan L, Winn P, Davis A, Barton E. Thyroid disease in the elderly. Part 2. Predictability of subclinical hypothyroidism. *The Journal of family practice.* 1994;38(6):583-588.
23. Biondi B, Cappola AR, Cooper DS. Subclinical Hypothyroidism: A Review. *Jama.* 2019;322(2):153-160.
24. Garmendia Madariaga A, Santos Palacios S, Guillen-Grima F, Galofre JC. The incidence and prevalence of thyroid dysfunction in Europe: a meta-analysis. *The Journal of clinical endocrinology and metabolism.* 2014;99(3):923-931.
25. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *The Journal of clinical endocrinology and metabolism.* 2002;87(2):489-499.
26. Duntas LH. Thyroid Function in Aging: A Discerning Approach. *Rejuvenation research.* 2018;21(1):22-28.
27. Ajish TP, Jayakumar RV. Geriatric thyroidology: An update. *Indian journal of endocrinology and metabolism.* 2012;16(4):542-547.
28. Somwaru LL, Rariy CM, Arnold AM, Cappola AR. The natural history of subclinical hypothyroidism in the elderly: the cardiovascular health study. *The Journal of clinical endocrinology and metabolism.* 2012;97(6):1962-1969.
29. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. *Endocrine reviews.* 2008;29(1):76-131.
30. Hennessey JV, Espaillat R. Diagnosis and Management of Subclinical Hypothyroidism in Elderly Adults: A Review of the Literature. *Journal of the American Geriatrics Society.* 2015;63(8):1663-1673.
31. Barbesino G. Thyroid Function Changes in the Elderly and Their Relationship to Cardiovascular Health: A Mini-Review. *Gerontology.* 2019;65(1):1-8.
32. Leng O, Razvi S. Hypothyroidism in the older population. *Thyroid research.* 2019;12:2.
33. den Elzen WP, Lefèbre-van de Fliert AA, Virgini V, et al. International variation in GP treatment strategies for subclinical hypothyroidism in older adults: a case-based survey. *The British journal of general practice : the journal of the Royal College of General Practitioners.* 2015;65(631):e121-132.



Chapter 2

Outcomes of thyroid dysfunction in people aged 80 years and older: an individual patient data meta-analysis of four prospective studies (TULIPS consortium)

Robert S Du Puy
Rosalinde KE Poortvliet
Simon P Mooijaart
Wendy PJ den Elzen
Carol Jagger
Simon HS Pearce
Yasumichi Arai
Nobuyoshi Hirose
Ruth Teh
Oliver Menzies
Anna Rolleston
Ngaire Kerse
Jacobijn Gussekloo for the TULIPS Consortium.

Thyroid, 2020; ahead of print

DOI: 10.1089/thy.2020.0567

ABSTRACT

Background

Subclinical and overt thyroid dysfunction is easily detectable, often modifiable and, in younger age groups, has been associated with clinically relevant outcomes. Robust associations in very old persons however are currently lacking. This study aimed to investigate the associations between (sub-)clinical thyroid dysfunction and disability in daily living, cognitive function, depressive symptoms, physical function and mortality in people aged 80 years and older.

Methods

Four prospective cohorts participating in the Towards Understanding Longitudinal International older People Studies (TULIPS) consortium were included. We performed a two-step Individual Participant Data meta-analysis on source data from community-dwelling participants aged 80 years and older from the Netherlands, New Zealand, United Kingdom and Japan. Outcome measures included disability in daily living (disability in activities of daily living questionnaires), cognitive function (MMSE), depressive symptoms (GDS), physical function (grip strength) at baseline and after 5 years of follow-up, and all- cause five-year mortality.

Results

Of the total 2,116 participants at baseline (mean age 87 years, range 80-109 years), 105 participants (5.0%) were overtly hypothyroid, 136 (6.4%) subclinically hypothyroid, 1,811 (85.6%) euthyroid, 60 (2.8%) subclinically hyperthyroid and 4 (0.2%) overtly hyperthyroid. Participants with thyroid dysfunction at baseline had non-significantly different activities of daily living scores compared to euthyroid participants at baseline and had similar MMSE scores, GDS scores and grip strength. There was no difference in the change of any of these functional measures in participants with thyroid dysfunction during five years of follow-up. Compared to the euthyroid participants, no 5-year survival differences were identified in participants with overt hypothyroidism (Hazard Ratio [HR] 1.0, 95% Confidence Interval [95%CI] 0.6 to 1.6), subclinical hypothyroidism (HR 0.9, 95%CI 0.7 to 1.2), subclinical hyperthyroidism (HR 1.1, 95%CI 0.8 to 1.7) and overt hyperthyroidism (HR 1.5, 95%CI 0.4 to 5.9). Results did not differ after excluding participants using thyroid-influencing medication.

Conclusions

In community-dwelling people aged 80 years and older, (sub-)clinical thyroid dysfunction was not associated with functional outcomes or mortality and may therefore be of limited clinical significance.

INTRODUCTION

Circulating levels of thyroid hormones influence cell metabolism of nearly all types of tissue in the human body.[1] Consequently, in young and middle-aged individuals both hypo- and hyperthyroidism have been associated with several adverse clinical outcomes.[2-5]

However, evidence is starting to accumulate that adverse health associations of thyroid dysfunction, found in younger age groups, cannot be extrapolated to persons aged 80 years and over.[6] Possibly due to changes in thyroid metabolism associated with aging. For example, in a recent meta-analysis of 13 studies investigating the association between subclinical hypothyroidism and cognitive decline or the risk of dementia, a significant risk was identified in participants younger than 75 years of age but not in participants older than 75 years.[7] In addition, clinical outcomes such as cardiovascular risk [8-10] and activities of daily living,[11] associated with thyroid dysfunction in younger age groups, have not been replicated in older adults.[12]

Researchers are currently unable to reach robust conclusions for the people aged 80 years and over due to methodological or logistic difficulties, such as low numbers of older participants included. In this study, we combine the individual participant data from four large, international prospective cohorts of community-dwelling very old persons, to investigate associations between (sub-)clinical thyroid dysfunction and disabilities in activities of daily living, functioning (including cognitive function, depressive symptoms and physical function) and mortality in people aged 80 years and older.

MATERIALS AND METHODS

Study Population

The Towards Understanding Longitudinal International older People Studies (TULIPS) Consortium was established in 2014 and set out to investigate determinants of successful aging and health outcomes in older persons using the data from large-scale, prospective, population-based cohort studies. The data for the analyses in this manuscript were provided by four international cohort studies: The Leiden 85-plus Study, The Life and Living in Advanced Age: a cohort study in New Zealand (LiLACS NZ) Study, Newcastle 85+ Study and Tokyo Centenarian Study (TCS).

The Leiden 85-plus Study

All inhabitants of Leiden, The Netherlands, reaching the age of 85 between September 1997 and September 1999 (N=705) were eligible for participation of which 599 participated.[13]

During a 5-year follow-up, participants were visited for interviews, performing of functional tests and the collection of venous blood samples. Thyroid stimulating hormone (TSH) and free thyroxine (fT4) at baseline were available for 553 participants.

LiLACS NZ Study

LiLACS NZ consists of two cohorts running in parallel, one with exclusively Māori (the indigenous people of New Zealand) and one with non-Māori participants.[14] All potential participants living in the Lakes or Bay of Plenty District Health Board areas in 2010 (aged 85 in 2010 for non-Māori or aged 80-90 in 2010 for Māori) were eligible (N=1,636) of which 937 were enrolled.[15] Participants were visited annually for a structured questionnaire, a physical health assessment and venous blood samples. TSH and fT4 at baseline were available for 195 participants in the Māori cohort and 352 participants in the non-Māori cohort.

Newcastle 85+ Study

All citizens born in 1921 registered with a participating general practitioner in Newcastle upon Tyne or North Tyneside primary care trusts, the United Kingdom, were approached in 2006 for participation (N=1,470), of which 851 were enrolled with complete health assessment and record review.[16] Participants received visits at baseline, at 18 months, at 36 months and at 60 months of follow-up, for the questionnaire collection, measurements and function tests and blood tests.[17] TSH and fT4 at baseline were available for 763 participants.

Tokyo Centenarian Study

Out of an estimated 1,735 centenarians living in Tokyo, Japan, 1,194 centenarians (68.8%) were randomly selected and approached with a survey between 2000 and 2002, of which 304 agreed to participate.[18] All participants were visited for blood sample collection, cognitive, mental and physical functioning assessments. TSH and fT4 at baseline were available for 251 participants.

Data collected

For all cohorts, baseline data per participant on socio-demographic characteristics (age, sex, ethnicity, current tobacco and alcohol use and education level) were collected. Follow-up data for thyroid-influencing medication use (levothyroxine, anti-thyroid medication, amiodarone, lithium, glucocorticosteroids or interferon-α) and the outcome variables described below were collected at every visit.

Thyroid Status

In the Leiden 85-plus study non-fasting blood samples were drawn between 8 and 11 am. and plasma TSH and fT4 levels were analysed using a Elecsys 2010 automated system (Hitachi, Tokyo, Japan).[19] For the LiLACS NZ cohort fasting blood samples were either drawn at time

of interview or in a local public laboratory. Thyroid assays were analysed using a Cobas immunoassay system (Roche, Switzerland). In the Newcastle 85+ Study, blood was drawn after an overnight fast before 10.30 am. and analysed using a Centaur chemiluminescent immunoassay system (Siemens, UK).[20] In the TCS study, non-fasting blood samples were obtained at time of interview and were stored at -80°C until the subsequent assay was performed. Serum TSH and fT4 concentrations were measured using a chemiluminescent enzyme immunoassay with a Lumipulse Forte fully automated system (Fuji Rebio, Tokyo, Japan).

For all studies, the reference ranges for TSH were defined as 0.3 mIU/L to 4.8 mIU/L for serum TSH and as 13 pmol/L to 23 pmol/L (1.01 to 1.79 ng/dL) for fT4. Five clinical strata of thyroid function at baseline were used to classify participants; overt hypothyroidism (TSH > 4.8 and fT4 < 13), subclinical hypothyroidism (TSH > 4.8 and fT4 within reference range), euthyroidism (0.3 ≤ TSH ≤ 4.8), subclinical hyperthyroidism (TSH < 0.3 and fT4 within reference range) and overt hyperthyroidism (TSH < 0.3 and fT4 > 23).

Outcome measures

Disability in activities of daily living

Disability in activities of daily living was measured in the Leiden 85-plus Study using the Groningen Activity Restriction Scale (GARS [21], range 18-72, higher summed scores more disability) consisting of 18 items of daily function and independent living. In the Newcastle 85+ study a summed disability score derived from 17 activities of daily living was used (range 0-17, higher summed scores more disability).[20] For the LiLACS NZ study the Nottingham Extended Activities of Daily Living Index (NEADL, range 0-22, higher summed scores less disability) [22] was used. For the TCS study, activities of daily living was assessed at baseline using the Barthel Index (BI, range 0-100, higher summed index scores less disability).[23] All disability in activities of daily living questionnaires used have been shown to be valid and reliable indicators of physical disability in older persons.[24-27]

Cognitive function

In all studies, cognitive function was assessed with the Mini-Mental State Examination (MMSE) questionnaire at every visit (maximum score of 30). The MMSE questionnaire has previously been validated in this age group.[28] Higher scores indicate better cognitive functioning. MMSE scores in TCS were available at baseline.

Depressive symptoms

The Geriatric Depression Scale (GDS-15) was used to evaluate a presence of depressive symptoms in the Leiden 85-plus, LiLACS NZ, and Newcastle 85+ cohort.[29] This questionnaire contains 15 items adding up to maximum summed score of 15, with higher scores indicating more depressive symptoms. The GDS-15 questionnaire has previously been validated in this

age group.[29] In the Leiden 85-plus and Newcastle 85+ cohort the GDS was not undertaken in individuals with established cognitive impairment. Depressive symptoms were not assessed in TCS.

Physical function

Hand grip strength measurements in kilogram were used as a proxy for overall muscle strength [30] and hand-held dynamometry has been validated in this age group for measuring muscle strength.[31] A Jamar hand dynamometer (Sammons Preston INC., Illinois, USA) was used in the Leiden 85-plus study. A Takei hand dynamometer Grip-D (Takei Scientific Instruments Co., Niigata-City, Japan) was used in the LiLACS NZ and Newcastle 85+ cohort. Grip strength was not assessed in TCS.

Mortality

Date of death was collected from the records of Statistics Netherlands (Centraal Bureau voor de Statistiek, The Hague) for the Leiden 85-plus Study. General practitioner records and mortality administrative data (held by the Ministry of Health) were consulted for the LiLACS NZ study. Information on mortality for the Newcastle 85+ Study was provided by NHS Digital, UK. For TCS, all-cause mortality was ascertained by telephone contact or mail survey conducted every 12 months.[32] Survival time was calculated as the time between the baseline visit and date of death or censored at 5-years if death had not yet occurred.

Statistical analyses

All measurements are analysed at yearly intervals from baseline (i.e. after 1, 2, 3, 4 and 5 years). Measurements for Phase 2 of the Newcastle 85+ Study, however, were taken 18 months post baseline and, to allow linear mixed model analyses, were assumed constant at the 2-year analysis. Results were reported separately for the Māori and non-Māori cohorts to do justice to potential differences in Māori and non-Māori preferences and practices in research.[33]

A two-stage Individual Participant Data (IPD) meta-analysis approach was employed for all analyses.[34,35] All variables for the analyses were coded similarly for each study, analysed on a study level using the same syntax and finally pooled for meta-analysis.

Study level analyses

Continuous variables were described as mean with standard deviation or as median with interquartile range (IQR), where appropriate, and categorical variables were presented as frequency with percentage of the total.

Five-year survival analysis was carried out using Cox proportional-hazards regression models, corrected for age and sex, and presented as a hazard ratio (HR) with 95% confidence intervals

(95%CI). The proportionality of hazards assumption was checked. Repeated measurements linear mixed-effect models using restricted maximum likelihood estimation and an unstructured covariance matrix, corrected for age and sex, were used to assess the associations between thyroid function and MMSE, GDS, grip strength and disability in activities of daily living.

Disability in activities of daily living scores were inverted for TCS (BI), and for LiLACS NZ, the mean was subtracted from the scores (NEADL), to standardise the direction of the scales, a prerequisite for pooling.[36] Considering that for disability in daily living the same outcome was measured using different scales in the cohorts, standardised mean differences (SMDs) were calculated for each study using independent-groups pretest- posttest calculations.[37] By dividing the mean outcome difference between a thyroid dysfunction group and the euthyroid group with the standard deviation among the participants, the size of the between group effect in each study is normalised relative to the variability observed in the study (See supplementary material).

All analyses used the euthyroid participants as the reference group or used TSH levels as a continuous independent variable.

Pooled analyses

All cohort effect estimates were subsequently pooled using random-effects models with inverse-variance weighting and summarised in forest plots. Inconsistency between cohorts due to heterogeneity was quantified using the I^2-statistic and found not to be important for all summarised values ($I^2 < 40\%$).

Sensitivity analysis

All participants (n=236, 11%) using any form or combination of thyroid influencing medication (levothyroxine, anti-thyroid medication, amiodarone, lithium, glucocorticosteroids or interferon-a) were excluded for additional sensitivity analyses. The study-level analyses were performed using IBM SPSS Statistics version 22.0 (IBM, Armond, NY, USA). All pooled analyses were performed using Review Manager 5.3 (The Cochrane Collaboration, Copenhagen, Denmark).[38]

RESULTS

Table 1 displays the baseline characteristics of the combined study population and for the cohorts separately. The combined cohort consisted of 2,116 participants. Mean participant age was 87 (range 79 to 109) years and 1,311 (62%) were female. The median follow-up time over all cohorts was 5.3 (IQR 2.6 to 6.9) years.

Table 1. Baseline characteristics of study participants.

Study	Total	Leiden 85-plus Study	LiLACS NZ Study		Newcastle 85+ Study	Tokyo Centenarian Study
			Māori	Non-Māori		
N=	2,116	555	195 (35.7%)	352 (64.3%)	763	251
Study Population:		Longitudinal cohort of adults aged 85 (born between 1912 and 1914) in Leiden, the Netherlands	Longitudinal cohort of Māori aged 80-90 (born between 1920 and 1930) in New Zealand	Longitudinal cohort of Non-Māori aged 85 (born in 1925) in New Zealand	Longitudinal cohort in over-85s (born in 1921) in Newcastle upon Tyne, UK	Dynamic cohort study of centenarians (born before 1902) living in Tokyo, Japan
Socio-demographic characteristics:						
Age, mean (range)	87 (80-109)	85	82 (80-84)	85 (84-86)	85 (85 -86)	102 (100-109)
Female, no. (%)	1,311 (62.0)	366 (65.9)	108 (55.4)	177 (50.3)	465 (60.3)	195 (77.7)
Follow-up time, median (IQR), years	5.3 (2.6-6.9)	5.4 (2.7-8.6)	6.0 (4.1-6.2)	6.0 (4.6-6.2)	5.4 (2.6-7.7)	2.1 (1.0-3.8)
Current smoker, no. (%)	169 (8.0)	88 (15.9)	20 (10.3)	16 (4.5)	42 (5.5)	3 (1.2)
Current alcohol user, no. (%)	1,156 (54.6)	279 (50.3)	106 (54.4)	257 (73.0)	464 (60.8)	50 (19.9)
Laboratory measurements, median (IQR):						
TSH, mIU/L	2.1 (1.3-3.2)	1.8 (1.2-2.9)	2.1 (1.3-3.2)	2.6 (1.7-3.6)	2.0 (1.2-3.1)	2.3 (1.6-3.9)
fT4, pmol/L	14.9 (13.2-16.5)	14.3 (12.8-15.8)	15.2 (14.0-16.6)	15.1 (13.9-16.9)	15.0 (14.0-17.0)	13.4 (11.7-14.8)
Thyroid function, no. (%):						
Overt hypothyroidism	105 (5.0)	37 (6.7)	4 (2.1)	10 (2.8)	19 (2.5)	35 (13.9)
Subclinical hypothyroidism	136 (6.4)	29 (5.2)	12 (6.2)	30 (8.5)	55 (7.2)	10 (4.0)
Euthyroidism	1,811 (85.6)	471 (84.9)	172 (88.2)	307 (87.2)	656 (86.0)	205 (81.7)
Subclinical hyperthyroidism	60 (2.8)	16 (2.9)	7 (3.6)	5 (1.4)	31 (4.1)	1 (0.4)
Overt hyperthyroidism	4 (0.2)	2 (0.4)	-	-	2 (0.3)	-

Functional parameters, median (IQR):						
Disability in activities of daily living		28 (22-40)	19 (16-20)	19 (17-20)	3.0 (1-7)	45 (15-80)
Disability in activities of daily living questionnaire		GARS	NEADL	NEADL	ADL sum scores	BI
Mini Mental State Examination score	27 (23-29)	26 (22-28)	28 (26-29)	28 (27-29)	28 (25-29)	15 (7-21)
Geriatric Depression Scale score	1 (1-3)	2 (1-3)	2 (1-3)	2 (1-3)	1 (1-2)	-
Grip Strength (Kg)	21 (16-28)	21 (17-28)	25 (19-30)	23 (18-31)	19 (15-26)	-

Abbreviations: IQR, interquartile range; GARS, Groningen Activity Restriction Scale; NEADL, Nottingham Extended Activities of Daily Living; BI, Barthel Index.

The median serum TSH level in the combined cohort was 2.1 (IQR 1.3 to 3.2) mIU/L and median serum fT4 was 14.9 (IQR 13.2 to 16.5) pmol/L, both within the normal reference ranges. One thousand eight hundred and eleven (85.6%) participants were euthyroid. Thyroid dysfunction was found in 305 (14.4%) participants; 136 (6.4%) participants with subclinical hypothyroidism, 105 (5.0%) with overt hypothyroidism, 60 (2.8%) with subclinical hyperthyroidism and 4 (0.2%) with overt hyperthyroidism. In the hypothyroidism categories 18 (17.1%) participants with overt hypothyroidism and 41 (30.1%) of participants with subclinical hypothyroidism used a preparation of thyroid hormone at any point during follow-up.

The median MMSE score was 27 points (IQR 23 to 29) and median GDS score 1 point (IQR 1 to 2). Median disability in activities of daily living scores were 28 (IQR 22 to 40) on the GARS questionnaire (Leiden 85-plus Study); 19 (IQR 16 to 20) for the Māori subgroup and 19 (IQR 17 to 20) for the non-Māori subgroup on the NEADL questionnaire (LiLACS NZ); 3.0 (IQR 1 to7) on the ADL sum scores (Newcastle 85+ Study) and 45 (IQR 15 to 80) on the Barthel Index questionnaire (TCS).

Figure 1 shows activities of daily living scores according to clinical thyroid state at baseline and after 5 years of follow-up. At baseline no significant standardised mean differences in activity of daily living scores were identified in the clinical strata with thyroid dysfunction compared to the euthyroid participants; mean difference 0.01 (95%CI -0.19 to 0.21) points for overt hypothyroidism, -0.00 (95%CI -0.18 to 0.17) points for subclinical hypothyroidism, 0.00 (95%CI -0.25 to 0.26) points for subclinical hyperthyroidism and 0.01 (95%CI -0.97 to 1.00) points for overt hypothyroidism. Although all groups showed decline in activities of daily living, over time there were no significant standardised mean differences of decline in activities of daily living scores (SMD between -0.01 and 0.00, IQR between −0.99 and 0.97) between euthyroid participants and participants with thyroid dysfunction. In this study population a calculated SMD of -0.01 corresponds with a 3.5-point decrease on the GARS scale.

At baseline, estimated marginal means of MMSE scores did not differ significantly between participants with thyroid dysfunction (means between 23.6 points [95%CI 18.2 to 29.0] and 26.3 points [95%CI 25.5 to 27.4]) and participants with euthyroidism (mean 23.7 points [95%CI 21.0 to 26.5]). Similarly, no differences were found for depressive symptom scores (means between 1.9 points [95%CI 1.5 to 2.3] and 5.1 points [95%CI 2.7 to 7.5] for participants with thyroid dysfunction, and 2.1 points [95%CI 1.4 to 2.7] for euthyroid participants) and grip strength (means between 18.1 kg [95%CI 9.8 to 26.4] and 26.0 kg [95%CI 22.6 to 29.5] for participants with thyroid dysfunction, and 24.1 kg [95%CI 23.1 to 25.1] for euthyroid participants) (Table 2).

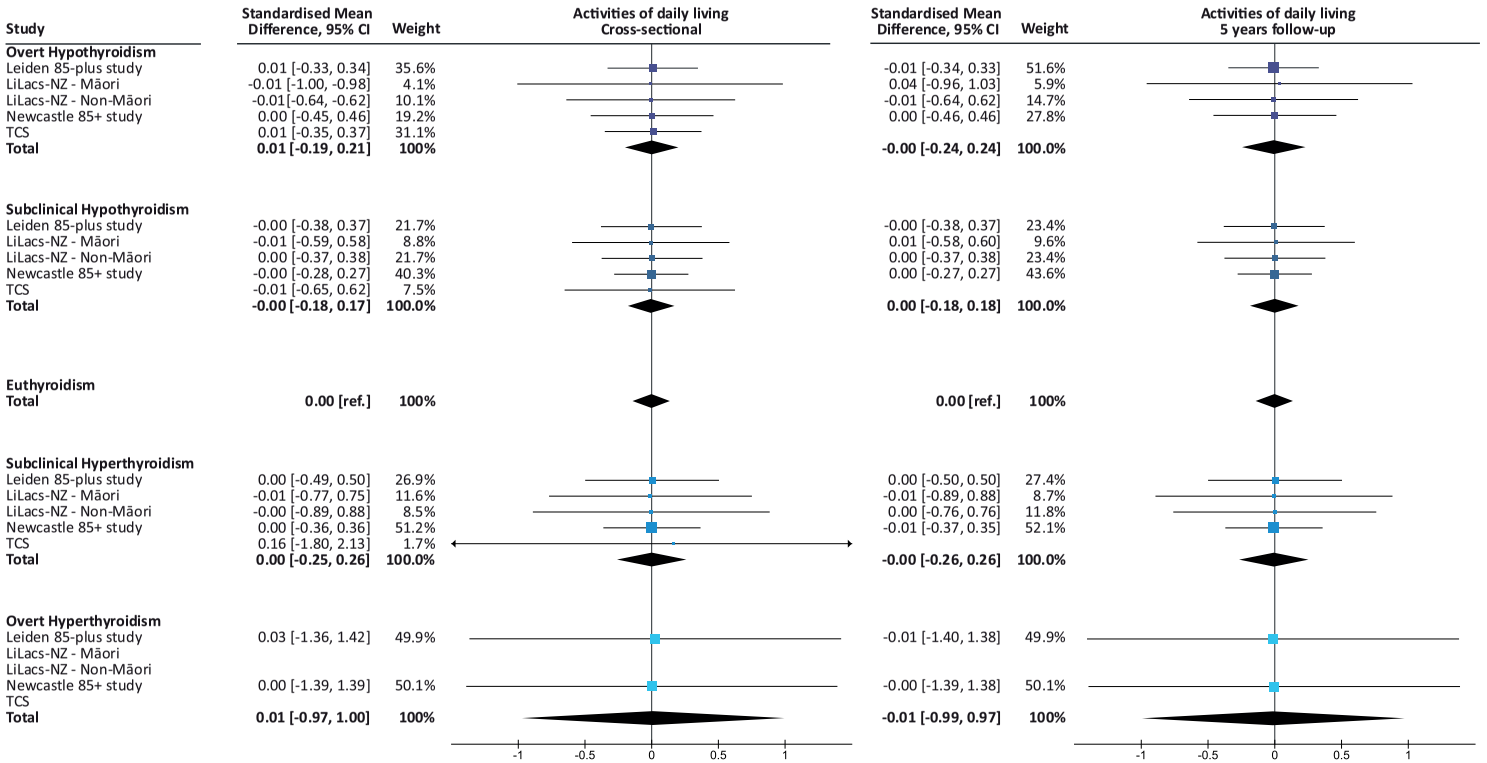


Figure 1. Activities of daily living per clinical thyroid stratum at baseline and after five years of follow-up. Standardized mean differences calculated per clinical thyroid stratum and per cohort using independent-groups pretest–post-test calculations compared with the euthyroid participants. Higher scores represent worse outcomes. Abbreviations: IQR, interquartile range; GARS, Groningen Activity Restriction Scale; NEADL, Nottingham Extended Activities of Daily Living; BI, Barthel Index.

Table 2. Functional parameters per clinical thyroid stratum at baseline.

	Overt Hypothyroidism (N=105)	**Subclinical hypothyroidism** (N=136)	**Euthyroidism** (N=1181)	**Subclinical hyperthyroidism** (N=60)	**Overt Hyperthyroidism** (N=4)
Baseline[a]					
Cognitive function (MMSE)	23.7 (20.4, 27.0)	25.5 (23.6, 27.5)	23.7 (21.0, 26.5)	26.3 (25.2, 27.4)	23.6 (18.2, 29.0)
Depressive symptoms (GDS)	2.1 (1.6, 2.7)	1.9 (1.5, 2.3)	2.1 (1.4, 2.7)	2.4 (1.9, 3.0)	5.1 (2.7, 7.5)
Physical function (grip strength)	26.0 (22.6, 29.5)	25.3 (24.0, 25.3)	24.1 (23.1, 25.1)	23.3 (21.3, 25.3)	18.1 (9.8, 26.4)
5-year follow-up[b]					
Cognitive function (MMSE)	-0.4 (-1.4, 0.5)	0.4 (-0.3, 1.0)	0 (ref.)	0.1 (-0.9, 1.1)	0.6 (-0.9, 2.1)
Depressive symptoms (GDS)	-0.1 (-0.6, 0.3)	0.0 (-0.4, 0.3)	0 (ref.)	0.3 (-0.2, 0.7)	1.7 (-0.2, 3.5)
Physical function (grip strength)	1.7 (-1.9, 5.2)	1.4 (-0.4, 2.1)	0 (ref.)	-0.7 (-2.5, 1.0)	-3.3 (-10.0, 11.3)

[a] Estimated Marginal Means (95% Confidence Intervals) estimated using Linear Mixed models per cohort, adjusted for baseline age and sex, and pooled using random-effects models with inverse-variance weighting.

[b] Beta's (95% Confidence Intervals) estimated using repeated measures Linear Mixed models per cohort, adjusting for baseline age and sex, and pooled using random effects models with inverse-variance weighting, representing change in outcome measure over five years of follow-up. Euthyroidism is the reference group.

Over time, no differences were identified in decline in MMSE scores for participants with thyroid dysfunction (means between -0.4 points less decline [95%CI -1.4 to 0.5] and 0.4 points more decline [95%CI -0.9 to 2.1]) compared to euthyroid patients. Nor were differences identified for depressive symptoms scores (means between -0.1 points less decline [95%CI -0.6 to 0.3] and 1.7 points more decline [95%CI -0.2 to 3.5] for thyroid dysfunction compared to euthyroid participants) and grip strength (means between -3.3 [95%CI -10.0 to 11.3] and 1.7 [95%CI -1.9 to 5.2] for thyroid dysfunction compared to euthyroid participants) (Table 2). Forest plots for all thyroid categories stratified per cohort are presented in the supplementary material.

Pooled HRs were calculated to estimate five-year mortality (Figure 2). No difference in survival risks were found for overt hypothyroidism (HR 0.98, 95%CI 0.63 to 1.51), subclinical hypothyroidism (HR 0.89, 95%CI 0.67 to 1.17), subclinical hyperthyroidism (HR 1.11, 95%CI 0.75 to 1.65) or overt hypothyroidism (HR 1.46, 95%CI 0.36 to 5.88). Similarly, in a cox-regression survival analysis across the clinical thyroid strata, TSH as a continuous variable did not predict 5-year mortality (HR 0.99, 95%CI 0.96 to 1.02).

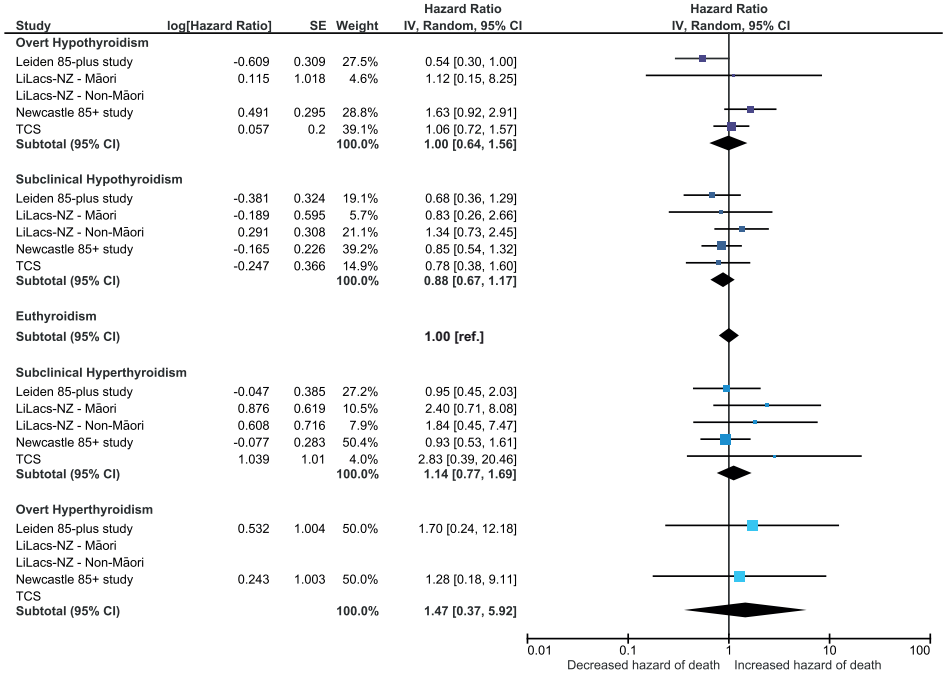


Figure 2. All-cause five-year mortality per clinical thyroid stratum at baseline. Pooled hazard ratios [95% CI] per clinical thyroid stratum estimated using Cox proportional hazards regression models adjusting for sex and age at baseline compared with participants with euthyroidism. When mortality rates were too, low hazard ratios could not be estimated. Abbreviations: CI, confidence interval.

In sensitivity analyses, excluding all participants (n=236, 11%) using any or a combination of thyroid influencing medications at any point during follow-up, did not alter the association estimates substantially in either direction of effect or significance (data not shown).

DISCUSSION

The findings of this large IPD meta-analysis of 4 prospective observational cohorts among community-dwelling participants aged 80 years and older are two-fold. First, there was no association of thyroid dysfunction with disability in activities of daily living, nor with cognition, physical function or depressive symptoms, at baseline or during follow-up. Second, thyroid dysfunction was not associated with increased mortality during an average 5 years of follow-up.

Earlier research investigating the outcomes of thyroid dysfunction in old age yielded greatly conflicting data. For example, some studies demonstrated decreased cognitive function,[39-41] increased depressive symptoms,[42,43] decreased physical function [44] and increased cardiovascular or all-cause mortality [45-47] in participants with subclinical or overt thyroid dysfunction. Others, however, were unable to replicate these findings.[10,48-58] Due to differences in, for example, the thyroid reference and age-ranges used, follow-up time chosen, the participants selected, study design employed, and outcome measures used, the aforementioned studies could not be pooled without suffering from significant methodological heterogeneity. This study was optimised to address these specific limitations, by pooling the results of four of the most comparable cohort studies of community-dwelling oldest old.

Our findings are in line with most previous publications from the individual datasets used in this study, such as an absence of association between thyroid function and disabilities in daily living, cognitive function and depressive symptoms.[19] Additionally no association was found between fT4 or TSH and all-cause mortality in earlier studies.[20,59] However, some nonunanimous results have been described using these datasets as well. For instance, in the Leiden 85-plus study increasing levels of TSH were associated with lower mortality rates.[19] It should be noted that the aggregate negative findings in our pooled analyses do not invalidate any earlier identified associations in any of the individual cohorts or other publications, as these may be a representation of associations in that specific population being sampled, at that time, using those specific methods, study designs and analyses. Nevertheless, the findings from this IPD meta-analysis, do more accurately reflect an estimation of effects for a more universal population of older people, with uniform definitions and follow-up times, by pooling the results and allowing statistically for sampling error and biases. These findings

suggest that the influence of thyroid dysfunction on disabilities in daily living, cognitive function, depressive symptoms, physical function and mortality in persons aged 80 years and older is limited, particularly in the subclinical thyroid disease subgroups.

For older age groups with subclinical thyroid dysfunction, evidence supporting beneficial effects from treatment has generally been of lower grade and contradictory, as acknowledged in international guidelines. Currently, these recommend routine or trial treatment in patients older than 65 with subclinical thyroid dysfunction, albeit individualised, gradual and closely monitored.[60,61] However, they simultaneously acknowledge the lack of evidence in scientific literature and call for large prospective studies with a long-term follow-up in older persons. Recently two international multicentre trials demonstrated no beneficial effects of levothyroxine treatment for subclinical hypothyroidism in persons aged 65 years and over, and in persons aged 80 years and over.[62,63]

The lack of associations between (mild) thyroid dysfunction and relevant outcomes described in this manuscript, and the absence of beneficial effects of levothyroxine treatment for subclinical hypothyroidism in older persons, reinforces the notion that for older persons with subclinical hypothyroidism, routine treatment does not seem indicated and that the thresholds for treatment initiation (i.e. currently the normal reference ranges) ought to be redefined. This is not necessarily the same as the definition of the reference ranges. Whether all thyroid reference ranges in aging need to be re-evaluated to distinguish physiologic states from pathophysiologic ones is still a much-debated issue and will require further evaluation. Future research, focusing on how thyroid-related (patho–) physiological processes change with aging and how these affect outcomes, are needed, including the influence of other established thyroid markers and treatment modalities. This study used data from some of the largest, unselected, population-based observational cohorts of community-dwelling persons aged 80 years and older to date, with considerable follow-up time and a substantial international representation of countries, cultures and persons. The IPD meta-analysis design allowed for harmonization of thyroid function categories, standardization of definitions and variables, and consistency in analyses. The two-step design allowed the presentation of overall results on an aggregate level, while maintaining maximum transparency on any potential inter-cohort variation, for example between Māori and non-Māori populations.

Some limitations do need to be acknowledged. First, TSH and fT4 were only available for the baseline measurements and it could be possible that thyroid function changed during the observational periods. However, a recent study in almost 3,000 community dwelling older persons has demonstrated high stability of thyroid function over 5 years of follow-up and therefore we estimate the odds of erroneous misclassification as small.[64] Second, the activities of daily living questionnaires were too heterogenous in scale and structure to pool in any

traditional way. Standardization with the Standardised Mean Difference- method allowed for methodologically sound pooling but may have introduced additional uncertainty to the analyses leading to underestimation of true associations.

In participants aged 80 years and older, (sub-)clinical thyroid dysfunction is not associated with functional outcomes or mortality and may therefore be of limited clinical significance.

REFERENCES

1. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism. *Physiol Rev.* 2014;94(2):355-382.
2. Beydoun MA, Beydoun HA, Kitner-Triolo MH, Kaufman JS, Evans MK, Zonderman AB. Thyroid hormones are associated with cognitive function: moderation by sex, race, and depressive symptoms. *J Clin Endocrinol Metab.* 2013;98(8):3470-3481.
3. Hage MP, Azar ST. The Link between Thyroid Function and Depression. *J Thyroid Res.* 2012;2012:590648.
4. Laulund AS, Nybo M, Brix TH, Abrahamsen B, Jorgensen HL, Hegedus L. Duration of thyroid dysfunction correlates with all-cause mortality. the OPENTHYRO Register Cohort. *PLoS One.* 2014;9(10):e110437.
5. Mainenti MR, Vigario PS, Teixeira PF, Maia MD, Oliveira FP, Vaisman M. Effect of levothyroxine replacement on exercise performance in subclinical hypothyroidism. *J Endocrinol Invest.* 2009;32(5):470-473.
6. Ruggeri RM, Trimarchi F, Biondi B. MANAGEMENT OF ENDOCRINE DISEASE: l-Thyroxine replacement therapy in the frail elderly: a challenge in clinical practice. *European journal of endocrinology / European Federation of Endocrine Societies.* 2017;177(4):R199-R217.
7. Pasqualetti G, Pagano G, Rengo G, Ferrara N, Monzani F. Subclinical Hypothyroidism and Cognitive Impairment: Systematic Review and Meta-Analysis. *J Clin Endocrinol Metab.* 2015;100(11):4240-4248.
8. Chaker L, Baumgartner C, den Elzen WP, et al. Subclinical Hypothyroidism and the Risk of Stroke Events and Fatal Stroke: An Individual Participant Data Analysis. *J Clin Endocrinol Metab.* 2015;100(6):2181-2191.
9. Pasqualetti G, Tognini S, Polini A, Caraccio N, Monzani F. Is subclinical hypothyroidism a cardiovascular risk factor in the elderly? *J Clin Endocrinol Metab.* 2013;98(6):2256-2266.
10. Razvi S, Shakoor A, Vanderpump M, Weaver JU, Pearce SH. The influence of age on the relationship between subclinical hypothyroidism and ischemic heart disease: a metaanalysis. *J Clin Endocrinol Metab.* 2008;93(8):2998-3007.
11. SO ES, Chan IT, Lobo Santos MA, et al. Impact of thyroid status and age on comprehensive geriatric assessment. *Endocrine.* 2014;47(1):255-265.
12. Mariotti S. Thyroid function and aging: do serum 3,5,3'-triiodothyronine and thyroid-stimulating hormone concentrations give the Janus response? *J Clin Endocrinol Metab.* 2005;90(12):6735-6737.
13. der Wiel AB, van Exel E, de Craen AJ, et al. A high response is not essential to prevent selection bias: results from the Leiden 85-plus study. *J Clin Epidemiol.* 2002;55(11):1119-1125.
14. Hayman KJ, Kerse N, Dyall L, et al. Life and living in advanced age: a cohort study in New Zealand--e Puawaitanga o Nga Tapuwae Kia Ora Tonu, LiLACS NZ: study protocol. *BMC Geriatr.* 2012;12:33.
15. Kerse N, Teh R, Moyes SA, et al. Cohort Profile: Te Puawaitanga o Nga Tapuwae Kia Ora Tonu, Life and Living in Advanced Age: a Cohort Study in New Zealand (LiLACS NZ). *Int J Epidemiol.* 2015;44(6):1823-1832.
16. Collerton J, Davies K, Jagger C, et al. Health and disease in 85 year olds: baseline findings from the Newcastle 85+ cohort study. *BMJ.* 2009;339:b4904.
17. Collerton J, Barrass K, Bond J, et al. The Newcastle 85+ study: biological, clinical and psychosocial factors associated with healthy ageing: study protocol. *BMC Geriatr.* 2007;7:14.

18. Gondo Y, Hirose N, Arai Y, et al. Functional status of centenarians in Tokyo, Japan: developing better phenotypes of exceptional longevity. *J Gerontol A Biol Sci Med Sci.* 2006;61(3):305-310.
19. Gussekloo J, van Exel E, de Craen AJ, Meinders AE, Frolich M, Westendorp RG. Thyroid status, disability and cognitive function, and survival in old age. *JAMA.* 2004;292(21):2591-2599.
20. Pearce SH, Razvi S, Yadegarfar ME, et al. Serum Thyroid Function, Mortality and Disability in Advanced Old Age: The Newcastle 85+ Study. *J Clin Endocrinol Metab.* 2016;101(11):4385-4394.
21. Kempen GI, Miedema I, Ormel J, Molenaar W. The assessment of disability with the Groningen Activity Restriction Scale. Conceptual framework and psychometric properties. *Social science & medicine (1982).* 1996;43(11):1601-1610.
22. Essink-Bot ML, Krabbe PF, Bonsel GJ, Aaronson NK. An empirical comparison of four generic health status measures. The Nottingham Health Profile, the Medical Outcomes Study 36-item Short-Form Health Survey, the COOP/WONCA charts, and the EuroQol instrument. *Med Care.* 1997;35(5):522-537.
23. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland state medical journal.* 1965;14:61-65.
24. Hopman-Rock M, van Hirtum H, de Vreede P, Freiberger E. Activities of daily living in older community-dwelling persons: a systematic review of psychometric properties of instruments. *Aging Clin Exp Res.* 2019;31(7):917-925.
25. Kingston A, Collerton J, Davies K, Bond J, Robinson L, Jagger C. Losing the ability in activities of daily living in the oldest old: a hierarchic disability scale from the Newcastle 85+ study. *PLoS One.* 2012;7(2):e31665.
26. Lin SY, Kerse N, McLean C, Moyes SA. Validation of quality of life and functional measures for older people for telephone administration. *J Prim Health Care.* 2010;2(1):35-42.
27. van Houwelingen AH, Cameron ID, Gussekloo J, et al. Disability transitions in the oldest old in the general population. The Leiden 85-plus study. *Age (Dordr).* 2014;36(1):483-493.
28. Heeren TJ, Lagaay AM, von Beek WC, Rooymans HG, Hijmans W. Reference values for the Mini-Mental State Examination (MMSE) in octo- and nonagenarians. *J Am Geriatr Soc.* 1990;38(10):1093-1096.
29. de Craen AJ, Heeren TJ, Gussekloo J. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. *Int J Geriatr Psychiatry.* 2003;18(1):63-66.
30. Bohannon RW. Dynamometer measurements of hand-grip strength predict multiple outcomes. *Percept Mot Skills.* 2001;93(2):323-328.
31. Abizanda P, Navarro JL, Garcia-Tomas MI, Lopez-Jimenez E, Martinez-Sanchez E, Paterna G. Validity and usefulness of hand-held dynamometry for measuring muscle strength in community-dwelling older persons. *Arch Gerontol Geriatr.* 2012;54(1):21-27.
32. Arai Y, Inagaki H, Takayama M, et al. Physical independence and mortality at the extreme limit of life span: supercentenarians study in Japan. *J Gerontol A Biol Sci Med Sci.* 2014;69(4):486-494.
33. Bishop R. Kaupapa Māori research: An indigenous approach to creating knowledge. *Robertson, N (Ed).* 1999;Maori and psychology : research and practice(The proceedings of a symposium sponsered by the Maori and Psychology Research Unit):1-6.
34. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ.* 2010;340:c221.
35. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof.* 2002;25(1):76-97.
36. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. In: The Cochrane Collaboration; 2011: www.handbook.cochrane.org. Accessed updated March 2011.

37. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods.* 2002;7(1):105-125.
38. *Review Manager (RevMan)* [computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre; 2014.
39. Kalmijn S, Mehta KM, Pols HA, Hofman A, Drexhage HA, Breteler MM. Subclinical hyperthyroidism and the risk of dementia. The Rotterdam study. *Clin Endocrinol (Oxf).* 2000;53(6):733-737.
40. Tan ZS, Beiser A, Vasan RS, et al. Thyroid function and the risk of Alzheimer disease: the Framingham Study. *Arch Intern Med.* 2008;168(14):1514-1520.
41. van Osch LA, Hogervorst E, Combrinck M, Smith AD. Low thyroid-stimulating hormone as an independent risk factor for Alzheimer disease. *Neurology.* 2004;62(11):1967-1971.
42. Blum MR, Wijsman LW, Virgini VS, et al. Subclinical Thyroid Dysfunction and Depressive Symptoms among the Elderly: A Prospective Cohort Study. *Neuroendocrinology.* 2016;103(3-4):291-299.
43. Medici M, Direk N, Visser WE, et al. Thyroid function within the normal range and the risk of depression: a population-based cohort study. *J Clin Endocrinol Metab.* 2014;99(4):1213-1219.
44. Bano A, Chaker L, Darweesh SK, et al. Gait patterns associated with thyroid function: The Rotterdam Study. *Sci Rep.* 2016;6:38912.
45. Haentjens P, Van Meerhaeghe A, Poppe K, Velkeniers B. Subclinical thyroid dysfunction and mortality: an estimate of relative and absolute excess all-cause mortality based on time-to-event data from cohort studies. *European journal of endocrinology / European Federation of Endocrine Societies.* 2008;159(3):329-341.
46. Iervasi G, Molinaro S, Landi P, et al. Association between increased mortality and mild thyroid dysfunction in cardiac patients. *Arch Intern Med.* 2007;167(14):1526-1532.
47. Sgarbi JA, Matsumura LK, Kasamatsu TS, Ferreira SR, Maciel RM. Subclinical thyroid dysfunctions are independent risk factors for mortality in a 7.5-year follow-up: the Japanese-Brazilian thyroid study. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(3):569-577.
48. de Jong FJ, den Heijer T, Visser TJ, et al. Thyroid hormones, dementia, and atrophy of the medial temporal lobe. *J Clin Endocrinol Metab.* 2006;91(7):2569-2573.
49. de Jongh RT, Lips P, van Schoor NM, et al. Endogenous subclinical thyroid disorders, physical and cognitive function, depression, and mortality in older individuals. *European journal of endocrinology / European Federation of Endocrine Societies.* 2011;165(4):545-554.
50. Grabe HJ, Volzke H, Ludemann J, et al. Mental and physical complaints in thyroid disorders in the general population. *Acta Psychiatr Scand.* 2005;112(4):286-293.
51. Ochs N, Auer R, Bauer DC, et al. Meta-analysis: subclinical thyroid dysfunction and the risk for coronary heart disease and mortality. *Ann Intern Med.* 2008;148(11):832-845.
52. Roberts LM, Pattison H, Roalfe A, et al. Is subclinical thyroid dysfunction in the elderly associated with depression or cognitive dysfunction? *Ann Intern Med.* 2006;145(8):573-581.
53. Simonsick EM, Chia CW, Mammen JS, Egan JM, Ferrucci L. Free Thyroxine and Functional Mobility, Fitness, and Fatigue in Euthyroid Older Men and Women in the Baltimore Longitudinal Study of Aging. *J Gerontol A Biol Sci Med Sci.* 2016;71(7):961-967.
54. Simonsick EM, Newman AB, Ferrucci L, et al. Subclinical hypothyroidism and functional mobility in older adults. *Arch Intern Med.* 2009;169(21):2011-2017.
55. Singh S, Duggal J, Molnar J, Maldonado F, Barsano CP, Arora R. Impact of subclinical thyroid disorders on coronary heart disease, cardiovascular and all-cause mortality: a meta-analysis. *Int J Cardiol.* 2008;125(1):41-48.

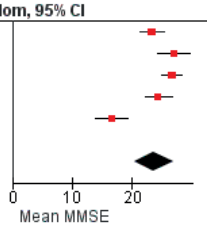
56. van der Cammen TJ, Mattace-Raso F, van Harskamp F, de Jager MC. Lack of association between thyroid disorders and Alzheimer's disease in older persons: a cross-sectional observational study in a geriatric outpatient population. *J Am Geriatr Soc.* 2003;51(6):884.
57. Virgini VS, Wijsman LW, Rodondi N, et al. Subclinical thyroid dysfunction and functional capacity among elderly. *Thyroid.* 2014;24(2):208-214.
58. Volzke H, Schwahn C, Wallaschofski H, Dorr M. Review: The association of thyroid dysfunction with all-cause and circulatory mortality: is there a causal relationship? *J Clin Endocrinol Metab.* 2007;92(7):2421-2429.
59. van Vliet NA, van der Spoel E, Beekman M, et al. Thyroid status and mortality in nonagenarians from long-lived families and the general population. *Aging (Albany NY).* 2017;9(10):2223-2234.
60. Biondi B, Bartalena L, Cooper DS, Hegedus L, Laurberg P, Kahaly GJ. The 2015 European Thyroid Association Guidelines on Diagnosis and Treatment of Endogenous Subclinical Hyperthyroidism. *Eur Thyroid J.* 2015;4(3):149-163.
61. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. *Eur Thyroid J.* 2013;2(4):215-228.
62. Mooijaart SP, Du Puy RS, Stott DJ, et al. Association Between Levothyroxine Treatment and Thyroid-Related Symptoms Among Adults Aged 80 Years and Older With Subclinical Hypothyroidism. *JAMA.* 2019:1-11.
63. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.
64. Roberts L, McCahon D, Johnson O, Haque MS, Parle J, Hobbs FR. Stability of thyroid function in older adults: the Birmingham Elderly Thyroid Study. *Br J Gen Pract.* 2018;68(675):e718-e726.

Supplementary data. Cross-sectional forest plots: MMSE

Estimated Marginal Means (95% Confidence Intervals) estimated using Linear Mixed models per cohort, adjusted for baseline age and sex, and pooled using random-effects models with inverse-variance weighting.

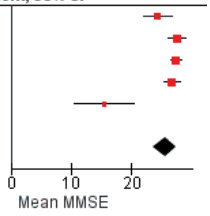
Overt hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	23.306	1.048	20.4%	23.31 [21.25, 25.36]
LiLacs-NZ - Māori	27	1.348	19.4%	27.00 [24.36, 29.64]
LiLacs-NZ - Non-Māori	26.7	0.873	20.9%	26.70 [24.99, 28.41]
Newcastle 85+ Study	24.379	1.151	20.1%	24.38 [22.12, 26.63]
TCS	16.556	1.378	19.3%	16.56 [13.86, 19.26]
Total (95% CI)			**100.0%**	**23.65 [20.35, 26.94]**



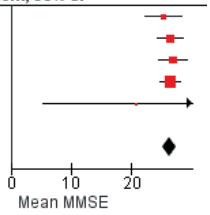
Subclinical hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	24.37	1.185	19.3%	24.37 [22.05, 26.69]
LiLacs-NZ - Māori	27.524	0.796	22.6%	27.52 [25.96, 29.08]
LiLacs-NZ - Non-Māori	27.419	0.504	24.7%	27.42 [26.43, 28.41]
Newcastle 85+ Study	26.694	0.667	23.6%	26.69 [25.39, 28.00]
TCS	15.476	2.547	9.7%	15.48 [10.48, 20.47]
Total (95% CI)			**100.0%**	**25.52 [23.56, 27.48]**



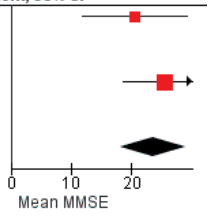
Subclinical hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	25.228	1.579	12.7%	25.23 [22.13, 28.32]
LiLacs-NZ - Māori	26.359	1.117	25.3%	26.36 [24.17, 28.55]
LiLacs-NZ - Non-Māori	26.863	1.233	20.8%	26.86 [24.45, 29.28]
Newcastle 85+ Study	26.43	0.882	40.7%	26.43 [24.70, 28.16]
TCS	20.764	7.898	0.5%	20.76 [5.28, 36.24]
Total (95% CI)			**100.0%**	**26.32 [25.22, 27.42]**



Overt hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	20.5	4.466	37.9%	20.50 [11.75, 29.25]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	25.474	3.49	62.1%	25.47 [18.63, 32.31]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**23.59 [18.20, 28.98]**

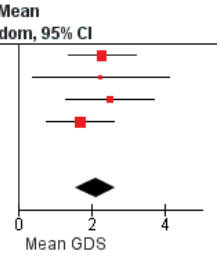


Supplementary data. Cross-sectional forest plots: GDS

Estimated Marginal Means (95% Confidence Intervals) estimated using Linear Mixed models per cohort, adjusted for baseline age and sex, and pooled using random-effects models with inverse-variance weighting.

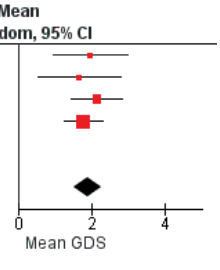
Overt hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	2.288	0.466	35.6%	2.29 [1.37, 3.20]
LiLacs-NZ - Māori	2.25	0.947	8.6%	2.25 [0.39, 4.11]
LiLacs-NZ - Non-Māori	2.5	0.611	20.7%	2.50 [1.30, 3.70]
Newcastle 85+ Study	1.696	0.47	35.0%	1.70 [0.77, 2.62]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**2.12 [1.58, 2.67]**



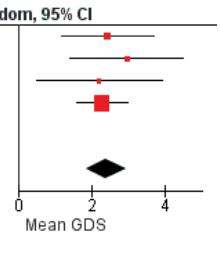
Subclinical hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	1.98	0.514	13.2%	1.98 [0.97, 2.99]
LiLacs-NZ - Māori	1.666	0.572	10.7%	1.67 [0.54, 2.79]
LiLacs-NZ - Non-Māori	2.151	0.35	28.5%	2.15 [1.47, 2.84]
Newcastle 85+ Study	1.777	0.271	47.6%	1.78 [1.25, 2.31]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**1.90 [1.53, 2.27]**



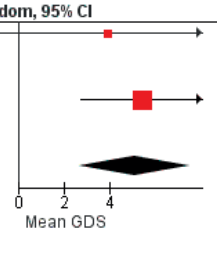
Subclinical hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	2.447	0.65	18.0%	2.45 [1.17, 3.72]
LiLacs-NZ - Māori	2.966	0.785	12.3%	2.97 [1.43, 4.50]
LiLacs-NZ - Non-Māori	2.214	0.871	10.0%	2.21 [0.51, 3.92]
Newcastle 85+ Study	2.287	0.357	59.6%	2.29 [1.59, 2.99]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**2.39 [1.85, 2.93]**



Overt hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	3.914	2.509	23.5%	3.91 [-1.00, 8.83]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	5.442	1.389	76.5%	5.44 [2.72, 8.16]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**5.08 [2.70, 7.47]**

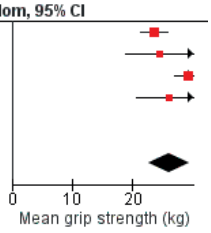


Supplementary data. Cross-sectional forest plots: Grip strength

Estimated Marginal Means (95% Confidence Intervals) estimated using Linear Mixed models per cohort, adjusted for baseline age and sex, and pooled using random-effects models with inverse-variance weighting.

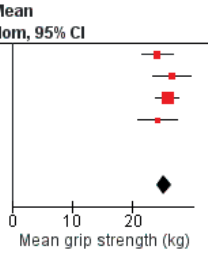
Overt hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	23.641	1.143	31.4%	23.64 [21.40, 25.88]
LiLacs-NZ - Māori	24.525	2.852	18.5%	24.52 [18.94, 30.11]
LiLacs-NZ - Non-Māori	29.21	1.143	31.4%	29.21 [26.97, 31.45]
Newcastle 85+ Study	26.176	2.827	18.6%	26.18 [20.64, 31.72]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**26.03 [22.60, 29.45]**



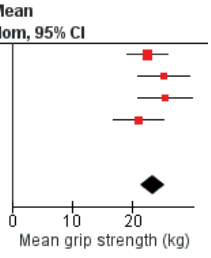
Subclinical hypothyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	24.09	1.32	24.2%	24.09 [21.50, 26.68]
LiLacs-NZ - Māori	26.557	1.618	16.1%	26.56 [23.39, 29.73]
LiLacs-NZ - Non-Māori	25.782	0.976	44.3%	25.78 [23.87, 27.69]
Newcastle 85+ Study	24.194	1.657	15.4%	24.19 [20.95, 27.44]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**25.25 [23.98, 26.53]**



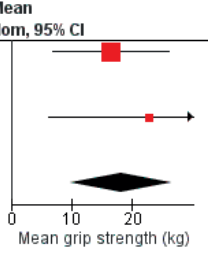
Subclinical hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	22.474	1.738	35.0%	22.47 [19.07, 25.88]
LiLacs-NZ - Māori	25.17	2.2	21.9%	25.17 [20.86, 29.48]
LiLacs-NZ - Non-Māori	25.383	2.312	19.8%	25.38 [20.85, 29.91]
Newcastle 85+ Study	21.048	2.129	23.3%	21.05 [16.88, 25.22]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**23.31 [21.29, 25.32]**



Overt hyperthyroidism

Study or Subgroup	Mean	SE	Weight	Mean IV, Random, 95% CI
Leiden 85-plus Study	16.5	4.918	74.9%	16.50 [6.86, 26.14]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	22.869	8.499	25.1%	22.87 [6.21, 39.53]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**18.10 [9.75, 26.44]**



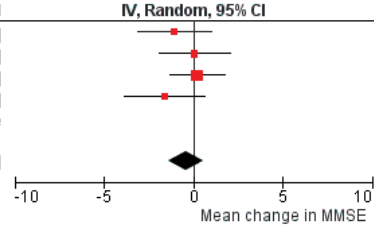
Supplementary data. Longitudinal forest plots: MMSE

Beta's (95% Confidence Intervals) estimated using repeated measures Linear Mixed models per cohort, adjusting for baseline age and sex, and pooled using random effects models with inverse-variance weighting, representing change in outcome measure over five years of follow-up. Euthyroidism is the reference group.

Overt hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-1.062	1.064	21.1%	-1.06 [-3.15, 1.02]
LiLacs-NZ - Māori	0.067	1.031	22.5%	0.07 [-1.95, 2.09]
LiLacs-NZ - Non-Māori	0.213	0.788	38.5%	0.21 [-1.33, 1.76]
Newcastle 85+ Study	-1.629	1.159	17.8%	-1.63 [-3.90, 0.64]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**-0.42 [-1.38, 0.54]**

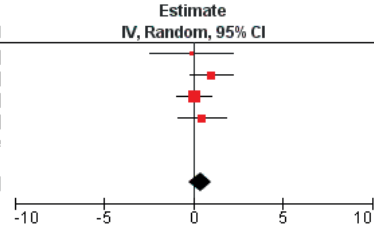
Heterogeneity: Tau² = 0.00; Chi² = 2.32, df = 3 (P = 0.51); I² = 0%
Test for overall effect: Z = 0.85 (P = 0.39)



Subclinical hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-0.116	1.198	7.2%	-0.12 [-2.46, 2.23]
LiLacs-NZ - Māori	1.013	0.62	27.0%	1.01 [-0.20, 2.23]
LiLacs-NZ - Non-Māori	0.025	0.489	43.4%	0.03 [-0.93, 0.98]
Newcastle 85+ Study	0.484	0.681	22.4%	0.48 [-0.85, 1.82]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**0.38 [-0.25, 1.02]**

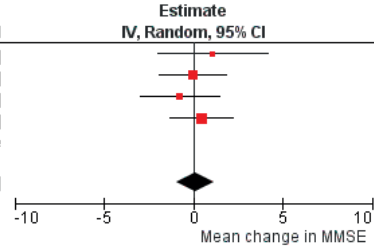
Heterogeneity: Tau² = 0.00; Chi² = 1.76, df = 3 (P = 0.62); I² = 0%
Test for overall effect: Z = 1.19 (P = 0.23)



Subclinical hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	1.081	1.578	11.3%	1.08 [-2.01, 4.17]
LiLacs-NZ - Māori	-0.047	0.945	31.5%	-0.05 [-1.90, 1.81]
LiLacs-NZ - Non-Māori	-0.775	1.138	21.7%	-0.78 [-3.01, 1.46]
Newcastle 85+ Study	0.456	0.891	35.5%	0.46 [-1.29, 2.20]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**0.10 [-0.94, 1.14]**

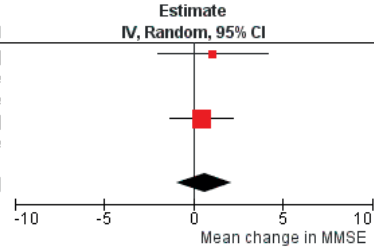
Heterogeneity: Tau² = 0.00; Chi² = 1.16, df = 3 (P = 0.76); I² = 0%
Test for overall effect: Z = 0.19 (P = 0.85)



Overt hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	1.081	1.578	24.2%	1.08 [-2.01, 4.17]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	0.456	0.891	75.8%	0.46 [-1.29, 2.20]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**0.61 [-0.91, 2.13]**

Heterogeneity: Tau² = 0.00; Chi² = 0.12, df = 1 (P = 0.73); I² = 0%
Test for overall effect: Z = 0.78 (P = 0.43)



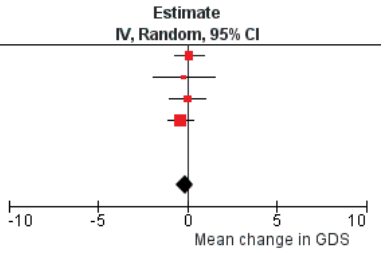
Supplementary data. Longitudinal forest plots: GDS

Beta's (95% Confidence Intervals) estimated using repeated measures Linear Mixed models per cohort, adjusting for baseline age and sex, and pooled using random effects models with inverse-variance weighting, representing change in outcome measure over five years of follow-up. Euthyroidism is the reference group.

Overt hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	0.108	0.437	28.8%	0.11 [-0.75, 0.96]
LiLacs-NZ - Māori	-0.204	0.881	7.1%	-0.20 [-1.93, 1.52]
LiLacs-NZ - Non-Māori	0.011	0.514	20.8%	0.01 [-1.00, 1.02]
Newcastle 85+ Study	-0.374	0.356	43.4%	-0.37 [-1.07, 0.32]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**-0.14 [-0.60, 0.32]**

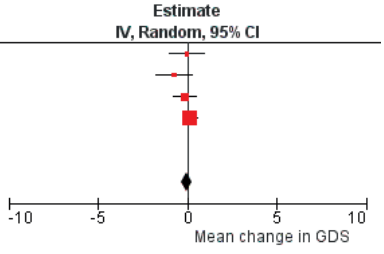
Heterogeneity: Tau² = 0.00; Chi² = 0.85, df = 3 (P = 0.84); I² = 0%
Test for overall effect: Z = 0.61 (P = 0.54)



Subclinical hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-0.018	0.495	10.4%	-0.02 [-0.99, 0.95]
LiLacs-NZ - Māori	-0.741	0.522	9.3%	-0.74 [-1.76, 0.28]
LiLacs-NZ - Non-Māori	-0.159	0.327	23.8%	-0.16 [-0.80, 0.48]
Newcastle 85+ Study	0.115	0.212	56.5%	0.12 [-0.30, 0.53]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**-0.04 [-0.36, 0.27]**

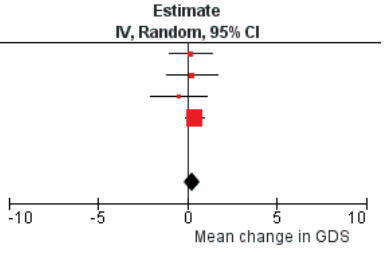
Heterogeneity: Tau² = 0.00; Chi² = 2.47, df = 3 (P = 0.48); I² = 0%
Test for overall effect: Z = 0.27 (P = 0.78)



Subclinical hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	0.182	0.619	13.8%	0.18 [-1.03, 1.40]
LiLacs-NZ - Māori	0.243	0.734	9.8%	0.24 [-1.20, 1.68]
LiLacs-NZ - Non-Māori	-0.458	0.809	8.1%	-0.46 [-2.04, 1.13]
Newcastle 85+ Study	0.395	0.278	68.3%	0.40 [-0.15, 0.94]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**0.28 [-0.17, 0.73]**

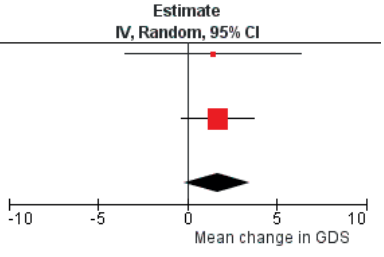
Heterogeneity: Tau² = 0.00; Chi² = 1.03, df = 3 (P = 0.79); I² = 0%
Test for overall effect: Z = 1.23 (P = 0.22)



Overt hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	1.425	2.504	14.5%	1.43 [-3.48, 6.33]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	1.705	1.031	85.5%	1.71 [-0.32, 3.73]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**1.66 [-0.20, 3.53]**

Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.92); I² = 0%
Test for overall effect: Z = 1.75 (P = 0.08)



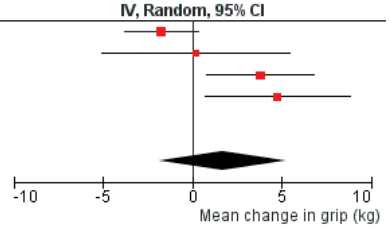
Supplementary data. Longitudinal forest plots: Grip strength

Beta's (95% Confidence Intervals) estimated using repeated measures Linear Mixed models per cohort, adjusting for baseline age and sex, and pooled using random effects models with inverse-variance weighting, representing change in outcome measure over five years of follow-up. Euthyroidism is the reference group.

Overt hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-1.739	1.069	30.2%	-1.74 [-3.83, 0.36]
LiLacs-NZ - Māori	0.21	2.692	19.3%	0.21 [-5.07, 5.49]
LiLacs-NZ - Non-Māori	3.805	1.541	27.1%	3.81 [0.78, 6.83]
Newcastle 85+ Study	4.771	2.055	23.5%	4.77 [0.74, 8.80]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**1.67 [-1.87, 5.20]**

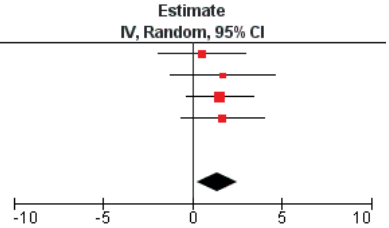
Heterogeneity: Tau² = 9.63; Chi² = 13.22, df = 3 (P = 0.004); I² = 77%
Test for overall effect: Z = 0.92 (P = 0.36)



Subclinical hypothyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	0.537	1.26	21.9%	0.54 [-1.93, 3.01]
LiLacs-NZ - Māori	1.718	1.498	15.5%	1.72 [-1.22, 4.65]
LiLacs-NZ - Non-Māori	1.544	0.965	37.4%	1.54 [-0.35, 3.44]
Newcastle 85+ Study	1.69	1.176	25.2%	1.69 [-0.61, 3.99]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**1.39 [0.23, 2.54]**

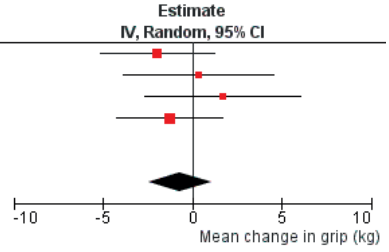
Heterogeneity: Tau² = 0.00; Chi² = 0.60, df = 3 (P = 0.90); I² = 0%
Test for overall effect: Z = 2.35 (P = 0.02)



Subclinical hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-1.943	1.627	30.5%	-1.94 [-5.13, 1.25]
LiLacs-NZ - Māori	0.341	2.134	17.7%	0.34 [-3.84, 4.52]
LiLacs-NZ - Non-Māori	1.709	2.222	16.4%	1.71 [-2.65, 6.06]
Newcastle 85+ Study	-1.256	1.51	35.4%	-1.26 [-4.22, 1.70]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**-0.70 [-2.46, 1.06]**

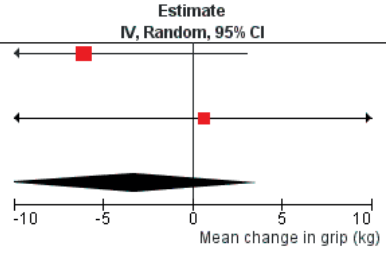
Heterogeneity: Tau² = 0.00; Chi² = 2.13, df = 3 (P = 0.55); I² = 0%
Test for overall effect: Z = 0.78 (P = 0.44)



Overt hyperthyroidism

Study or Subgroup	Estimate	SE	Weight	Estimate IV, Random, 95% CI
Leiden 85-plus Study	-6.074	4.641	58.0%	-6.07 [-15.17, 3.02]
LiLacs-NZ - Māori	0	0		Not estimable
LiLacs-NZ - Non-Māori	0	0		Not estimable
Newcastle 85+ Study	0.655	5.45	42.0%	0.66 [-10.03, 11.34]
TCS	0	0		Not estimable
Total (95% CI)			**100.0%**	**-3.25 [-10.17, 3.68]**

Heterogeneity: Tau² = 0.00; Chi² = 0.88, df = 1 (P = 0.35); I² = 0%
Test for overall effect: Z = 0.92 (P = 0.36)



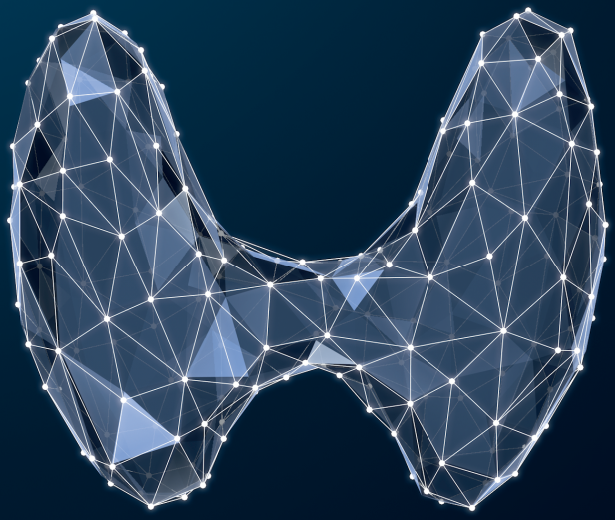
Supplementary method. SMD calculation

Standardised Mean Difference (SMD) and corresponding standard errors (SE) were calculated using an independent-groups pretest-posttest formula:

$$SMD = \frac{M_{D,E} - M_{D,C}}{SD_{pooled}} \qquad \text{with } SD_{pooled} = \sqrt{\frac{(N_E - 1)SD_E^{\,2} + (N_C - 1)SD_C^{\,2}}{N_{tot}}}$$

$$\text{and } SE = \sqrt{\frac{N_{tot}}{N_e * N_c} + \frac{SMD^2}{2(N_{tot} - 2)}}$$

SMD = standardised mean difference, $M_{D,E}$ = mean difference post - pre-test in thyroid dysfunction group, $M_{D,C}$ = mean difference post - pre-test in euthyroid group, N_E = participants in thyroid dysfunction category, SD_E = standard deviation in thyroid dysfunction group, N_C = participants in euthyroid category, SD_C = standard deviation in euthyroid category, N_{tot} = total number of participants.



Chapter 3

Associations of elevated antithyroperoxidase antibodies with thyroid function, survival, functioning, and depressive symptoms in the oldest old: the Leiden 85-plus study

Robert S Du Puy
Rosalinde KE Poortvliet
Marieke Snel
Wendy PJ den Elzen
Bart EPB Ballieux
Olaf M Dekkers
Simon P Mooijaart
Jacobijn Gussekloo

Thyroid, 2019;29(9):1201-1208

DOI: 10.1089/thy.2019.0129

ABSTRACT

Background

Elevated levels of antithyroperoxidase antibodies (TPOAbs) have been associated with progression of subclinical thyroid dysfunction, extrathyroidal diseases, and decrease in functional status. However, TPOAb as determinant of future thyroid dysfunction and other clinical outcomes has not been studied well for adults aged 85 years and over. This study aimed to assess associations of TPOAb levels with thyroid function, survival, physical function, disability in activities of daily living (ADL), cognitive function, and depressive symptoms in the oldest old.

Methods

Data from a population-based cohort study (Leiden 85-plus Study) of residents of Leiden, the Netherlands, aged 85 and older were used. Baseline serum TPOAb levels were available for 488 participants (82% of the total cohort). We considered levels ≥ 35 IU/mL as elevated. Thyroid function (thyrotropin [TSH] and free thyroxine) was assessed at age 85 (baseline), 87, and 88 years. Survival, physical function, disability in ADL, cognitive function, and depressive symptoms were assessed from age 85 through 90 years.

Results

At baseline, 64 of the 85-year-old participants (13.1%) had elevated TPOAb levels. They were more often female, had higher TSH levels, and a higher prevalence of overt or subclinical hypothyroidism than participants with normal TPOAb levels. Over time, elevated TPOAb levels were independently associated with a lower mortality risk (hazard ratio 0.72, [95% confidence interval 0.53–0.99]), but were not associated with changes in thyroid function, nor with physical function, disability in ADL, cognitive function, or depressive symptoms.

Conclusions

In community-dwelling oldest old, elevated TPOAb levels are cross-sectionally associated with higher TSH levels. Over time, elevated TPOAb levels are associated with a survival benefit but are not associated with changes in thyroid function, functional status, or depressive symptoms in old age. The added clinical value of TPOAb tests in oldest old persons with thyroid dysfunction is limited.

INTRODUCTION

In recent decades, the interest in autoimmune antibodies as predictive biomarkers has risen.[1,2] Since the discovery of antithyroperoxidase antibodies (TPOAbs), its usefulness as a determinant of health status in thyroid dysfunction has been investigated, but results have been ambiguous. Guidelines are starting to cautiously recommend assessing TPOAb status to assess risk of thyroid disease progression, choose follow-up modalities, and weigh treatment options.[3]

The relationship between elevated TPOAbs and the progression of subclinical hypothyroidism to overt hypothyroidism is well established,[4-7] but estimates of progression range between 4.3% and 80% for TPOAb-positive patients.[6,7] In addition, elevated TPOAb levels have been found in the presence of several other debilitating extrathyroidal diseases.[8-11] It is currently unclear whether elevated TPOAbs are causally interfering with extrathyroidal disease processes, or whether they are merely a marker of a common pathway associated with autoimmune diseases. However, the co-occurrence of elevated TPOAbs with comorbidities may suggest associations between thyroid antibody status and clinical outcomes that could be of interest to clinicians and patients. There are ample data to suggest that thyroid (dys-)function is associated with age.[12] Accordingly, potential effects of elevated TPOAb levels may be profoundly different in oldest old persons. Although associations have been found in younger age groups, prediction of thyroid dysfunction progression using TPOAbs has not been studied well for adults aged 85 years and over, nor have associations with other clinical outcomes.

Therefore, in this study, we aimed to investigate the associations of TPOAb levels with thyroid function, survival, physical function, disability in activities of daily living (ADL), cognitive function, and depressive symptoms in a community-dwelling prospective cohort of oldest old persons.

MATERIALS AND METHODS

Study population

This study was performed using data from the Leiden 85-plus Study, a population-based cohort study of 85-year-old residents of Leiden, the Netherlands. Eligible participants were all inhabitants of Leiden who reached the age of 85 years between September 1997 and September 1999. Participants were visited in residential settings annually for interviews with trained research nurses, performing of functional tests, and the collection of nonfasting venous blood samples during a five-year follow-up period. Full details of the study protocol

and procedures are described elsewhere.[13] The Medical Ethical Committee of the Leiden University Medical Center approved the study. Informed consent was obtained from all participants. For severely cognitively impaired participants, informed consent was obtained by proxy.

Laboratory measurements of thyroid function

To reduce diurnal variation, nonfasting venous blood samples were drawn at each annual visit (85–90 years) between 8 and 11 a.m. Biobank samples were stored at –80°C.

TPOAb levels were measured once (age 86 years) and were assessed from biobank samples in a single batch by a solid-phase enzyme-labeled chemiluminescent sequential immunometric assay, using an Immulite 2000 XPi Immunoassay System (Siemens AG, Berlin, Germany; WHO 1st IRP 66/387 standard compliant). The lower and upper limits of detection were 5 and 1,000 IU/mL, respectively (within-run coefficient of variation 5.5%, total 6.9%). TPOAb levels below the detection limit of 5 IU/mL were set at 5 IU/mL for the aim of the present analyses.

TPOAb levels ≥ 35 IU/mL were considered elevated in compliance with local laboratory reference ranges. Serum TPOAb levels were available for 488 participants (of a total 599 participants in the total cohort) at the age of 86 years, that is, 1 year after start of the study. Earlier research demonstrated little to no change in TPOAb levels over time, even when underlying thyroid disease was treated.[14] Therefore, TPOAb levels at age 86 years were assumed to be constant over time and representative for the levels at age 85 years. A more detailed description of the included participants is included as a flowchart (Supplementary Fig. S1).

Plasma levels of thyrotropin (TSH) and free thyroxine (fT4) were measured 3 times for a 4-year period (85, 87, and 88) and were analysed in batches using a fully automated Elecsys 2010 system (Hitachi, Tokyo, Japan). Reference ranges used for TSH were 0.3 to 4.8 mIU/L and 13.0 pmol/L (1.00 ng/dL) to 23.0 pmol/L (1.79 ng/dL) for fT4 in accordance with manufacturer specifications and local laboratory cut-offs. TSH and fT4 measurements were available at age 85 (baseline), 87, and 88 years. Participants were grouped into 5 clinical strata of thyroid function based on TSH and fT4 levels: overt hypothyroidism (TSH > 4.8 and fT4 < 13), subclinical hypothyroidism (TSH > 4.8 and 13 ≤ fT4 ≤ 23), euthyroidism (0.3 ≤ TSH ≤ 4.8), subclinical hyperthyroidism (TSH < 0.3 and 13 ≤ fT4 ≤ 23), and overt hyperthyroidism (TSH < 0.3 and fT4 > 23) at age 85, 87, and 88 years.

Mortality

Mortality information was available for a 10-year follow-up period and obtained from records of Statistics Netherlands (Centraal Bureau voor de Statistiek, The Hague) up to September 2008 (available for all participants).

Functional parameters

Outcomes of physical function, disability in ADL, cognitive function, or depressive symptoms were available for a 5-year follow-up period (age 85 to 90 years) unless otherwise specified.

Physical function

Hand grip strength measurements in kilograms using a Jamar hand dynamometer (Sammons Preston, Inc., IL) were used as a proxy for overall muscle strength.[15] Hand grip strength measurements were performed at age 85 and 89 years. At ages 86 through 90 years, gait speed was measured over a dual 3 m walk with a turn halfway and expressed in meters per second.[16]

Activities of daily living

Disability in ADL was assessed using the self-reported Groningen Activity Restriction Scale (GARS).[17] This questionnaire consisted of 18 items of daily functioning and independent living such as "getting in and out of bed" and "ability to fully dress yourself." Each item could be awarded a maximum of four points, with one point indicating "able to do so independently" and four points indicating "requiring assistance." A higher summed score indicates more disability in ADL.

Cognitive function

Mini Mental State Examination (MMSE) questionnaires [18] were taken for assessing overall cognitive functioning as well as Stroop-card examinations (concentration) [19] and Letter Digit Coding Tests (LDCTs, cognitive processing speed) [20]. The MMSE questionnaire contains 19 items adding up to a total score of 30. Higher scores indicate better cognitive functioning. For the Stroop-card examination, the time in seconds it took to complete the third card with 40 words of colours printed in a different colour was used. The Stroop test demonstrates the interference effect in the reaction time for a colour-word task. Shorter times indicate better concentration. For the LDCT, participants completed digits corresponding to a random set of pre-printed letters according to a key. The number of correct substitutions in 60 seconds was used for analysis.

Depressive symptoms

A presence of depressive symptoms was evaluated with the Geriatric Depression Scale (GDS-15) [21] consisting of 15 individual items in all subjects with a MMSE score ≥ 18 points. Higher scores indicate the presence of more depressive symptoms.

Demographics and medication

Information on highest completed level of education and independent living status was gathered during the interview. Use of any or a combination of thyroid influencing medica-

tion (levothyroxine, antithyroid medication, amiodarone, lithium, glucocorticosteroids, and/or interferon) was ascertained annually through registries of pharmacy records.

Statistical analyses

Continuous variables were described as mean with standard deviation or as median with interquartile range, where appropriate, and categorical variables were presented as frequency with percentage of the total. Participants were stratified into a normal (negative) TPOAb group (< 35 IU/L) or elevated (positive) TPOAb group (≥ 35 IU/L) at baseline. Between-group comparisons of baseline characteristics were performed with Mann–Whitney U tests for continuous variables or χ^2 tests for categorical data.

Association of TPOAbs with all-cause 10-year cumulative survival was estimated with the Kaplan–Meier method in all participants for which a TPOAb measurement was available, with follow-up starting at age 86 years. Survival analysis was performed using multivariable Cox proportional hazards regression model, adjusted for sex and fT4 levels, and excluding participants with thyroid influencing medication.

Associations of elevated TPOAb levels with TSH and fT4 levels over time were assessed with repeated measurements linear mixed effect models with an unstructured covariance structure, adjusted for sex, and reported as an estimated mean difference over time for both TSH and fT4. Slope divergence was assessed through the inclusion of a TPOAb × time-interaction covariate.

Incident thyroid dysfunction was defined as newly developed thyroid dysfunction between age 85 and 88 years in euthyroid participants at age 85 years. Changes in TSH and fT4 over 3 years' time (ΔTSH and ΔfT4) were calculated by subtracting the baseline thyroid function levels from levels at age 88 years and modelling these changes in linear mixed effect models adjusting for sex and using an unstructured covariance structure. Thyroid dysfunction at age 85 years was cross-tabulated, and $\chi 2$ tested, with thyroid dysfunction at age 87 and 88 years for all clinical thyroid strata, provided the participant was still alive at age 88 years, to determine whether TPOAb levels are associated with change from one clinical thyroid stratum to another.

Longitudinal associations between TPOAbs and physical function, disability in ADL, cognitive function, and depressive symptoms were assessed with repeated measurements linear mixed models with an unstructured covariance structure, adjusted for sex, and reported as estimated mean change and difference in both TPOAb groups in function over five years of follow-up.

For all analyses, SPSS Statistics Software version 22.0 for Windows (IBM, Armond, NY) was used. A p-value < 0.05 was considered statistically significant.

Sensitivity analysis

All analyses were repeated excluding all participants (n = 26) using any form of thyroid influencing medication at baseline (levothyroxine, antithyroid medication, amiodarone, lithium, glucocorticosteroids, or interferon).

RESULTS

Study population characteristics at baseline

Table 1 shows the baseline characteristics for all study participants in whom TPOAb was measured. In total, 64 (13.1%) of 488 study participants aged 85 years had elevated TPOAb levels. Participants in the elevated TPOAb group more often were female compared with the normal TPOAb group (81.3% vs. 65.8%, p = 0.008).

At baseline, elevated TPOAb was associated with an increased percentage of overt hypothyroidism (14.5% vs. 6.4%, p = 0.012) and subclinical hypothyroidism (16.1% vs. 3.1%, p < 0.001) compared with normal levels of TPOAbs. Median TSH levels were higher in participants with elevated TPOAbs (2.5 mIU/L vs. 1.8 mIU/L, p < 0.001). Level of education and independent living status were comparable between the groups. No clinically meaningful differences were observed for assessments of depressive symptoms and physical or cognitive functioning.

TPOAbs and survival

A Kaplan–Meier curve shows a small 10-year cumulative survival benefit in participants with elevated TPOAb levels compared with normal TPOAb levels (p = 0.032, Log-Rank test, Fig. 1).

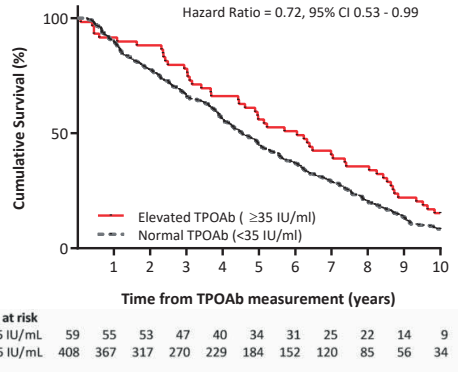


Figure 1. Association of baseline TPOAbs with all-cause 10-year cumulative survival from 86 years of age. Survival was estimated with the Kaplan–Meier method in all participants for which a TPOAb measurement was available, with follow-up starting at age 86 years. Hazard ratio was assessed using multivariable Cox proportional hazards regression models adjusted for sex and excluding participants with thyroid influencing medication. Abbreviation: TPOAbs, thyreoperoxidase

Table 1. Study population characteristics at 85 years of age.

Characteristic	TPOAb levels		p-value
	Normal (< 35IU/L)	Elevated (≥ 35IU/L)	
N (%)	424 (86.9)	64 (13.1)	
Demographics			
Female (%)	279 (65.8)	52 (81.3)	0.008
Secondary education (%)	157 (37.0)	22 (34.4)	0.386
Living independently (%)	351 (82.8)	55 (84.4)	0.457
Using thyroid influencing medication (%)[a]	20 (4.7)	6 (9.4)	0.112
Clinical thyroid function strata			
Overt Hypothyroidism (%)	27 (6.4)	9 (14.5)	0.012
Subclinical Hypothyroidism (%)	13 (3.1)	10 (16.1)	< 0.001
Euthyroidism (%)	366 (87.1)	41 (66.1)	< 0.001
Subclinical Hyperthyroidism (%)	14 (3.3)	1 (1.6)	0.549
Overt Hyperthyroidism (%)	0 (0)	1 (1.6)	0.103
Thyroid function parameters			
TSH (mIU/L)	1.8 (1.2, 2.7)	2.5 (1.6, 5.2)	< 0.001
fT4 (pmol/L)	14.3 (12.8, 15.8)	14.3 (12.4, 15.4)	0.574
Functional parameters			
Grip Strength (kg)	23 (9)	22 (8)	0.305
Gait speed (m/s)	0.4 (0.3, 0.6)	0.4 (0.3, 0.5)	0.834
GARS	28 (21, 39)	27 (21, 36)	0.728
MMSE	26 (22, 28)	27 (22, 29)	0.236
LDCT (subs/m)	17 (7)	19 (8)	0.073
Stroop (s)	73 (60, 98)	71 (56, 90)	0.362
GDS	2 (1, 3)	1 (0, 3)	0.420

The Stroop test demonstrates the interference effect in the reaction time for a color-word task. Data are presented as counts (%) for demographics and clinical thyroid function strata, as median (IQR) for thyroid function parameters and functional status with exception of LDTC and grip strength, which are mean (SD).

[a] Use of any or a combination of levothyroxine, antithyroid medication, amiodarone, lithium, glucocorticosteroids or interferons.

Abbreviations: TSH, Thyroid Stimulating Hormone; fT4, free Thyroxine; GARS, Groningen Activity Restriction Scale; MMSE, Mini Mental State Examination; LDCT, Letter-Digit Coding Test; GDS, Geriatric Depression Scale.

In a multivariable Cox-regression model, adjusted for sex and fT4, and excluding participants with thyroid influencing medication, elevated TPOAb levels were associated with improved survival (hazard ratio 0.72, [95% confidence interval, CI, 0.53–0.99], p = 0.044).

TPOAbs and changes in thyroid function

Median fT4 levels were comparable at ages 85, 87, and 88 years between participants with elevated or normal TPOAb levels (estimated mean difference between participants with

elevated TPOAbs and normal TPOAbs −0.15 mIU/L, [95% CI −0.96 to 0.66], p = 0.715) (Fig. 2). Participants with elevated TPOAb levels had higher median TSH levels at baseline than participants with normal TPOAb levels (age 85 years, respectively, 2.46 mIU/L vs. 1.78 mIU/L, p < 0.001) and this difference remained over time in linear mixed regression models (estimated mean difference 1.45 mIU/L, [95% CI 0.862.04], p < 0.001). Differences in TSH levels between the groups, however, did not increase or decrease with time (estimated mean difference −0.08 mIU/L, [95% CI −0.65 to 0.48], p-value for time interaction = 0.645).

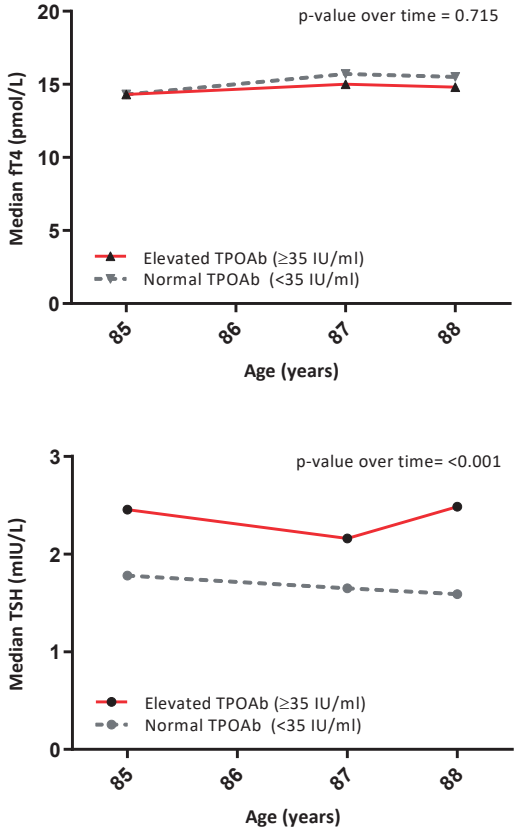


Figure 2. Association of baseline TPOAbs with TSH and fT4 during three-year follow-up. Associations over time were estimated using repeated measures linear mixed models, correcting for sex. Abbreviations: fT4, free thyroxine; TSH, thyrotropin.

TPOAbs and incident thyroid disease

Incident cases of thyroid disease during follow-up in all euthyroid participants at baseline, who were still alive at age 88 years, are presented in Table 2. Of the surviving participants who were euthyroid at baseline, 97.1% remained euthyroid after 3 years of follow-up regardless of TPOAb status. Incidence of thyroid disease in the entire cohort was low with six new cases

in the elevated TPOAb group versus one new case in the elevated TPOAb group. Participants with elevated TPOAb levels showed a very slight annual increase in TSH levels compared with participants with normal TPOAb levels (0.1 mIU/L vs. −0.1 mIU/L, p = 0.013). Participants with thyroid dysfunction at baseline either remained in their respective clinical thyroid stratum or regressed to a more euthyroid state (i.e., overt hypothyroidism regressed to subclinical hypothyroidism) after three years of follow-up in equal proportions in the normal and elevated TPOAb groups (data not shown).

Table 2. Association of TPOAb in euthyroid participants at baseline with thyroid function during 3-year follow-up.

	TPOAb levels		p-value
	Normal (< 35IU/L)	Elevated (≥ 35IU/L)	
Alive at age 88 (%)	206 (89.2)	26 (10.8)	
Clinical thyroid function strata at 88			
Incident overt Hypothyroidism (%)	2 (1.0)	0 (0.0)	0.788
Incident Subclinical Hypothyroidism (%)	3 (1.5)	1 (3.8)	0.380
Persistent Euthyroidism (%)	200 (97.1)	25 (96.2)	0.570
Incident Subclinical Hyperthyroidism (%)	1 (0.5)	0 (0.0)	0.888
Incident overt Hyperthyroidism (%)	0 (0.4)	0 (0.0)	-
Thyroid function parameters			
Δ TSH (mIU/L)	-0.1 (-0.5, 0.2)	0.1 (-1.4, 0.7)	0.013
Δ fT4 (pmol/L)	1.2 (-0.1, 2.5)	0.6 (-0.2, 2.1)	0.239

Data are presented as counts (%) and median (IQR). Incident thyroid dysfunction was assessed by reclassifying all euthyroid participants at baseline per TSH and fT4 status after 3 years and assessed using X^2 tests. Δ TSH and fT4 was calculated as mean change in thyroid function after 3 years using linear mixed effect models adjusting for sex. **Abbreviations:** TSH, Thyroid Stimulating Hormone; fT4, free Thyroxine.

TPOAbs and functioning

Table 3 shows associations between TPOAbs and functional status over a five-year follow-up period. No clear differences were observed in physical functioning (grip strength or gait speed), activity restriction scores (GARS), cognitive function (MMSE, LDCT, and Stroop card examination) and depressive symptoms scores (GDS) over time.

Sensitivity analysis

Excluding all participants (n = 26) using any form of thyroid influencing medication at baseline did not change the results considerably (data not shown).

Table 3. Longitudinal associations of TPOAb levels and physical function, disability in activities of daily living, cognitive function and depressive symptoms over 5 years of follow-up.

	TPOAb levels		Difference (95% CI)	p-value
	Normal (< 35IU/L)	Elevated (≥ 35IU/L)		
	Mean change (95% CI)	Mean change (95% CI)		
Grip strength (kg)	-4.37 (-4.95, -3.79)	-3.49 (-4.92, -2.06)	0.88 (-0.67, 2.43)	0.264
Gait speed (m/s)	-0.00 (-0.01, 0.01)	0.02 (-0.00, 0.04)	0.02 (-0.00, 0.04)	0.103
GARS	3.87 (3.58, 4.17)	3.18 (2.45, 3.90)	-0.70 (-1.48, 0.09)	0.081
MMSE	-0.89 (-1.00, -0.75)	-0.83 (-1.13, -0.52)	0.05 (-0.28, 0.38)	0.762
LDCT (subs./m)	-0.82 (-1.22, -0.42)	-0.76 (-1.76, 0.23)	0.05 (-1.02, 1.12)	0.912
Stroop (s)	2.27 (1.50, 3.04)	1.62 (-0.29, 3.53)	-0.64 (-2.70, 1.41)	0.537
GDS	0.34 (0.26, 0.43)	0.30 (0.09, 0.50)	-0.05 (-0.27, 0.17)	0.661

The Stroop test demonstrates the interference effect in the reaction time for a colour-word task. Data are presented as estimated mean change in physical or cognitive function calculated over 5 years of follow up per TPOAb group. The difference is estimated using repeated measures linear mixed models with normal TPOAb levels as reference group, adjusted for sex and baseline scores, with corresponding confidence intervals. Higher scores in grip strength, gait speed, MMSE scores and LDCT scores, and lower scores in GARS scores, Stroop examination scores and GDS scores, indicate better function. **Abbreviations:** GARS, Groningen Activity Restriction Scale; MMSE, Mini Mental State Examination; LDCT, Letter-Digit Coding Test; GDS, Geriatric Depression Scale.

DISCUSSION

The main findings of this observational prospective cohort study in oldest old community dwelling participants were first that elevated TPOAb levels were associated with higher TSH both at baseline and during follow-up. Second, elevated TPOAbs were associated with a decreased 10-year mortality risk. Third, elevated TPOAbs were not associated with an increased risk of incident thyroid disease. Fourth, elevated TPOAbs were not associated with changes in physical function, disability in ADL, cognitive function, or depressive symptoms.

The baseline associations between elevated TPOAb levels and female sex, elevated TSH levels, and prevalent overt and subclinical hypothyroidism described in this study are in line with earlier studies. Elevated TPOAb levels can be found in up to 10% of euthyroid older persons [22] and are more frequently found in Hashimoto's thyroiditis (90%), Graves' disease (80%), and nonautoimmune thyroid disease.[23] Such conditions are more prevalent in female patients, can give rise to elevated TSH levels, and may lead to subclinical and overt thyroid disease through a myriad of disease mechanisms.

In this study, elevated TPOAb levels were associated with a 10-year cumulative survival benefit. Moreover, this association was independent of sex or fT4 levels, suggesting that any

effects on survival are probably not mediated through an increased or decreased thyroid function. Earlier research investigating associations between TPOAbs and survival yielded contradicting results, with some suggesting that mild thyroid dysfunction may be associated with favourable health status and longevity in old age,[24-26] while subsequent large-scale individual participant data meta-analysis failed to demonstrate similar effects.[27] The increased overall all-cause survival associated with elevated TPOAbs is an interesting epidemiological finding, particularly in light of earlier identified prolonged survival in participants aged ≥ 85 years with thyroid dysfunction; however, etiologic research is needed to further explain any potential beneficial effects from a physiologic standpoint.[25]

Thyroid dysfunction usually progresses from euthyroid and subclinical thyroid disease to overt thyroid pathology.[28] Although number of incidences were low, in this study, no differences in disease progression were found when participants were stratified by TPOAb levels, comparable with results of recent studies in older age groups.[29,30] Although the TPOAb-positive participants demonstrated increased survival time, and both groups demonstrated a decline in functional status not uncommon of this age group, associations between TPOAbs and physical function, disability in ADL, cognitive function, or the presence of depressive symptoms during follow-up were not identified. Earlier research has provided ambiguous findings for most of the aforementioned associations. However, when examined more carefully, in general these earlier results are in line with the findings of this study. For instance, it has been suggested that the presence of elevated TPOAbs is related with an increased risk of depressive symptoms and depression, regardless of thyroid status.[31-34] However, these studies were performed in small, younger, or highly selected populations such as pregnant or perimenopausal women. When investigated in large scale or slightly older populations, such associations were no longer clearly demonstrable,[35,36] and in older patients these associations were clearly absent.[37] Any potential extrathyroidal effects of TPOAbs that contribute to functional decline may be subtle and clinically of little importance in old age. Our study corroborates the opinion that the relevance of TPOAb measurement for the prediction of progression of thyroid disease is possibly lower in the older age group than in middle or younger age groups.[38,39]

The findings of our study suggest a limited role for adding TPOAb measurements to the standard thyroid function tests (TSH and fT4) in the oldest old, which have been well established for estimating and monitoring the severity, pathophysiology, and progression of thyroid (dys-)function,[40] even in the oldest old. However, future research with repeated TPOAbs, thyroid function, and physical and cognitive function measurements over a longer period of time and in cohorts of older persons is needed for proper causal inferences and to establish whether elevated TPOAbs causally influence clinical outcomes.

Strengths and weaknesses

The prospective follow-up design of the study with a full 5-year follow-up of clinical data and 10-year follow-up for mortality data is a strong point of the study. The unselected, population-based, and sizable sample is an accurate representation of the older community-dwelling population in the Leiden district and the Netherlands as a whole (Supplementary Fig. S1). The highly standardised study protocol with validated assessments and a repeated measurements design ensured high precision and reliability of the collected data. Moreover, iodine intake is considered sufficient in the Netherlands.[41]

TPOAb levels were assessed only once after 15 years of frozen storage at −80°C. It is currently unknown how and to what extent TPOAb stability is affected during prolonged freezing periods. In an earlier study, no statistically significant changes were found in TPOAb samples stored at +4°C for 6 days or in samples exposed to 50 freezing and thawing cycles.[42] Hence, we assume little to no changes in TPOAb levels during storage. Even if degradation has occurred, we expect the degradation to be similar across the groups. Unfortunately, TPOAb measurements were only available at age 86 years, while other thyroid function measures were available from age 85 years onward. We cannot exclude the possibility that TPOAb levels at age 85 years or from age 87 years onward were different or changed over time. However, the demonstrated persistence of having elevated or normal TPOAb levels over time, even when underlying thyroid disease is being treated,[14] reinforces our premise that TPOAb levels remained constant at all visits in all but a negligible number of participants.

Until now, earlier research for TPOAbs in older age groups has mainly focused on estimating TPOAb prevalence and incidence estimates.[4-7] To the authors' knowledge, this is one of the first studies to cross-sectionally and longitudinally assess associations between TPOAbs and clinically relevant health outcomes in an unselected cohort of community-dwelling oldest old.

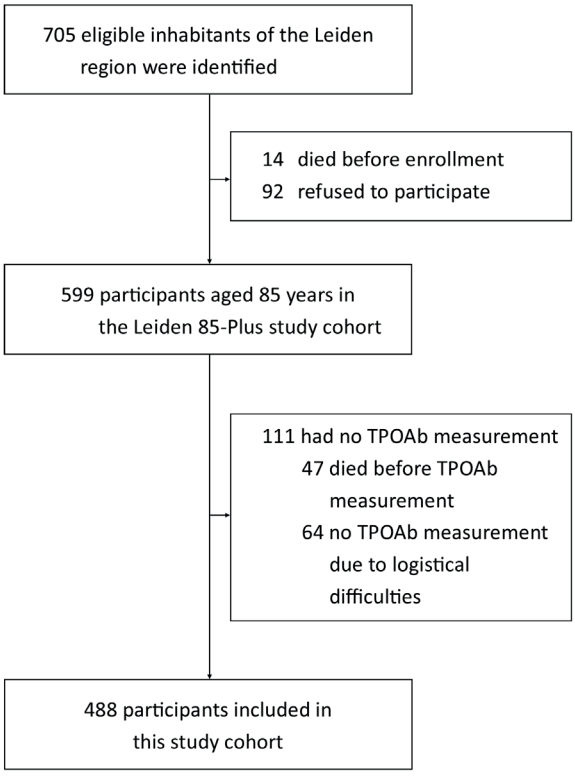
In conclusion, our study confirms the cross-sectional associations of elevated TPOAb levels with high TSH levels and a higher prevalence of subclinical and overt hypothyroidism in the oldest old. Over time, elevated TPOAb levels predict a survival benefit for a 10-year follow-up but are not associated with thyroid dysfunction progression, a decline in presence of depressive symptoms, or a decline in physical or cognitive functioning.

REFERENCES

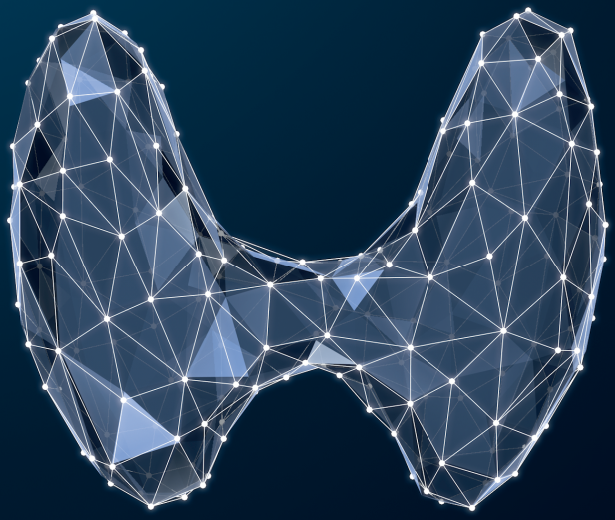
1. Rose NR. Prediction and prevention of autoimmune disease: a personal perspective. *Annals of the New York Academy of Sciences.* 2007;1109:117-128.
2. Scofield RH. Autoantibodies as predictors of disease. *Lancet (London, England).* 2004;363(9420):1544-1546.
3. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2012;18(6):988-1028.
4. Diez JJ, Iglesias P. Spontaneous subclinical hypothyroidism in patients older than 55 years: an analysis of natural course and risk factors for the development of overt thyroid failure. *J Clin Endocrinol Metab.* 2004;89(10):4890-4897.
5. Huber G, Staub JJ, Meier C, et al. Prospective study of the spontaneous course of subclinical hypothyroidism: prognostic value of thyrotropin, thyroid reserve, and thyroid antibodies. *J Clin Endocrinol Metab.* 2002;87(7):3221-3226.
6. Rosenthal MJ, Hunt WC, Garry PJ, Goodwin JS. Thyroid failure in the elderly. Microsomal antibodies as discriminant for therapy. *JAMA.* 1987;258(2):209-213.
7. Vanderpump MP, Tunbridge WM, French JM, et al. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. *Clin Endocrinol (Oxf).* 1995;43(1):55-68.
8. Benvenga S, Pintaudi B, Vita R, Di Vieste G, Di Benedetto A. Serum thyroid hormone autoantibodies in type 1 diabetes mellitus. *J Clin Endocrinol Metab.* 2015;100(5):1870-1878.
9. Cardenas Roldan J, Amaya-Amaya J, Castellanos-de la Hoz J, et al. Autoimmune thyroid disease in rheumatoid arthritis: a global perspective. *Arthritis.* 2012;2012:864907.
10. Chan JC, Liu HS, Kho BC, et al. Pattern of thyroid autoimmunity in chinese patients with pernicious anemia. *Am J Med Sci.* 2009;337(6):432-437.
11. Nakamura H, Usa T, Motomura M, et al. Prevalence of interrelated autoantibodies in thyroid diseases and autoimmune disorders. *J Endocrinol Invest.* 2008;31(10):861-865.
12. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. *Endocrine reviews.* 2008;29(1):76-131.
13. Weverling-Rijnsburger AW, Blauw GJ, Lagaay AM, Knook DL, Meinders AE, Westendorp RG. Total cholesterol and risk of mortality in the oldest old. *Lancet (London, England).* 1997;350(9085):1119-1123.
14. Schmidt M, Voell M, Rahlff I, et al. Long-term follow-up of antithyroid peroxidase antibodies in patients with chronic autoimmune thyroiditis (Hashimoto's thyroiditis) treated with levothyroxine. *Thyroid.* 2008;18(7):755-760.
15. Bohannon RW. Dynamometer measurements of hand-grip strength predict multiple outcomes. *Percept Mot Skills.* 2001;93(2):323-328.
16. Bloem BR, Haan J, Lagaay AM, van Beek W, Wintzen AR, Roos RA. Investigation of gait in elderly subjects over 88 years of age. *Journal of geriatric psychiatry and neurology.* 1992;5(2):78-84.
17. Bootsma-van der Wiel A, Gussekloo J, de Craen AJ, et al. Disability in the oldest old: "can do" or "do do"? *J Am Geriatr Soc.* 2001;49(7):909-914.
18. Heeren TJ, Lagaay AM, von Beek WC, Rooymans HG, Hijmans W. Reference values for the Mini-Mental State Examination (MMSE) in octo- and nonagenarians. *J Am Geriatr Soc.* 1990;38(10):1093-1096.

19. Klein M, Ponds RW, Houx PJ, Jolles J. Effect of test duration on age-related differences in Stroop interference. *J Clin Exp Neuropsychol.* 1997;19(1):77-82.
20. Houx PJ, Shepherd J, Blauw GJ, et al. Testing cognitive function in elderly populations: the PROSPER study. PROspective Study of Pravastatin in the Elderly at Risk. *J Neurol Neurosurg Psychiatry.* 2002;73(4):385-389.
21. de Craen AJ, Heeren TJ, Gussekloo J. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. *Int J Geriatr Psychiatry.* 2003;18(1):63-66.
22. Roti E, Gardini E, Minelli R, Bianconi L, Braverman LE. Prevalence of anti-thyroid peroxidase antibodies in serum in the elderly: comparison with other tests for anti-thyroid antibodies. *Clinical chemistry.* 1992;38(1):88-92.
23. Carvalho GA, Perez CL, Ward LS. The clinical use of thyroid function tests. *Arquivos brasileiros de endocrinologia e metabologia.* 2013;57(3):193-204.
24. Atzmon G, Barzilai N, Hollowell JG, Surks MI, Gabriely I. Extreme longevity is associated with increased serum thyrotropin. *J Clin Endocrinol Metab.* 2009;94(4):1251-1254.
25. Gussekloo J, van Exel E, de Craen AJ, Meinders AE, Frolich M, Westendorp RG. Thyroid status, disability and cognitive function, and survival in old age. *JAMA.* 2004;292(21):2591-2599.
26. Jansen SW, Akintola AA, Roelfsema F, et al. Human longevity is characterised by high thyroid stimulating hormone secretion without altered energy metabolism. *Sci Rep.* 2015;5:11525.
27. Collet TH, Bauer DC, Cappola AR, et al. Thyroid antibody status, subclinical hypothyroidism, and the risk of coronary heart disease: an individual participant data analysis. *J Clin Endocrinol Metab.* 2014;99(9):3353-3362.
28. Kabadi UM. 'Subclinical hypothyroidism'. Natural course of the syndrome during a prolonged follow-up study. *Arch Intern Med.* 1993;153(8):957-961.
29. Imaizumi M, Sera N, Ueki I, et al. Risk for progression to overt hypothyroidism in an elderly Japanese population with subclinical hypothyroidism. *Thyroid.* 2011;21(11):1177-1182.
30. Somwaru LL, Rariy CM, Arnold AM, Cappola AR. The natural history of subclinical hypothyroidism in the elderly: the cardiovascular health study. *J Clin Endocrinol Metab.* 2012;97(6):1962-1969.
31. Haggerty JJ, Jr., Silva SG, Marquardt M, et al. Prevalence of antithyroid antibodies in mood disorders. *Depression and anxiety.* 1997;5(2):91-96.
32. Harris B, Othman S, Davies JA, et al. Association between postpartum thyroid dysfunction and thyroid antibodies and depression. *BMJ.* 1992;305(6846):152-156.
33. Kuijpens JL, Vader HL, Drexhage HA, Wiersinga WM, van Son MJ, Pop VJ. Thyroid peroxidase antibodies during gestation are a marker for subsequent depression postpartum. *European journal of endocrinology / European Federation of Endocrine Societies.* 2001;145(5):579-584.
34. Pop VJ, Maartens LH, Leusink G, et al. Are autoimmune thyroid dysfunction and depression related? *J Clin Endocrinol Metab.* 1998;83(9):3194-3197.
35. Engum A, Bjoro T, Mykletun A, Dahl AA. Thyroid autoimmunity, depression and anxiety; are there any connections? An epidemiological study of a large population. *Journal of psychosomatic research.* 2005;59(5):263-268.
36. van de Ven AC, Muntjewerff JW, Netea-Maier RT, et al. Association between thyroid function, thyroid autoimmunity, and state and trait factors of depression. *Acta Psychiatr Scand.* 2012;126(5):377-384.
37. Medici M, Direk N, Visser WE, et al. Thyroid function within the normal range and the risk of depression: a population-based cohort study. *J Clin Endocrinol Metab.* 2014;99(4):1213-1219.
38. Lazarus JH, Burr ML, McGregor AM, et al. The prevalence and progression of autoimmune thyroid disease in the elderly. *Acta Endocrinol (Copenh).* 1984;106(2):199-202.

39. Sawin CT, Thomas Bigos S, Land S, Bacharach P. The aging thyroid. Relationship between elevated serum thyrotropin level and thyroid antibodies in elderly patients. *Am J Med.* 1985;79(5):591-595.
40. Sheehan MT. Biochemical Testing of the Thyroid: TSH is the Best and, Oftentimes, Only Test Needed - A Review for Primary Care. *Clinical medicine & research.* 2016;14(2):83-92.
41. World Health Organisation. WHO Global Database on Iodine Deficiency. 2007; http://who.int/vmnis/iodine/data/database/countries/nld_idd.pdf?ua=1. Accessed Sept 2018.
42. Mannisto T, Surcel HM, Bloigu A, et al. The effect of freezing, thawing, and short- and long-term storage on serum thyrotropin, thyroid hormones, and thyroid autoantibodies: implications for analyzing samples stored in serum banks. *Clinical chemistry.* 2007;53(11):1986-1987.



Supplementary Figure S1. Flowchart of included participants



Chapter 4

The relation between thyroid function and anemia: a pooled analysis of individual participant data

Daisy M Wopereis
Robert S Du Puy
Diana van Heemst
John P Walsh
Alexandra Bremner
Stephan J L Bakker
Douglas C Bauer
Anne R Cappola
Graziano Ceresini
Jean Degryse
Robin PF Dullaart
Martin Feller
Luigi Ferrucci
Carmen Floriani
Oscar H Franco
Massimo Iacoviello
Georgio Iervasi
Misa Imaizumi
J Wouter Jukema
Kay-Tee Khaw
Robert N Luben
Sabrina Molinaro
Matthias Nauck
Kushang V Patel
Robin P Peeters
Bruce M Psaty
Salman Razvi
Roger K Schindhelm
Natasja M van Schoor
David J Stott
Bert Vaes
Mark PJ Vanderpump
Henry Völzke
Rudi GJ Westendorp
Nicolas Rodondi
Christa M Cobbaert
Jacobijn Gussekloo
Wendy PJ den Elzen for the

Thyroid Studies Collaboration

The Journal of Clinical Endocrinology & Metabolism, 2018;103(10):3658-3667

DOI: 10.1210/jc.2018-00481

ABSTRACT

Context

Anaemia and thyroid dysfunction often co-occur, and both increase with age. Human data on relationships between thyroid disease and anaemia are scarce.

Objective

To investigate the cross-sectional and longitudinal associations between clinical thyroid status and anaemia.

Design

Individual participant data meta-analysis.

Setting

Sixteen cohorts participating in the Thyroid Studies Collaboration (N = 42,162).

Main Outcome Measures

Primary outcome measure was anaemia (haemoglobin < 130 g/L in men and < 120 g/L in women).

Results

Cross-sectionally, participants with abnormal thyroid status had an increased risk of having anaemia compared with euthyroid participants [overt hypothyroidism, pooled OR 1.84 (95% CI 1.35 to 2.50), subclinical hypothyroidism 1.21 (1.02 to 1.43), subclinical hyperthyroidism 1.27 (1.03 to 1.57), and overt hyperthyroidism 1.69 (1.00 to 2.87)]. Haemoglobin levels were lower in all groups compared with participants with euthyroidism. In the longitudinal analyses (N = 25,466 from 14 cohorts), the pooled hazard ratio for the risk of development of anaemia was 1.38 (95% CI 0.86 to 2.20) for overt hypothyroidism, 1.18 (1.00 to 1.38) for subclinical hypothyroidism, 1.15 (0.94 to 1.42) for subclinical hyperthyroidism, and 1.47 (0.91 to 2.38) for overt hyperthyroidism. Sensitivity analyses excluding thyroid medication or high levels of C-reactive protein yielded similar results. No differences in mean annual change in haemoglobin levels were observed between the thyroid hormone status groups.

Conclusion

Higher odds of having anaemia were observed in participants with both hypothyroid function and hyperthyroid function. In addition, reduced thyroid function at baseline showed a trend of increased risk of developing anaemia during follow-up. It remains to be assessed in a randomised controlled trial whether treatment is effective in reducing anaemia.

INTRODUCTION

Thyroid diseases and anaemia are common disorders, and their prevalence increases with age.[1-4] Hypothyroidism and anaemia can each cause nonspecific symptoms of ill health like fatigue, and both lead to decreased quality of life. The combination of anaemia and abnormal thyroid function may therefore be accompanied by serious morbidity and further effects on quality of life.

The co-occurrence of anaemia and hypothyroidism is not only a challenging diagnostic problem in allocating symptoms to one of the diseases, but may also point to a causal relationship between thyroid disease and anaemia.[5] Indeed, relationships between thyroid disease and anaemia have already been documented in experimental animal studies in the distant past. [5] For instance, hypophysectomised mammals were found to have decreased red blood cell counts that corrected after administration of thyroid hormones.[6,7] Additionally, mice deficient in the thyroid hormone receptor TRα have been found to have decreased haematocrit values.[8]

However, human data regarding relationships between thyroid disease and hematologic anomalies are scarce. Researchers investigating potential altered erythropoiesis as a result of thyroid dysfunction found red cell abnormalities and a reduced proliferative potential of hematopoietic progenitor cells in both patients with hypothyroidism and hyperthyroidism, but the total number of studied participants was low.[9,10]

In addition, a higher prevalence of anaemia was identified in older male patients with subclinical hypothyroidism [11] and in patients with clinical hypothyroidism [12], but incidence estimates were not available due to the cross-sectional study design. Additionally, a rise of thyroid hormone levels or a decrease in levels of TSH within the reference ranges was associated with higher erythropoietic activity,[13] but the low number of studied participants precluded stratification by hyperthyroid subgroups. In one population-based cohort, both hypothyroidism and hyperthyroidism were associated with decreased haemoglobin in cross-sectional analyses but not in longitudinal analyses.[14]

Clinical experimental evidence on the causal relation between low thyroid function and anaemia is currently limited to a number of small case series in which treatment of hypothyroidism with levothyroxine resulted in a considerable increase in haemoglobin and resolution of anaemia.[12,15,16] Alternatively, and in line with the observational data, in a cohort of patients with hyperthyroidism, a high prevalence of anaemia was found, which returned to normal following antithyroid therapy.[17]

Despite the myriad of smaller studies hinting at a potential relationship between thyroid dysfunction and anaemia, methodologically sound pooled estimates drawn from large and representative populations are missing. In the current study, we sought to determine the association between thyroid hormone status and anaemia in cross-sectional and longitudinal analyses by performing an individual participant data meta-analysis on data from 16 independent observational cohort studies participating in the Thyroid Studies Collaboration.

METHODS

Study population

We performed an individual participant data meta-analysis of cohorts participating in the Thyroid Studies Collaboration. The cohorts are summarised in Table 1 and described elsewhere in detail.[4,18-21] For the current project, we included the 16 cohorts in which thyroid function tests and haemoglobin were measured at baseline.

Anaemia

Anaemia was defined according to the World Health Organization criteria (haemoglobin concentration < 130 g/L in men and < 120 g/L in women).[22] In 14 cohorts, a follow-up measurement of haemoglobin was available.

Thyroid function

TSH and free T4 concentrations were measured at baseline in all cohorts. Cohort-specific cut-off values were applied for free T4 concentrations (Supplemental Table 1). Participants with a TSH level of 0.45 to 4.5 mIU/L were categorised as euthyroid. Overt hypothyroidism was defined as a TSH level > 4.5 mIU/L in combination with reduced free T4 concentration. Subclinical hypothyroidism was defined as a TSH level > 4.5 mIU/L in combination with a normal free T4 concentration. A TSH level < 0.45 mIU/L with normal free T4 levels was defined as subclinical hyperthyroidism. Overt hyperthyroidism was defined as a TSH level < 0.45 mIU/L with an elevated free T4 concentration.[4]

Statistical analyses

We performed a two-stage individual participant data meta-analysis to allow for consistent definitions and analyses across the cohorts, increased analytical flexibility, and decreased complexity of the analyses.[18,23-26] In the first step, the cross-sectional and longitudinal associations between thyroid hormone status and anaemia in each study cohort were estimated separately from supplied original study datasets with data on the participant level. In the second step, all effect estimates found in step one were pooled using random-effects models (DerSimonian and Laird) with inverse variance weighting.

For the cross-sectional association between thyroid hormone status and anaemia at baseline, logistic regression models were constructed. Prospectively, we investigated the risk of developing anaemia during follow-up using Cox regression models; participants with pre-existing anaemia were excluded. The analyses were based on the thyroid function category at baseline. If a new case of anaemia was identified, it was assumed that the anaemia had developed halfway through the follow-up period.

Thyroid status was included as a categorical variable (overt hypothyroidism, subclinical hypothyroidism, subclinical hyperthyroidism, and overt hyperthyroidism), with euthyroidism as the reference group. All models were adjusted for age and sex. A p value for trend was obtained for both overt and subclinical hypothyroid and hyperthyroid categories. Subgroup analyses, including calculations of a p value for interaction, were performed separately for sex, age groups, and ethnicity.

In sensitivity analyses, we excluded all participants who used antithyroid medication or thyroid hormone replacement therapy at baseline or during follow-up. We also compared mean haemoglobin levels at baseline between thyroid status groups and differences in mean annual change in haemoglobin levels during follow-up between thyroid status groups using linear regression models. Additionally, we excluded all participants with a high level of C-reactive protein [(CRP); > 20 mg/L] as a proxy for chronic inflammatory disease.

Data analyses were performed using IBM SPSS Statistics Version 23 and Review Manager 5.3 from the Cochrane Collaboration.

RESULTS

For this study, individual participant data of 56,297 participants from 16 different cohorts participating in the Thyroid Studies Collaboration were available. At baseline, thyroid function (TSH and free T4) and haemoglobin measurements were available from 42,162 participants, of whom 459 (1.1%) had overt hypothyroidism, 2,930 (6.9%) had subclinical hypothyroidism, 36,081 (85.6%) were euthyroid, 2,386 (5.7%) had subclinical hyperthyroidism, and 306 (0.7%) had overt hyperthyroidism.

Baseline characteristics of the cohorts are presented in Table 1. The overall median age of each cohort ranged from 46 to 85 years, and the overall percentage of women was 51.0%. More detailed information about the study participants is presented in Supplemental Tables 2 and 3. The participants excluded because their thyroid function or haemoglobin measure-

Table 1. Baseline characteristics of individuals in included studies (N=42,162).

Study	Study population	Total number of participants: baseline/ follow-up	Age, Median (Range), y	Women (%)	Antithyroid or thyroid medication at baseline (%)	Anaemia at baseline (%)	Anaemia during follow-up (%)	Duration of follow-up, Median (IQR), y	Total person years
Total		42,162/25,466	14-103	22,308 (52.9)	1,067 (2.5)	4,274 (10.1)	2,423 (5.7)	5.7 (3-9.5)	162,583
Bari study	Outpatients with heart failure followed up by Cardiology Department in Bari, Italy	337/206	66 (21-92)	78 (20.5)	23 (6.8)	69 (20.5)	30 (8.9)	1.4 (0.7-1.9)	273
BELFRAIL	Subjects aged 80 years and older in three well-circumscribed areas of Belgium.	524/331	84 (80-100)	331 (63.2)	52 (9.9)	106 (20.2)	52 (9.9)	1.6 (1.4-1.8)	521
Busselton Health study	Adults living in Busselton, Western Australia	2,074/1245	51 (17-90)	1,030 (49.7)	27 (1.3)	76 (3.7)	54 (2.6)	14.0 (14.0-14.0)	17,164
Cardiovascular Health study	Community-dwelling adults with Medicare eligibility in 4 US communities.	3,106/2314	71 (64-100)	1,864 (60.0)	0	259 (8.3)	321 (10.3)	3.0 (3.0-3.0)	12,552
EPIC-Norfolk study	Adults aged 45-79 years living in Norfolk, England	13 1/27	59 (40-78)	7,276 (54.8)	NA	1,090 (8.2)	499 (3.8)	4.3 (3.4-12.3)	57,604
Health, Aging, and Body Composition study	Community dwelling adults aged 70-79 years with Medicare eligibility in 2 US communities	2,531/1236	74 (70-81)	1,305 (51.6)	253 (10.0)	384 (15.2)	195 (7.7)	7.5 (7.5-7.5)	8,543
InChianti study	Community dwelling from two small towns in Tuscany, Italy. Invecchiare in Chianti, "Aging in the Chianti Area" (InCHIANTI) study.	1,200/944	72 (21-103)	675 (56.3)	30 (2.5)	120 (10.0)	177 (14.8)	9.0 (6.0-9.2)	6,958
Longitudinal Aging Study Amsterdam (LASA)	Random sample of older men and women (aged 55-85) in Amsterdam, Zwolle, and Oss, the Netherlands.	766/329	68 (55-85)	393 (51.3)	14 (1.8)	43 (5.6)	28 (3.7)	3.0 (3.0-3.1)	974
Leiden 85-plus Study	All adults aged 85 years living in Leiden, the Netherlands.	555/397	85 (NA)	368 (66.3)	20 (3.6)	158 (28.5)	98 (17.7)	3.0 (0.5-5.0)	1,324
Nagasaki Adult Health study	Atomic bomb survivors in Nagasaki, Japan.	965/753	57 (38-92)	578 (59.9)	11 (1.1)	179 (18.5)	196 (20.3)	11.9 (7.4-12.0)	7,196

Table 1. Baseline characteristics of individuals in included studies (N=42,162). (*continued*)

Study	Study population	Total number of participants: baseline/ follow-up	Age, Median (Range), y	Women (%)	Antithyroid or thyroid medication at baseline (%)	Anaemia at baseline (%)	Anaemia during follow-up (%)	Duration of follow-up, Median (IQR), y	Total person years
Pisa cohort	Patients admitted to cardiology department in Pisa, Italy.	2,259/NA	68 (14-96)	785 (34.7)	NA	490 (21.7)	NA	NA	NA
PREVEND study	Inhabitants, aged 28–75 years, of the city of Groningen, The Netherlands.	934/779	60 (35-82)	397 (42.5)	NA	106 (11.3)	82 (8.8)	5.7 (5.7-5.7)	8,247
PROSPER study	Older community dwelling adults at high cardiovascular risk in the Netherlands, Ireland, and Scotland.	5,769/5138	75 (69-83)	2,983 (51.7)	256 (4.4)	402 (7.0)	203 (3.5)	0.25 (0.25-0.25)	1,261
Rotterdam Study	All inhabitants of the suburb Ommoord in Rotterdam, the Netherlands, aged 55 years and over.	1,835/1322	69 (55-93)	1,135 (61.9)	45 (2.5)	109 (5.9)	214 (11.7)	11.1 (6.6-17.4)	14,066
SHIP	Adults living in Western Pomerania, Germany.	4,214/2882	50 (20-81)	2,139 (50.8)	263 (6.2)	589 (14.0)	274 (6.5)	10.0 (5.0-11.0)	25,900
Whickham Survey	Adults living in and near Newcastle upon Tyne, England.	1,807/NA	46 (18-93)	971 (53.7)	73 (4.0)	94 (5.2)	NA	NA	NA

Abbreviations: EPIC, European Prospective Investigation of Cancer; Health ABC, Health, Aging and Body Composition; IQR, interquartile range (25th–75th percentiles); NA, not available; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; SHIP, Study of Health in Pomerania.

ment were not available had a median age ranging from 45 to 84 years; the percentage of women was 51.5%.

Cross-sectional analyses

At baseline, 4,274 (10.1%) participants had anaemia: 15.9% in the overt hypothyroid group, 11.6% in the subclinical hypothyroid group, 9.7% in the euthyroid group, 13.6% in the subclinical hyperthyroid group, and 11.1% in the overt hyperthyroid group. Participants with subclinical or overt hypothyroidism and subclinical or overt hyperthyroidism had increased odds of having anaemia compared with participants with euthyroidism (Table 2; Figure 1). The pooled OR for the overt hypothyroid group was 1.84 (95% CI 1.35 to 2.50), 1.21 (1.02 to 1.43) for the group with subclinical hypothyroidism, 1.27 (1.03 to 1.57) for those with subclinical hyperthyroidism, and 1.69 (1.00 to 2.87) for those in the overt hyperthyroid group. We observed statistically significant trends from euthyroidism to hypothyroidism (i.e., from subclinical hypothyroidism to overt hypothyroidism; $p = 0.01$) and from euthyroidism to hyperthyroidism (i.e., from subclinical hyperthyroidism to overt hyperthyroidism; $p = 0.04$). When the analyses were stratified by sex, we observed no statistically significant differences (all p values for interaction > 0.05) between men and women (Table 2). Also, no statistically noteworthy differences were observed among different age categories or among white, black, or Asian participants.

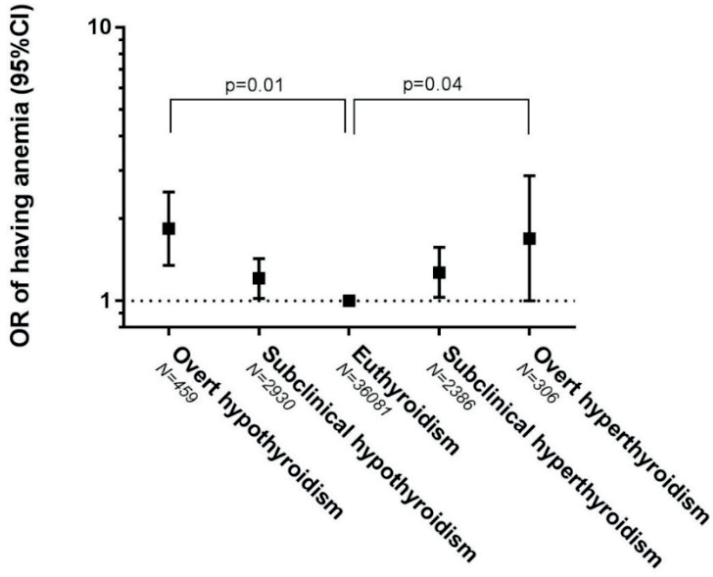


Figure 1. The pooled ORs of the risk of having anaemia at baseline with the 95% CI and p value for trend. Logistic regression models corrected for age and sex; reference group is euthyroidism.

Table 2. The risk of having anaemia at baseline according to thyroid hormone status (N=42,162 from 16 cohorts).

	Overt Hypothyroidism	Subclinical Hypothyroidism	Euthyroidism	Subclinical Hyperthyroidism	Overt Hyperthyroidism	N Overt Hypothyroidism/ Subclinical Hypothyroidism/ Subclinical Hyperthyroidism/ Overt Hyperthyroidism
All[a]	1.84 (1.35-2.50)	1.21 (1.02-1.43)	1 (ref)	1.27 (1.03-1.57)	1.69 (1.00-2.87)	459/2,930/36,081/2,386/306
Sex						
Male	2.45 (1.45-4.12)	1.27 (1.03-1.57)	1 (ref)	1.19 (0.95-1.49)	1.59 (0.80-3.14)	122/1,029/17,546/1,055/102
Female	1.79 (1.30-2.47)	1.23 (0.99-1.52)	1 (ref)	1.42 (1.11-1.81)	1.78 (0.99-3.21)	337/1,901/18,535/1,331/204
Age, y						
< 50[b]	2.25 (1.10-4.60)	1.15 (0.77-1.74)	1 (ref)	1.27 (0.73-2.21)	3.53 (0.26-48.39)	48/452/6,763/599/27
50-65	5.53 (0.93-33.03)	1.44 (0.94-2.21)	1 (ref)	1.88 (1.09-3.24)	4.71 (1.25-17.78)	132/677/9,719/670/63
65-80	2.02 (1.02-3.99)	1.40 (1.10-1.78)	1 (ref)	1.21 (0.85-1.73)	1.49 (0.89-2.51)	215/1,711/16,814/949/186
> 80	1.91 (1.01-3.62)	1.03 (0.68-1.54)	1 (ref)	1.49 (0.99-2.23)	2.66 (0.35-20.26)	65/234/2,646/168/27
Ethnicity						
White[c]	1.97 (1.37-2.82)	1.29 (1.11-1.51)	1 (ref)	1.30 (1.04-1.63)	1.56 (1.03-2.34)	431/2,687/34,154/2,339/303
Black[d]	0.96 (0.20-4.58)	1.51 (0.74-3.05)	1 (ref)	0.77 (0.31-1.89)	-	12/98/1,013/35/2
Asian[e]	2.01 (0.63-6.39)	0.87 (0.54-1.39)	1 (ref)	0.82 (0.10-6.83)	-	13/143/828/8/1
Other[f]	-	-	1 (ref)	-	-	0/0/22/1/0

Data are pooled hazard ratio (95% CI) unless otherwise noted. Results were obtained by logistic regression analysis, adjusted for age (if applicable) and sex.

[a] p for trend: overt hyperthyroidism to euthyroidism, p=0.01; euthyroidism to overt hyperthyroidism p=0.04

[b] Reference group is < 50 y.

[c] Reference group is white.

[d] Only data from CHS, HEALTH ABC and PREVEND.

[e] Only data from LASA, Nagasaki, PREVEND and Rotterdam.

[f] Only data from LASA, PREVEND and Rotterdam.

Longitudinal analyses

In the longitudinal analyses, 25,466 participants from 14 cohorts were included, with a median follow-up time of 5.7 years (interquartile range 3.0 to 9.5). A total of 2,423 participants developed anaemia during follow-up (14.9 per 1,000 person-years): 12.2% in the overt hypothyroid group, 12.0% in the subclinical hypothyroid group, 9.2% in the euthyroid group, 10.7% in the subclinical hyperthyroid group, and 8.7% in the overt hyperthyroid group (Table 3; Figure 2). The pooled hazard ratios for the risk of developing anaemia were 1.38 (95% CI 0.86 to 2.20) for the overt hypothyroid group, 1.18 (1.00 to 1.38) for the group with subclinical hypothyroidism, 1.15 (0.94 to 1.42) for the group with subclinical hyperthyroidism, and 1.47 (0.91 to 2.38) in the overt hyperthyroid group. We observed a statistically significant trend from euthyroidism to hypothyroidism (p = 0.02). No statistically significant trend was observed for euthyroidism to hyperthyroidism (p = 0.20). When the participants were stratified by sex, age, or ethnicity, these findings remained unchanged. Associations were more pronounced in those studies with a median follow-up ≥ 5 years (Supplemental Table 4).

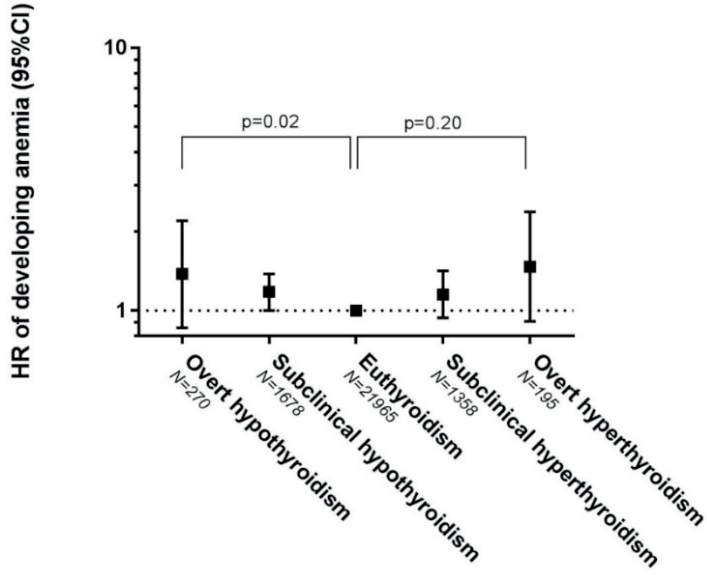


Figure 2. The pooled hazard ratios of developing anaemia during follow-up in the thyroid function groups with the 95% CI and p value for trend. Logistic regression models corrected for age and sex; reference group is euthyroidism.

Table 3. The risk of developing anaemia during follow-up according to thyroid hormone status at baseline (n=25,466 from 14 cohorts).

	Overt Hypothyroidism	Subclinical Hypothyroidism	Euthyroidism	Subclinical Hyperthyroidism	Overt Hyperthyroidism	N Overt Hypothyroidism/ Subclinical Hypothyroidism/ Subclinical Hyperthyroidism/ Overt Hyperthyroidism
All[a]	1.38 (0.86-2.20)	1.18 (1.00-1.38)	1 (ref)	1.15 (0.94-1.42)	1.47 (0.91-2.38)	270/1,678/21,965/1,358/195
Sex						
Male	2.14 (0.79-5.79)	1.05 (0.83-1.34)	1 (ref)	1.41 (0.92-2.18)	0.83 (0.26-2.61)	62/577/10,450/584/63
Female	1.19 (0.75-1.88)	1.37 (1.05-1.80)	1 (ref)	1.22 (0.95-1.57)	2.27 (1.30-3.93)	207/1,096/11,487/773/131
Age, y						
< 50[b]	17.81 (4.06-78.24)	1.48 (0.78-2.83)	1 (ref)	1.09 (0.68-1.73)	-	16/106/3,857/343/21
50-65	1.58 (0.14-18.00)	1.14 (0.82-1.58)	1 (ref)	1.02 (0.70-1.50)	2.97 (0.56-15.64)	85/384/5,872/414/32
65-80	1.39 (0.81-2.37)[c]	1.20 (0.96-1.51)	1 (ref)	1.14 (0.83-1.57)	1.51 (0.85-2.69)	130/1,069/10,737/521/124
> 80	1.48 (0.55-4.03)[c]	1.30 (0.80-2.10)	1 (ref)	1.57 (0.98-2.50)	3.59 (0.49-26.06)	39/119/1,499/80/18
Ethnicity						
White[c]	1.38 (0.75-2.54)	1.21 (1.00-1.45)	1 (ref)	1.16 (0.93-1.44)	1.52 (0.93-2.50)	257/1,527/20,849/1,336/193
Black[d]	2.80 (0.38-20.76)	1.30 (0.59-2.83)	1 (ref)	1.57 (0.57-4.36)	-	5/41/415/16/1
Asian[e]	1.68 (0.53-5.29)	1.03 (0.70-1.51)	1 (ref)	-	-	6/109/654/6/1
Other[f]	-	-	1 (ref)	-	-	0/0/12/0/0

Data are pooled hazard ratio (95% CI) unless otherwise noted. Results were obtained by logistic regression analysis, adjusted for age (if applicable) and sex.

[a] p for trend: overt hyperthyroidism to euthyroidism, p=0.02; euthyroidism to overt hyperthyroidism p=0.20

[b] Reference group is < 50 y.

[c] p value for interaction ($p < 0.05$).

[d] Reference group is white.

[e] Only data from CHS, HEALTH ABC and PREVEND.

[f] Only data from LASA, Nagasaki, PREVEND and Rotterdam.

[g] Only data from LASA, PREVEND and Rotterdam.

Additional analyses

Cross-sectionally, haemoglobin levels (as a continuous variable) were lower (mean difference between −0.06 and −0.19 g/dL) in all groups compared with participants with euthyroidism (Supplemental Table 5). Prospectively, no differences in mean annual change in haemoglobin levels were observed among the thyroid hormone status groups (Supplemental Table 6). Similar results were observed when analyses were stratified on sex. In addition, sensitivity analyses excluding participants who used thyroid hormone medication or with high levels of CRP yielded higher ORs in line with the unrestricted results but with wider CIs (Supplemental Tables 7 and 8).

For all main analyses, I^2 statistics remained < 40% (Supplemental Tables 9 and 10), and, in combination with size and direction of effects, statistical heterogeneity was deemed low to negligible.[27]

DISCUSSION

In this individual participant data meta-analysis, we observed a cross-sectional relation between thyroid function and anaemia; higher odds of anaemia were observed in participants with both overt and subclinical hypothyroidism as well as overt and subclinical hyperthyroidism. In addition, reduced thyroid function at baseline showed a trend of increased risk of developing anaemia during follow-up. The longitudinal association between overt and subclinical hyperthyroidism and the risk of developing anaemia did not reach statistical significance. Prospectively, no differences in mean annual change in haemoglobin levels were observed among the thyroid hormone status groups.

The findings in the current individual participant data meta-analysis build on findings from earlier studies in which thyroid dysfunction was associated with abnormal red blood cell indices.[11-13] In this study, thyroid dysfunction, whether overt or subclinical hypothyroidism and hypothyroidism, was associated with slightly lower haemoglobin levels. Given the small difference in haemoglobin levels among thyroid function groups, the contribution of thyroid dysfunction on low haemoglobin levels or anaemia may be small. It remains to be assessed in a randomised controlled trial whether treatment of (subclinical) hypothyroidism is effective in reducing anaemia to further decide whether the findings are thought to be clinically relevant and whether these should influence practice and policies. Christ-Crain *et al.* [28] showed that erythropoietin levels increased after thyroxin treatment in patients with subclinical hypothyroidism. In addition, a number of studies have also shown a beneficial effect of thyroid hormone treatment in patients with hypothyroidism on erythropoietin levels.[12,15,16]

There are numerous types of anaemia that can be classified according to whether the anaemia is primarily the result of blood loss, deficits in the production of healthy erythrocytes, or by reduced erythrocyte survival. Currently, it is unclear what mechanisms exactly allow thyroid function and erythropoiesis to be linked pathophysiologically and how both ends of the thyroid disease spectrum might lead to an anaemic state. However, for subclinical and overt hyperthyroidism, several pathways have been proposed. Hyperthyroidism might be associated with anaemia via reduced erythrocyte survival due to altered iron metabolism and utilization, enhanced oxidative stress, and increased haemolysis.[29,30] Thyroid hormones stimulate energy metabolism, resulting in an enhanced requirement of oxygen delivery to the tissues speeding up destructive processes.

For subclinical and overt hypothyroidism, there is accumulating evidence that indicates low thyroid function may be causally related to anaemia via deficits in the production of healthy erythrocytes, although the underlying mechanisms by which thyroid hormones and TSH may lead to anaemia are not fully understood.[31] T3, T4, and TSH may play a direct role in erythropoiesis.[32] For instance, both T3 and T4 are involved in the regulation of haematopoiesis by influencing erythroid precursor proliferative capacity.[33] In addition, a direct β2-adrenergic receptor–mediated stimulation of red cell precursors by T4 has been shown.[34] T4 has also been found to stimulate the initiation and completion of haemoglobin protein chains in vitro and to enhance red blood cell formation.[5] Thyroid hormones were also shown to promote erythropoiesis by increasing the production of erythropoietin by the kidneys.[35] Furthermore, there is evidence that thyroid hormones affect iron transport and utilization. TSH could affect haematopoiesis by binding to a functional TSH receptor, which can be found in erythrocytes and some extrathyroidal tissues.[10] Another explanation for the co-occurrence of low thyroid function and anaemia is that there are common causes for abnormal thyroid status and anaemia. Chronic (inflammatory) diseases, malnutrition, and malabsorption may all result in reduced thyroid status as an adaptive response to energetic deficits. In addition, malnutrition and malabsorption may cause deficiencies of micronutrients that are crucial for erythropoiesis, like iron, vitamin B12, and folate, as well as iodine deficiency, which is crucial for normal thyroid function. Interestingly, iron deficiency, which is the most common cause of anaemia, was also found to decrease the activity of thyroid peroxidase, an iron-containing enzyme involved in the synthesis of thyroid hormones.[36]

Strengths of the current individual participant data meta-analysis are the inclusion of individual participant data of large cohort studies from across the globe. The availability of individual participant data allowed us to choose clinically relevant categories of thyroid function and anaemia, standardise these definitions, and perform several standardised subgroup analyses.

An individual participant data meta-analysis of well-designed observational studies can be considered an important tool in assessing causality. When studying causality, the nine considerations of Hill in 1965 [37] can be used as a checklist. In our study, many of these considerations are met. Although the individual study cohorts and individual subgroups may have been small, we had sufficient power to study the associations in this pooled analysis because of the increased combined sample size. Because multiple studies were included, we could also study consistency in the results of the different cohorts (e.g., effect estimates all pointing in the same direction); the low level of heterogeneity also aids in considering a causal relation. In addition, the availability of the individual participant data allowed us to define identical subgroups for each study in a biological gradient, from overt hypothyroidism to overt hyperthyroidism. The availability of prospective observational data are also in compliance with the fourth consideration of temporality; in 14 studies, a baseline measurement of the determinant (thyroid function) and (baseline and) follow-up measurements of the outcome of interest (haemoglobin) were available. Therefore, our pooled analysis of observational studies satisfies multiple criteria of Hill. However, it remains to be assessed in a well-designed, randomised controlled trial with a considerable number of participants with (subclinical) hypothyroidism if treatment is effective in reducing anaemia. Further analysis of the data from two well-designed, randomised controlled trials for subclinical hypothyroidism in older persons (TRUST and IEMO Thyroid Trial [38,39]) could be a first attempt at uncovering the clinical relevance of thyroid influences on haemoglobin levels.

Some limitations of this study have to be acknowledged as well. First, a limitation of this pooled analysis is that TSH and free T4 were only measured once at baseline. Because subclinical hypothyroidism has been shown to normalise in one-third of cases,[40] in guidelines, it is often recommended that measurements of these parameters are repeated. Unfortunately, repeated TSH and free T4 measurements were not available in many cohorts. Erroneously classifying patients with euthyroidism based on one measurement may have led to an underestimation of the associations found. Second, the statistical power was more limited in the longitudinal models than in the baseline, cross-sectional analysis. The association between overt and subclinical hyperthyroidism and the risk of developing anaemia did not reach statistical significance, but the results of the longitudinal analyses followed a similar pattern. Third, we did not apply age-adjusted reference ranges as per current consensus and usual practice. However, evidence in favour of age-specific TSH reference ranges is starting to accumulate;[41] so, too, is evidence to the contrary.[42-44] This is an important topic of future research. Fourth, we performed sensitivity analyses excluding participants with high CRP levels as a proxy for chronic diseases that might predispose to anaemia, but this only excluded diseases associated with inflammation. Particularly in the group of participants with subclinical hypothyroidism, the possibility of the presence of nonthyroidal illness cannot be fully excluded. As a result, possible residual errors caused by residual bias and confounding

may have deflated the results. Unfortunately, information on additional potential confounding factors, like thyroid medication dose titrations, other diseases relating to anaemia (cancer, chronic kidney disease, leukaemia, gastric ulcers, arthritis, or chronic obstructive pulmonary disease), menopausal state, nonthyroidal illness, concomitant medications, and iron or vitamin supplements, was not available for most cohorts.

In conclusion, we observed higher odds of anaemia in both participants with hypothyroid and hyperthyroid function. In addition, reduced thyroid function at baseline showed a trend of increased risk of developing anaemia during follow-up. It remains to be assessed in a randomised controlled trial whether treatment of (subclinical) hypothyroidism is effective in reducing anaemia.

REFERENCES

1. Beghe C, Wilson A, Ershler WB. Prevalence and outcomes of anemia in geriatrics: a systematic review of the literature. *Am J Med.* 2004;116 Suppl 7A(7):3S-10S.
2. Cooper DS, Biondi B. Subclinical thyroid disease. *Lancet (London, England).* 2012;379(9821):1142-1154.
3. Gaskell H, Derry S, Andrew Moore R, McQuay HJ. Prevalence of anaemia in older persons: systematic review. *BMC Geriatr.* 2008;8:1.
4. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA.* 2010;304(12):1365-1374.
5. Fein HG, Rivlin RS. Anemia in thyroid diseases. *The Medical clinics of North America.* 1975;59(5):1133-1145.
6. Evans ES, Rosenberg LL, Simpson ME. Erythropoietic response to calorigenic hormones. *Endocrinology.* 1961;68(3):517-532.
7. Horsley V. The Brown Lectures on Pathology. *Br Med J.* 1885;1(1261):419-423.
8. Kendrick TS, Payne CJ, Epis MR, et al. Erythroid defects in TRalpha-/- mice. *Blood.* 2008;111(6):3245-3248.
9. Das KC, Mukherjee M, Sarkar TK, Dash RJ, Rastogi GK. Erythropoiesis and erythropoietin in hypo- and hyperthyroidism. *J Clin Endocrinol Metab.* 1975;40(2):211-220.
10. Kawa MP, Grymula K, Paczkowska E, et al. Clinical relevance of thyroid dysfunction in human haematopoiesis: biochemical and molecular studies. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(2):295-305.
11. den Elzen WP, de Craen AJ, Mooijaart SP, Gussekloo J. Low thyroid function and anemia in old age: the Leiden 85-plus study. *J Am Geriatr Soc.* 2015;63(2):407-409.
12. Horton L, Coburn RJ, England JM, Himsworth RL. The haematology of hypothyroidism. *The Quarterly journal of medicine.* 1976;45(177):101-123.
13. Bremner AP, Feddema P, Joske DJ, et al. Significant association between thyroid hormones and erythrocyte indices in euthyroid subjects. *Clin Endocrinol (Oxf).* 2012;76(2):304-311.
14. Floriani C, Feller M, Aubert CE, et al. Thyroid Dysfunction and Anemia: A Prospective Cohort Study and a Systematic Review. *Thyroid.* 2018;28(5):575-582.
15. Tudhope GR, Wilson GM. Anaemia in hypothyroidism. Incidence, pathogenesis, and response to treatment. *The Quarterly journal of medicine.* 1960;29:513-537.
16. Vitale G, Fatti LM, Prolo S, et al. Screening for hypothyroidism in older hospitalized patients with anemia: a new insight into an old disease. *J Am Geriatr Soc.* 2010;58(9):1825-1827.
17. Gianoukakis AG, Leigh MJ, Richards P, et al. Characterization of the anaemia associated with Graves' disease. *Clin Endocrinol (Oxf).* 2009;70(5):781-787.
18. Blum MR, Bauer DC, Collet TH, et al. Subclinical thyroid dysfunction and fracture risk: a meta-analysis. *JAMA.* 2015;313(20):2055-2065.
19. Huisman M, Poppelaars J, van der Horst M, et al. Cohort profile: the Longitudinal Aging Study Amsterdam. *Int J Epidemiol.* 2011;40(4):868-876.
20. Ittermann T, Haring R, Sauer S, et al. Decreased serum TSH levels are not associated with mortality in the adult northeast German population. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(3):579-585.
21. Meuwese CL, van Diepen M, Cappola AR, et al. Low thyroid function is not associated with an accelerated deterioration in renal function. *Nephrol Dial Transplant.* 2019;34(4):650-659.

22. Nutritional anaemias. Report of a WHO scientific group. *World Health Organization technical report series.* 1968;405:5-37.
23. Ankle Brachial Index C, Fowkes FG, Murray GD, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA.* 2008;300(2):197-208.
24. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-188.
25. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ.* 2010;340(feb05 1):c221.
26. Simmonds MC, Higgins JP, Stewart LA, Tierney JF, Clarke MJ, Thompson SG. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin Trials.* 2005;2(3):209-217.
27. Cochrane Handbook for Systematic Reviews of Interventions. In: www.handbook.cochrane.org.
28. Christ-Crain M, Meier C, Huber P, Zulewski H, Staub JJ, Muller B. Effect of restoration of euthyroidism on peripheral blood cells and erythropoietin in women with subclinical hypothyroidism. *Hormones (Athens, Greece).* 2003;2(4):237-242.
29. Asl SZ, Brojeni NK, Ghasemi A, Faraji F, Hedayati M, Azizi F. Alterations in osmotic fragility of the red blood cells in hypo- and hyperthyroid patients. *J Endocrinol Invest.* 2009;32(1):28-32.
30. Yucel R, Ozdemir S, Dariyerli N, Toplan S, Akyolcu MC, Yigit G. Erythrocyte osmotic fragility and lipid peroxidation in experimental hyperthyroidism. *Endocrine.* 2009;36(3):498-502.
31. Maggio M, De Vita F, Fisichella A, et al. The Role of the Multiple Hormonal Dysregulation in the Onset of "Anemia of Aging": Focus on Testosterone, IGF-1, and Thyroid Hormones. *Int J Endocrinol.* 2015;2015:292574.
32. Perrin MC, Blanchet JP, Mouchiroud G. Modulation of human and mouse erythropoiesis by thyroid hormone and retinoic acid: evidence for specific effects at different steps of the erythroid pathway. *Hematol Cell Ther.* 1997;39(1):19-26.
33. Golde DW, Bersch N, Chopra IJ, Cline MJ. Thyroid hormones stimulate erythropoiesis in vitro. *Br J Haematol.* 1977;37(2):173-177.
34. Sullivan PS, McDonald TP. Thyroxine suppresses thrombocytopoiesis and stimulates erythropoiesis in mice. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine (New York, NY).* 1992;201(3):271-277.
35. Fandrey J, Frede S, Jelkmann W. Role of hydrogen peroxide in hypoxia-induced erythropoietin production. *Biochem J.* 1994;303 (Pt 2)(Pt 2):507-510.
36. Khatiwada S, Gelal B, Baral N, Lamsal M. Association between iron status and thyroid function in Nepalese children. *Thyroid Res.* 2016;9:2.
37. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med.* 1965;58(5):295-300.
38. IEMO 80-plus Thyroid Trial: Nederlands Trial Register. www.trialregister.nl/trialreg/admin/rctview.asp?TC=3851. Accessed 30 July 2018.
39. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.
40. Diez JJ, Iglesias P, Burman KD. Spontaneous normalization of thyrotropin concentrations in patients with subclinical hypothyroidism. *J Clin Endocrinol Metab.* 2005;90(7):4124-4127.
41. Surks MI, Boucai L. Age- and race-based serum thyrotropin reference limits. *J Clin Endocrinol Metab.* 2010;95(2):496-502.
42. Fatourechi V. Upper limit of normal serum thyroid-stimulating hormone: a moving and now an aging target? *J Clin Endocrinol Metab.* 2007;92(12):4560-4562.

43. Kahapola-Arachchige KM, Hadlow N, Wardrop R, Lim EM, Walsh JP. Age-specific TSH reference ranges have minimal impact on the diagnosis of thyroid dysfunction. *Clin Endocrinol (Oxf)*. 2012;77(5):773-779.
44. Laurberg P, Andersen S, Carle A, Karmisholt J, Knudsen N, Pedersen IB. The TSH upper reference limit: where are we at? *Nat Rev Endocrinol*. 2011;7(4):232-239.

Supplemental table 1. Cohort specific cut-off values of free T4 and the total number of participants per thyroid category per study (baseline/follow-up).

Study	Free T4	Overt hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt hyperthyroidism
Total		459/270	2,930/1,678	36,081/21,965	2,386/1,358	306/195
Bari study	0.7 – 1.8 ng/dL	-	40/19	289/183	7/4	1/0
BELFRAIL	0.9 – 1.8 ng/dL	11/8	4/3	455/286	54/34	-
Busselton Health study	9.0 – 23.0 pmol/L	13/8	97/50	1,898/1,155	52/27	14/5
Cardiovascular Health study	0.7 – 1.7 ng/dL	38/29	503/351	2,519/1,906	44/27	2/1
EPIC-Norfolk study	9.0 – 20 pmol/L	217/115	731/414	11,888/6,874	367/219	83/35
Health, Aging, and Body Composition study	0.8 – 1.8 ng/dL	23/11	313/167	2,110/1,013	79/42	6/3
InChianti study	0.8 – 2.2 ng/dL	9/6	32/24	1,054/849	88/57	17/8
Longitudinal Aging Study Amsterdam (LASA)	11 – 22 pmol/L	9/4	15/6	695/302	42/15	5/2
Leiden 85-plus Study	13 – 23 pmol/L	40/26	35/26	455/330	23/13	2/2
Nagasaki Adult Health study	0.8 – 2.5 ng/dL	13/6	143/109	800/631	8/6	1/1
Pisa cohort	7.1 – 18.5 pg/mL	3/NA	117/NA	2,025/NA	107/NA	7/NA
PREVEND study	9.14 – 23.81 mmol/L	3/3	60/49	835/698	35/28	1/1
PROSPER study	12 – 18 pmol/L	31/30	445/386	5,056/4,506	132/118	105/98
Rotterdam Study	11 – 25 pmol/L	20/12	108/64	1,577/1,102	118/71	12/6
SHIP	8.3 – 18.9 pmol/L	17/12	15/10	3,097/2,130	1,035/697	50/33
Whickham Survey	3.6 – 13.6 pmol/L	12/NA	272/NA	1,328/NA	195/NA	-

SI unit free T4 is pmol/L. Conversion factor for ng/dL is 12.87, pg/mL is 1.29, mmol/L is 1x10^9. Molecular mass for thyroxine is 776.87 g/mol

Supplemental table 2. Baseline characteristics of individuals in included studies per thyroid category per study (N=42,162).

Characteristics	Overt Hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt Hyperthyroidism
Age, Median (Range), y	26-89	18-95	14-103	16-100	24-90
Bari study	-	70 (62-77)	65 (57-74)	62 (57-82)	64
BELFRAIL	84 (83-87)	85 (83-87)	84 (80-100)	84 (80-97)	-
Busselton Health study	61 (26-71)	58 (19-79)	51 (17-90)	51 (21-84)	59 (33-82)
Cardiovascular Health study	72 (65-84)	72 (65-95)	71 (64-100)	73 (65-98)	73 (68-78)
EPIC-Norfolk study	60 (41-78)	61 (40-78)	58 (40-78)	61 (40-77)	64 (42-77)
Health, Aging, and Body Composition study	73 (70-78)	75 (70-81)	74 (70-81)	76 (70-81)	74 (71-79)
InChianti study	74 (63-87)	76 (39-92)	71 (21-103)	73 (25-100)	70 (24-90)
Longitudinal Aging Study Amsterdam (LASA)	69 (59-79)	68 (56-85)	68 (55-85)	72 (58-85)	68 (63-85)
Leiden 85-plus Study	85	85	85	85	85
Nagasaki Adult Health study	65 (55-89)	58 (39-90)	57 (38-92)	54 (46-80)	83
Pisa cohort	77 (64-78)	71 (22-93)	68 (14-96)	69 (16-93)	70 (24-84)
PREVEND study	60 (56-65)	64 (36-82)	59 (35-82)	67 (45-81)	70
PROSPER study	76 (71-83)	75 (70-83)	75 (69-83)	76 (70-83)	75 (70-83)
Rotterdam Study	70 (55-81)	70 (56-86)	68 (55-93)	69 (55-89)	71 (57-82)
SHIP	56 (26-77)	52 (32-72)	47 (20-81)	57 (20-81)	62 (25-80)
Whickham Survey	55 (38-68)	52 (18-92)	45 (18-89)	43 (19-93)	-
Women (%)	337 (73.4)	1,901 (64.9)	18,535 (51.1)	1,331 (55.8)	204 (66.7)
Bari study	-	11 (27.5)	64 (22.1)	3 (42.9)	0
BELFRAIL	10 (57.9)	2 (50.0)	275 (60.4)	44 (81.5)	-
Busselton Health study	9 (69.2)	67 (69.1)	917 (48.3)	25 (48.1)	12 (85.7)
Cardiovascular Health study	24 (63.2)	327 (65.0)	1,484 (58.9)	28 (63.6)	1 (50.0)
EPIC-Norfolk study	160 (73.7)	515 (70.5)	6,314 (53.1)	233 (63.5)	54 (65.1)

Supplemental table 2. Baseline characteristics of individuals in included studies per thyroid category per study (N=42,162). (*continued*)

Characteristics	Overt Hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt Hyperthyroidism
Health, Aging, and Body Composition study	16 (69.6)	172 (55.0)	1,057 (50.1)	55 (69.6)	5 (83.3)
InChianti study	7 (77.8)	21 (65.6)	582 (55.2)	53 (60.2)	12 (70.6)
Longitudinal Aging Study Amsterdam (LASA)	7 (77.8)	13 (86.7)	342 (49.2)	26 (61.9)	5 (100.0)
Leiden 85-plus Study	31 (77.5)	28 (80.0)	293 (64.4)	15 (65.2)	1 (50.0)
Nagasaki Adult Health study	10 (76.9)	88 (61.5)	478 (59.8)	2 (25.0)	0
Pisa cohort	2 (66.7)	54 (46.2)	681 (33.6)	45 (42.1)	3 (42.9)
PREVEND study	1 (33.3)	36 (60.0)	343 (41.1)	17 (48.6)	0
PROSPER study	21 (67.7)	290 (65.2)	2,498 (49.4)	92 (69.7)	82 (78.1)
Rotterdam Study	16 (80.0)	88 (81.5)	939 (59.5)	82 (69.5)	10 (83.3)
SHIP	14 (82.4)	10 (66.7)	1,573 (50.8)	523 (50.5)	19 (38.0)
Whickham Survey	9 (75.0)	179 (65.8)	695 (52.3)	88 (45.1)	-

Abbreviations: IQR, interquartile range (25^{th}-75^{th} percentiles); NA, data not available.

Supplemental table 3. Overview of the number of participants who had anaemia at baseline or developed anaemia during follow-up per study.

Study	Overt hypothyroidism		Subclinical hypothyroidism		Euthyroidism		Subclinical hyperthyroidism		Overt hyperthyroidism	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Total cases	73 (15.9)	33 (7.2)	341 (11.6)	202 (6.9)	3,501 (9.7)	2,026 (5.6)	325 (13.6)	145 (6.1)	34 (11.1)	17 (5.6)
Bari study	-	-	11 (27.5)	4 (10.0)	56 (19.4)	25 (8.7)	2 (28.6)	1 (14.3)	0	-
BELFRAIL	0	1 (9.1)	1 (25.0)	0	94 (20.7)	44 (9.7)	11 (20.4)	7 (13.0)	-	-
Busselton Health study	2 (15.4)	2 (15.4)	6 (6.2)	2 (2.1)	62 (3.3)	48 (2.5)	5 (9.6)	2 (3.8)	1 (7.1)	0
Cardiovascular Health study	4 (10.5)	4 (10.5)	65 (12.9)	46 (9.1)	188 (7.5)	266 (10.6)	1 (2.3)	5 (11.4)	1 (50.0)	0
EPIC-Norfolk study	30 (13.8)	4 (1.8)	68 (9.3)	34 (4.7)	946 (8.0)	449 (3.8)	33 (9.0)	8 (2.2)	13 (15.7)	4 (4.8)
Health, Aging, and Body Composition study	3 (13.0)	1 (4.3)	47 (15.0)	29 (9.3)	325 (15.4)	157 (7.4)	9 (11.4)	7 (8.9)	0	1 (16.7)
InChianti study	2 (22.2)	1 (11.1)	4 (12.5)	10 (31.3)	93 (8.8)	147 (13.9)	17 (19.3)	18 (20.5)	4 (23.5)	1 (5.9)
Longitudinal Aging Study Amsterdam (LASA)	3 (33.3)	0	1 (6.7)	0	37 (5.3)	26 (3.7)	2 (4.8)	2 (48)	0	0
Leiden 85-plus Study	14 (35.0)	8 (20.0)	9 (25.7)	4 (11.4)	125 (27.5)	80 (17.6)	10 (43.5)	5 (21.7)	0	1 (50.0)
Nagasaki Adult Health study	5 (38.5)	3 (23.1)	25 (17.5)	31 (21.7)	148 (18.5)	162 (20.3)	1 (12.5)	0	0	0
Pisa cohort	1 (33.3)	NA	32 (27.4)	NA	422 (20.8)	NA	33 (30.8)	NA	2 (28.6)	NA
PREVEND study	0	2 (66.7)	4 (6.7)	10 (16.7)	98 (11.7)	67 (8.0)	4 (11.4)	3 (8.6)	0	0
PROSPER study	1 (3.2)	2 (6.5)	44 (9.9)	18 (4.0)	347 (6.9)	174 (3.4)	7 (5.3)	5 (3.8)	3 (2.9)	4 (3.8)
Rotterdam Study	3 (15.0)	3 (15.0)	9 (8.3)	14 (13.0)	91 (5.8)	183 (11.6)	4 (3.4)	12 (10.2)	2 (16.7)	2 (16.7)
SHIP	4 (23.5)	2 (11.8)	4 (26.7)	0	404 (13.0)	198 (6.4)	169 (16.3)	70 (6.8)	8 (16.0)	4 (8.0)
Whickham Survey	1 (8.3)	NA	11 (4.0)	NA	65 (4.9)	NA	17 (8.7)	NA	-	NA

Abbreviations: NA, data not available.

Supplemental table 4. The risk of developing anaemia during follow-up according to thyroid hormone status at baseline.

	Overt hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt hyperthyroidism
Pooled HR[a] (95% CI)					
All	1.38 (0.86-2.20)	1.18 (1.00-1.38)	1 (ref)	1.15 (0.94-1.42)	1.47 (0.91-2.38)
Median follow-up					
< 5 years[b]	0.95 (0.60-1.50)	1.06 (0.87-1.30)	1 (ref)	1.11 (0.77-1.61)	1.43 (0.74-2.77)
≥ 5 years[c]	2.22 (1.11-4.43)	1.37 (1.02-1.85)	1 (ref)	1.21 (0.91-1.61)	1.51 (0.75-3.05)

[a] Results were obtained by cox regression analysis, adjusted for age (if applicable) and sex

[b] Cohorts: Bari study, BELFRAIL, Cardiovascular Health study, EPIC-Norfolk study, Longitudinal Aging Study Amsterdam, Leiden 85-plus Study and PROSPER study

[c] Cohorts: Busselton Health study, Health, Aging, and Body composition study, InChianti study, Nagasaki Adult Health study, PREVEND study, Rotterdam Study and SHIP

Supplemental table 5. The mean difference in haemoglobin levels (continuous in g/dL) at baseline compared to euthyroidism depending on thyroid status.

	Overt hypothyroidism		Subclinical hypothyroidism		Euthyroidism		Subclinical hyperthyroidism		Overt hyperthyroidism	
	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)
All	459	-0.19 (-0.32- -0.06)	2,930	-0.07 (-0.13- -0.01)	36,081	(ref).	2,386	-0.06 (-0.13-0.01)	309	-0.07 (-0.20-0.05)
Sex										
Male	122	-0.31 (-0.52- -0.10)	1,029	-0.08 (-0.15-0.00)	17,546	(ref).	1,055	-0.06 (-0.13-0.02)	102	0.06 (-0.16-0.28)
Female	337	-0.18 (-0.34- -0.02)	1,901	-0.08 (-0.14- -0.02)	18,535	(ref).	1,331	-0.09 (-0.22-0.04)	204	-0.14 (-0.29-0.01)
Age										
< 50 years	47	-0.23 (-0.55-0.09)	308	-0.13 (-0.28-0.02)	6,902	(ref).	599	-0.01 (-0.09-0.08)	30	-0.01 (-0.64-0.62)
50-65 years	132	-0.24 (-0.42- -0.06)	677	-0.05 (-0.14-0.04)	9,715	(ref).	672	-0.12 (-0.25-0.01)	98	-0.06 (-0.26-0.15)
65-80 years	215	-0.21 (-0.45-0.03)	1,711	-0.10 (-0.17- -0.02)	16,814	(ref).	948	0.01 (-0.15-0.18)	186	-0.09 (-0.28-0.09)
> 80 years	65	-0.20 (-0.53-0.12)	234	0.01 (-0.18-0.21)	2,646	(ref).	168	-0.24 (-0.45- -0.03)	27	0.29 (-0.15-0.73)
Ethnicity										
White	431	-0.18 (-0.32- -0.05)	2,687	-0.09 (-0.16- -0.01)	34,154	(ref).	2,339	-0.06 (-0.13-0.01)	303	-0.08 (-0.21-0.04)
Black[b]	12	-0.19 (-0.88-0.50)	98	-0.20 (-0.45-0.05)	1,007	(ref).	35	0.11 (-0.31-0.52)	2	0.72 (-0.98-2.41)
Asian[c]	13	-0.53 (-1.22-0.16)	143	0.00 (-0.23-0.22)	800	(ref).	8	-0.68 (-1.55-0.19)	1	0.37 (-2.08-2.83)
Other[d]	-	-	-	-	13	(ref).	1	1.48 (-0.66-3.62)	-	-
Without medication	201	-0.15 (-0.32-0.01)	1,848	-0.08 (-0.17-0.01)	20,451	(ref).	1,625	-0.05 (-0.14-0.05)	138	0.08 (-0.10-0.26)

[a] Results were obtained by linear regression analysis, adjusted for age (if applicable) and sex

[b] Only data from CHS, HABC and PREVEND

[c] Only data from LASA, Nagasaki, PREVEND and Rotterdam

[d] Only data from LASA, PREVEND and Rotterdam

Supplemental table 6. The mean difference in annual change in haemoglobin levels (continuous in g/dL) compared to euthyroidism depending on thyroid hormone status at baseline.

	Overt hypothyroidism		Subclinical hypothyroidism		Euthyroidism		Subclinical hyperthyroidism		Overt hyperthyroidism	
	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)	n	Mean diff. (95% CI)
All	459	-0.19 (-0.32- -0.06)	2,930	-0.07 (-0.13- -0.01)	36,081	(ref).	2,386	-0.06 (-0.13-0.01)	309	-0.07 (-0.20-0.05)
Sex										
Male	122	-0.31 (-0.52- -0.10)	1,029	-0.08 (-0.15-0.00)	17,546	(ref).	1,055	-0.06 (-0.13-0.02)	102	0.06 (-0.16-0.28)
Female	337	-0.18 (-0.34- -0.02)	1,901	-0.08 (-0.14- -0.02)	18,535	(ref).	1,331	-0.09 (-0.22-0.04)	204	-0.14 (-0.29-0.01)
Age										
< 50 years	47	-0.23 (-0.55-0.09)	308	-0.13 (-0.28-0.02)	6,902	(ref).	599	-0.01 (-0.09-0.08)	30	-0.01 (-0.64-0.62)
50-65 years	132	-0.24 (-0.42- -0.06)	677	-0.05 (-0.14-0.04)	9,715	(ref).	672	-0.12 (-0.25-0.01)	98	-0.06 (-0.26-0.15)
65-80 years	215	-0.21 (-0.45-0.03)	1,711	-0.10 (-0.17- -0.02)	16,814	(ref).	948	0.01 (-0.15-0.18)	186	-0.09 (-0.28-0.09)
> 80 years	65	-0.20 (-0.53-0.12)	234	0.01 (-0.18-0.21)	2,646	(ref).	168	-0.24 (-0.45- -0.03)	27	0.29 (-0.15-0.73)
Ethnicity										
White	431	-0.18 (-0.32- -0.05)	2,687	-0.09 (-0.16- -0.01)	34,154	(ref).	2,339	-0.06 (-0.13-0.01)	303	-0.08 (-0.21-0.04)
Black[b]	12	-0.19 (-0.88-0.50)	98	-0.20 (-0.45-0.05)	1,007	(ref).	35	0.11 (-0.31-0.52)	2	0.72 (-0.98-2.41)
Asian[c]	13	-0.53 (-1.22-0.16)	143	0.00 (-0.23-0.22)	800	(ref).	8	-0.68 (-1.55-0.19)	1	0.37 (-2.08-2.83)
Other[d]	-	-	-	-	13	(ref).	1	1.48 (-0.66-3.62)	-	-
Without medication	201	-0.15 (-0.32-0.01)	1,848	-0.08 (-0.17-0.01)	20,451	(ref).	1,625	-0.05 (-0.14-0.05)	138	0.08 (-0.10-0.26)

[a] Results were obtained by linear regression analysis, adjusted for age (if applicable) and sex

[b] Only data from CHS, HABC and PREVEND

[c] Only data from LASA, Nagasaki, PREVEND and Rotterdam

[d] Only data from LASA, PREVEND and Rotterdam

Supplemental table 7. Sensitivity analyses: The risk of having anaemia at baseline depending on thyroid hormone status.

		Overt hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt hyperthyroidism
Pooled OR (95% CI)[a]						
Without antithyroid medication or thyroid hormone replacement therapy[b]		1.95 (1.27-3.01)	1.28 (1.06-1.54)	1 (ref.)	1.29 (0.92-1.81)	1.83 (0.64-5.26)
	N	*204*	*1,845*	*20,451*	*1,664*	*143*
Without elevated CRP levels[c]		2.54 (1.31-4.93)	0.93 (0.60-1.45)	1 (ref.)	1.31 (1.02-1.66)	1.36 (0.60-3.09)
	N	*87*	*242*	*5,547*	*684*	*55*

Excluded the following medications: levothyroxine, thiamazol, ATC codes H03AA, H03CA, H05AA, H03B, H05B and by the cohort's expertise classified as antithyroid medication or thyroid hormone replacement therapy

[a] Results were obtained by logistic regression analysis, adjusted for age (if applicable) and sex

[b] No data on thyroid medication from EPIC-Norfolk, PISA and PREVEND

[c] Analysed in studies containing CRP records: Bari, Belfrail, InChianti, LASA, Leiden 85-plus, PREVEND, Rotterdam and SHIP. CRP >20 mg/L was considered elevated.

Supplemental table 8. Sensitivity analyses: The risk developing anaemia during follow-up depending on thyroid hormone status at baseline.

		Overt hypothyroidism	Subclinical hypothyroidism	Euthyroidism	Subclinical hyperthyroidism	Overt hyperthyroidism
Pooled OR (95% CI)[a]						
Without antithyroid medication or thyroid hormone replacement therapy[b]		2.00 (0.65-6.19)	1.13 (0.86-1.48)	1 (ref)	1.58 (1.16-2.16)	2.08 (0.86-5.03)
	N	*66*	*836*	*10,792*	*292*	*64*
Without elevated CRP levels[c]		1.72 (0.81-3.68)	1.70 (1.20-2.41)	1 (ref)	1.31 (1.01-1.69)	2.14 (0.67-6.79)
	N	*57*	*168*	*3,960*	*434*	*32*

Excluded the following medications: levothyroxine, thiamazol, ATC codes H03AA, H03CA, H05AA, H03B, H05B and by the cohort's expertise classified as antithyroid medication or thyroid hormone replacement therapy

[a] Results were obtained by logistic regression analysis, adjusted for age (if applicable) and sex

[b] No data on thyroid medication from EPIC-Norfolk, PISA and PREVEND

[c] Analysed in studies containing CRP records: Bari, Belfrail, InChianti, LASA, Leiden 85-plus, PREVEND, Rotterdam and SHIP. CRP >20 mg/L was considered elevated.

Supplemental table 9. Individual cohort estimates for the cross-sectional analysis.

Thyroid status	Odds ratio	95% CI		Weight	I^2
		Lower	Upper		
Overt hypothyroidism	1.84	1.35	2.50		10%
Bari study	NA				
BELFRAIL	NA				
Busselton Health study	4.97	1.07	23.05	3.9%	
Cardiovascular Health study	1.50	0.52	4.33	7.7%	
EPIC-Norfolk study	1.62	1.09	2.40	34.4%	
Health- Aging- and Body Composition study	0.92	0.27	3.12	5.9%	
InChianti study	2.67	0.53	13.54	3.5%	
Longitudinal Aging Study Amsterdam (LASA)	14.10	3.06	64.90	3.9%	
Leiden 85-plus Study	1.63	0.81	3.28	15.7%	
Nagasaki Adult Health study	2.01	0.63	6.39	6.5%	
Pisa cohort	1.52	0.13	17.06	1.6%	
PREVEND study	NA				
PROSPER study	0.43	0.06	3.21	2.3%	
Rotterdam Study	3.12	0.88	11.10	5.5%	
SHIP	1.72	0.55	5.33	6.8%	
Whickham Survey	1.35	0.17	10.79	2.2%	
Subclinical hypothyroidism	1.21	1.02	1.43		28%
Bari study	1.37	0.64	2.95	4.0%	
BELFRAIL	1.19	0.12	11.86	0.5%	
Busselton Health study	1.72	0.59	4.99	2.2%	
Cardiovascular Health study	1.80	1.33	2.44	13.9%	
EPIC-Norfolk study	1.06	0.81	1.37	15.9%	
Health- Aging- and Body Composition study	0.98	0.70	1.37	12.7%	
InChianti study	1.02	0.34	3.05	2.1%	
Longitudinal Aging Study Amsterdam (LASA)	1.73	0.21	14.46	0.6%	
Leiden 85-plus Study	1.08	0.48	2.40	3.7%	
Nagasaki Adult Health study	0.87	0.54	1.39	8.4%	
Pisa cohort	1.38	0.90	2.11	9.5%	
PREVEND study	0.47	0.17	1.35	2.3%	
PROSPER study	1.48	1.06	2.07	12.7%	
Rotterdam Study	1.56	0.75	3.23	4.3%	
SHIP	2.20	0.69	7.00	1.9%	
Whickham Survey	0.69	0.36	1.35	5.1%	

Supplemental table 9. Individual cohort estimates for the cross-sectional analysis. (*continued*)

Thyroid status	Odds ratio	95% CI		Weight	I^2
		Lower	Upper		
Subclinical hyperthyroidism	1.27	1.03	1.57		35%
Bari study	1.71	0.31	9.42	1.4%	
BELFRAIL	1.12	0.55	2.29	6.3%	
Busselton Health study	3.13	1.20	8.17	4.0%	
Cardiovascular Health study	0.24	0.03	1.75	1.1%	
EPIC-Norfolk study	1.07	0.74	1.54	13.5%	
Health- Aging- and Body Composition study	0.72	0.36	1.47	6.4%	
InChianti study	2.01	1.11	3.64	8.2%	
Longitudinal Aging Study Amsterdam (LASA)	0.92	0.21	4.04	1.9%	
Leiden 85-plus Study	2.13	0.89	5.09	4.7%	
Nagasaki Adult Health study	0.82	0.10	6.83	1.0%	
Pisa cohort	1.63	1.06	2.51	11.7%	
PREVEND study	0.82	0.28	2.38	3.3%	
PROSPER study	0.75	0.34	1.61	5.6%	
Rotterdam Study	0.54	0.19	1.51	3.6%	
SHIP	1.24	1.01	1.51	18.8%	
Whickham Survey	2.11	1.20	3.71	8.6%	
Overt hyperthyroidism	1.69	1.00	2.87		34%
Bari study	NA				
BELFRAIL	NA				
Busselton Health study	1.95	0.25	15.41	5.6%	
Cardiovascular Health study	12.82	0.74	222.89	3.1%	
EPIC-Norfolk study	2.03	1.11	3.70	26.4%	
Health- Aging- and Body Composition study	NA				
InChianti study	3.18	0.95	10.71	12.9%	
Longitudinal Aging Study Amsterdam (LASA)	NA				
Leiden 85-plus Study	NA				
Nagasaki Adult Health study	NA				
Pisa cohort	1.66	0.30	9.14	7.8%	
PREVEND study	NA				
PROSPER study	0.40	0.12	1.26	13.7%	
Rotterdam Study	3.39	0.72	16.07	8.9%	
SHIP	1.27	0.59	2.75	21.6%	
Whickham Survey	NA				

Abbreviations: CI, Confidence interval; NA, data not available.

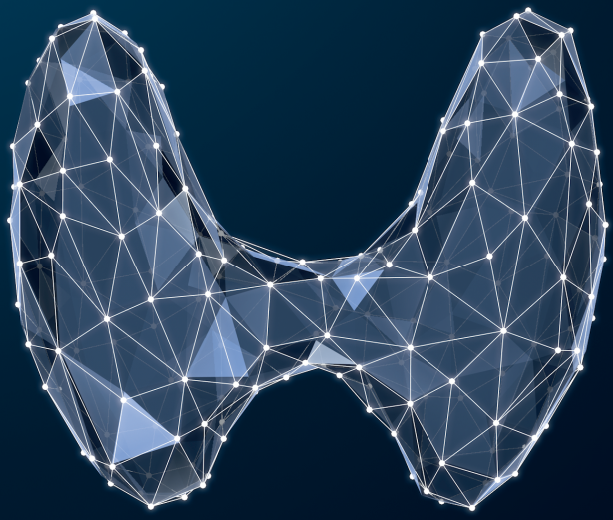
Supplemental table 10. Individual cohort estimates for the longitudinal analysis.

Thyroid status	Hazard ratio	95% CI		Weight	I^2
		Lower	Upper		
Overt hypothyroidism	1.38	0.86	2.20		33%
Bari study	NA				
BELFRAIL	0.96	0.13	7.02	4.7%	
Busselton Health study	4.84	1.17	20.04	8.0%	
Cardiovascular Health study	1.00	0.37	2.69	12.8%	
EPIC-Norfolk study	0.44	0.16	1.18	12.8%	
Health- Aging- and Body Composition study	0.72	0.10	5.17	4.8%	
InChianti study	0.96	0.13	6.85	4.8%	
Longitudinal Aging Study Amsterdam (LASA)	NA				
Leiden 85-plus Study	1.20	0.58	2.49	17.2%	
Nagasaki Adult Health study	1.68	0.53	5.29	10.7%	
Pisa cohort	NA				
PREVEND study	7.29	1.79	29.79	8.1%	
PROSPER study	1.81	0.45	7.33	8.2%	
Rotterdam Study	NA				
SHIP	1.42	0.35	5.75	8.1%	
Whickham Survey	NA				
Subclinical hypothyroidism	1.18	1.00	1.38		17%
Bari study	1.51	0.52	4.41	2.4%	
BELFRAIL	0.76	0.19	3.16	1.4%	
Busselton Health study	NA				
Cardiovascular Health study	0.93	0.68	1.28	19.0%	
EPIC-Norfolk study	1.19	0.84	1.69	16.3%	
Health- Aging- and Body Composition study	1.14	0.77	1.70	13.7%	
InChianti study	2.24	1.17	4.28	6.1%	
Longitudinal Aging Study Amsterdam (LASA)	NA				
Leiden 85-plus Study	0.70	0.34	1.46	4.9%	
Nagasaki Adult Health study	1.03	0.70	1.51	14.3%	
Pisa cohort	NA				
PREVEND study	2.16	1.10	4.25	5.6%	
PROSPER study	1.26	0.77	2.05	9.8%	
Rotterdam Study	1.29	0.70	2.39	6.7%	
SHIP	NA				
Whickham Survey	NA				

Supplemental table 10. Individual cohort estimates for the longitudinal analysis. (*continued*)

Thyroid status	Hazard ratio	95% CI		Weight	I^2
		Lower	Upper		
Subclinical hyperthyroidism	1.15	0.94	1.42		10%
Bari study	2.40	0.31	18.49	1.0%	
BELFRAIL	1.31	0.59	2.93	6.0%	
Busselton Health study	1.91	0.46	7.87	2.1%	
Cardiovascular Health study	1.61	0.66	3.92	5.0%	
EPIC-Norfolk study	0.52	0.26	1.05	7.8%	
Health- Aging- and Body Composition study	1.19	0.55	2.55	6.7%	
InChianti study	1.84	1.13	3.00	14.4%	
Longitudinal Aging Study Amsterdam (LASA)	1.65	0.38	7.07	1.9%	
Leiden 85-plus Study	1.44	0.58	3.56	4.9%	
Nagasaki Adult Health study	NA				
Pisa cohort	NA				
PREVEND study	1.02	0.32	3.25	3.0%	
PROSPER study	1.14	0.47	2.79	5.0%	
Rotterdam Study	1.30	0.70	2.40	9.8%	
SHIP	0.91	0.69	1.21	32.4%	
Whickham Survey	NA				
Overt hyperthyroidism	1.47	0.91	2.38		0%
Bari study	NA				
BELFRAIL	NA				
Busselton Health study	NA				
Cardiovascular Health study	NA				
EPIC-Norfolk study	1.43	0.53	3.84	23.8%	
Health- Aging- and Body Composition study	2.32	0.32	16.59	6.0%	
InChianti study	1.21	0.17	8.69	6.0%	
Longitudinal Aging Study Amsterdam (LASA)	NA				
Leiden 85-plus Study	3.59	0.49	26.06	5.9%	
Nagasaki Adult Health study	NA				
Pisa cohort	NA				
PREVEND study	NA				
PROSPER study	1.13	0.42	3.06	23.3%	
Rotterdam Study	2.91	0.72	11.76	11.8%	
SHIP	1.03	0.38	2.79	23.3%	
Whickham Survey	NA				

Abbreviations: CI, Confidence interval; NA, data not available.



Chapter 5

Thyroid hormone therapy for older adults with subclinical hypothyroidism

David J Stott
Nicolas Rodondi
Patricia M Kearney
Ian Ford
Rudi GJ Westendorp
Simon P Mooijaart
Naveed Sattar
Carole E Aubert
Drahomir A Aujesky
Doug C Bauer
Christine Baumgartner
Manuel R Blum
John P Browne
Stephen Byrne
Tinh-Hai Collet
Olaf M Dekkers
Wendy PJ den Elzen
Robert S Du Puy
Graham Ellis
Martin Feller
Carmen Floriani
Kirsty Hendry
Caroline Hurley
J Wouter Jukema
Sharon Kean
Maria Kelly
Danielle Krebs
Peter Langhorne
Gemma McCarthy
Vera McCarthy
Alex McConnachie
Mairi McDade
Martina Messow
Annemarie O'Flynn
David O'Riordan
Rosalinde KE Poortvliet
Terry J Quinn
Audrey Russell
Carol Sinnott
Jan WA Smit
H Anette Van Dorland
Kieran A Walsh
Elaine K Walsh
Torquil Watt
Robbie Wilson
Jacobijn Gussekloo for the
TRUST Study Group

New England Journal of Medicine, 2017;376(26):2534-2544

DOI: 10.1056/NEJMoa1603825

ABSTRACT

Background

The use of levothyroxine to treat subclinical hypothyroidism is controversial. We aimed to determine whether levothyroxine provided clinical benefits in older persons with this condition.

Methods

We conducted a double-blind, randomised, placebo-controlled, parallel-group trial involving 737 adults who were at least 65 years of age and who had persisting subclinical hypothyroidism (thyrotropin level, 4.60 to 19.99 mIU per litre; free thyroxine level within the reference range). A total of 368 patients were assigned to receive levothyroxine (at a starting dose of 50 µg daily, or 25 µg if the body weight was < 50 kg or the patient had coronary heart disease), with dose adjustment according to the thyrotropin level; 369 patients were assigned to receive placebo with mock dose adjustment. The two primary outcomes were the change in the Hypothyroid Symptoms score and Tiredness score on a thyroid-related quality-of-life questionnaire at 1 year (range of each scale is 0 to 100, with higher scores indicating more symptoms or tiredness, respectively; minimum clinically important difference, 9 points).

Results

The mean age of the patients was 74.4 years, and 396 patients (53.7%) were women. The mean (±SD) thyrotropin level was 6.40±2.01 mIU per litre at baseline; at 1 year, this level had decreased to 5.48 mIU per litre in the placebo group, as compared with 3.63 mIU per litre in the levothyroxine group ($p < 0.001$), at a median dose of 50 µg. We found no differences in the mean change at 1 year in the Hypothyroid Symptoms score (0.2±15.3 in the placebo group and 0.2±14.4 in the levothyroxine group; between-group difference, 0.0; 95% confidence interval [CI], −2.0 to 2.1) or the Tiredness score (3.2±17.7 and 3.8±18.4, respectively; between-group difference, 0.4; 95% CI, −2.1 to 2.9). No beneficial effects of levothyroxine were seen on secondary-outcome measures. There was no significant excess of serious adverse events prespecified as being of special interest.

Conclusions

Levothyroxine provided no apparent benefits in older persons with subclinical hypothyroidism. (Funded by European Union FP7 and others; TRUST ClinicalTrials.gov number, NCT01660126.)

INTRODUCTION

Subclinical hypothyroidism is defined as an elevated serum thyrotropin level and a serum free thyroxine level within the reference range.[1] Between 8% and 18% of adults 65 years of age or older have these biochemical features, and the prevalence is higher among women than among men.[2]

Subclinical hypothyroidism is a possible contributor to many problems in older persons. Thyroid hormones have multiple effects, since they act as an essential regulatory factor in numerous physiological systems, including the vascular tree and the heart,[3] the brain (including cognition),[4] skeletal muscle, and bone.[5] Tiredness is the most important symptom of overt hypothyroidism,[6] but most patients with subclinical hypothyroidism have no symptoms or have nonspecific symptoms.[7] There is a convincing epidemiologic association with subsequent coronary heart disease.[8]

Randomised, controlled trials of levothyroxine replacement for the treatment of subclinical hypothyroidism have been small [9,10] and have yielded only limited evidence regarding the possible benefits and risks of treatment.[1] We aimed to determine whether there are clinical benefits from levothyroxine replacement in older persons with subclinical hypothyroidism.

METHODS

Trial overview

The trial protocol, which was published previously [11] and is available with the full text of this article at NEJM.org, was approved by the relevant ethics committees and regulatory authorities in all the countries involved in the trial. Participants provided written informed consent.

The trial was conducted in accordance with the principles of the Declaration of Helsinki [12] and Good Clinical Practice guidelines.[13] The Robertson Centre for Biostatistics at the University of Glasgow was the trial data and biostatistics centre.

The European Union FP7 provided primary financial support for the conduct of the trial. Supplies of levothyroxine and matching placebo were provided free of charge by Merck (Darmstadt, Germany). The funder, the trial sponsors (NHS Greater Glasgow and Clyde Health Board and University of Glasgow, United Kingdom; University College Cork, Ireland; Leiden University Medical Center, the Netherlands; and University of Bern and Bern University Hospital, Switzerland), and Merck played no role in the design, analysis, or reporting of the trial.

The main sponsor (NHS Greater Glasgow and Clyde Health Board) contributed to the writing of the protocol. None of the sponsors had any involvement in the analysis or the reporting of the results. The authors vouch for the accuracy and completeness of the data and analyses reported and for the fidelity of the trial to the protocol.

Participants

Participants were identified from clinical laboratory and general practice databases and records. The inclusion criteria were an age of 65 years or more and persistent subclinical hypothyroidism, defined as an elevated thyrotropin level (4.60 to 19.99 mIU per litre) that was measured on at least two occasions that were 3 months to 3 years apart, with the free thyroxine level within the reference range. The main exclusion criteria for the trial were a current prescription for levothyroxine, antithyroid drugs, amiodarone, or lithium; thyroid surgery or receipt of radioactive iodine within the previous 12 months; dementia; hospitalization for a major illness or an elective surgery within the previous 4 weeks; an acute coronary syndrome (including myocardial infarction or unstable angina) within the previous 4 weeks; and terminal illness.[11]

Trial design and regimen

We conducted a randomised, double-blind, parallel-group trial of levothyroxine versus placebo. Patients underwent randomization in a 1:1 ratio, with stratification according to country, sex, and starting dose, with the use of randomly permuted blocks.

The active intervention started with levothyroxine at a dose of 50 µg daily (or 25 µg in patients with a body weight of < 50 kg or with known coronary heart disease [previous myocardial infarction or symptoms of angina pectoris]) or matching placebo. Dose adjustment in the levothyroxine group was aimed to result in a thyrotropin level within the reference range (0.40 to 4.59 mIU per litre). Details regarding how the dose was adjusted and the mock adjustment in the placebo group are provided in the Supplementary Appendix, available at NEJM.org. All dose adjustments were generated and executed by means of computer without the intervention of a physician. The participants, investigators, and treating physicians were unaware of the results of thyrotropin measurements throughout the course of the trial.

Procedures and outcomes

The two primary outcomes for the trial were the change from baseline to 12 months in the Thyroid-Related Quality-of-Life Patient-Reported Outcome measure (ThyPRO) Hypothyroid Symptoms score (4 items) and Tiredness score (7 items); each scale ranges from 0 to 100, with higher scores indicating more symptoms and tiredness, respectively.[14] A recent systematic review recommended ThyPRO as the preferred measurement tool for the assessment of health-related quality of life in patients with benign thyroid disease.[15] The ThyPRO and

other instruments were administered in English, French, German, or Dutch as appropriate. We had initially planned for cardiovascular events and thyroid-specific quality of life to be the two primary outcomes. However, this plan was modified during the trial to thyroid-specific quality-of-life scores as the two primary outcomes and cardiovascular events as a secondary outcome when it became apparent that the trial would be underpowered for cardiovascular events owing to delays and difficulties in recruitment.[11]

The secondary outcomes included changes from baseline in generic health-related quality of life (as assessed by the EuroQoL [EQ] Group 5-Dimension Self-Report Questionnaire [EQ-5D]; scores on the EQ-5D descriptive index range from −0.59 to 1.00, and scores on the EQ visual-analogue scale range from 0 to 100, with higher scores indicating better quality of life),[16] comprehensive thyroid-related quality of life (as assessed by the ThyPRO-39 score, a shorter version of the ThyPRO measure,[17] at final follow-up only), hand-grip strength (as assessed by means of the Jamar isometric dynamometer, with the recorded score as the best of three measures in the dominant hand),[18] executive cognitive function (as assessed with the letter–digit coding test, which indicates the speed of processing according to the number of correct responses in matching nine letters with nine digits in 90 seconds; minimum score, 0, with higher scores indicating better executive cognitive function; there is no maximum score),[19] blood pressure (systolic and diastolic), weight, body-mass index, waist circumference, activities of daily living (as assessed by the Barthel Index of functional levels in activities of daily living, on a scale ranging from 0 to 20, with higher scores indicating better performance),[20] the Instrumental Activities of Daily Living score (on a scale from 0 to 14, with higher scores indicating better performance in activities of daily living),[21] and fatal and nonfatal cardiovascular events. The minimum follow-up was 1 year, and the maximum follow-up was 3 years.

Safety and recording of adverse events

Adverse events were assessed, managed, recorded, reported, and analysed in accordance with the Medicines for Human Use (Clinical Trials) Regulations 2004 (as amended). Adverse events of special interest included new atrial fibrillation, heart failure, fracture, and new diagnosis of osteoporosis. The score on the ThyPRO Hyperthyroid Symptoms scale was recorded as a measure of possible adverse effects (on a scale from 0 to 100, with higher scores indicating more symptoms; minimum clinically important difference has been estimated as 9 points).[14]

Statistical analysis

The Hypothyroid Symptoms and Tiredness scores from the ThyPRO14 were the two primary outcomes, with the required p value for statistical significance split equally to each test (0.05/2=0.025 for each test). We assumed standard deviations for data at 1 year of 13.3 and

18.3 on the 100-point scales, respectively, after adjustment for baseline values. These calculations provided the trial with 80% power to detect a change with levothyroxine treatment (vs. placebo) of 3.0 points on the Hypothyroid Symptoms score and 4.1 points on the Tiredness score with our revised maximum expected number of recruited participants of 750, and with changes of 3.5 points and 4.9 points, respectively, with our minimum expected number of 540 participants. Justification for these power calculations is provided in the trial protocol.[11]

The methods of analysis of the continuous efficacy outcomes involving measurements at baseline and follow-up were analysed at each time point for the comparison of the two trial groups, with adjustment for stratification variables (country, sex, and starting dose of levothyroxine) and baseline levels of the same variable with the use of multivariate linear regression (see the Supplementary Appendix). The efficacy and safety analyses were carried out in a modified intention-to-treat population, which included participants with data on the outcome of interest. Patients who discontinued the trial regimen continued to be followed for the modified intention-to-treat analysis. These analyses were supported with sensitivity analyses that used mixed-effects models and multiple imputations for missing data. The primary and secondary outcomes at 12 months were also analysed in prespecified subgroups according to sex and baseline thyrotropin level.[11] Analyses were repeated in the per-protocol population, which included participants who continued to take the trial regimen per the trial protocol.

RESULTS

We screened 2,647 community-dwelling persons who were at least 65 years of age and who were identified as having biochemical subclinical hypothyroidism. A total of 737 participants underwent randomization, 369 of whom were assigned to receive placebo and 368 to receive levothyroxine (Figure 1). The characteristics at baseline were similar in the two groups (Table 1, and Supplemental table 1). The mean age of the patients was 74.4 years, and 396 patients (53.7%) were women. A score of 0 (indicating no symptoms) at baseline was observed in 199 of 737 participants (27.0%) on the Hypothyroid Symptoms scale and in 64 (8.7%) on the Tiredness scale; 36 participants (4.9%) had a score of 0 in both domains.

A total of 337 participants (91.3%) who were randomly assigned to the placebo group completed 12-month follow-up, as did 332 (90.2%) in the levothyroxine group. The median follow-up for all the participants who underwent randomization (including participants who discontinued the trial regimen) was 17.3 months (interquartile range, 12.0 to 24.4) in the placebo group and 18.0 months (interquartile range, 11.0 to 25.4) in the levothyroxine group. The median dose of levothyroxine at 1 year was 50 µg. The numbers of patients who were included in the analyses are presented in Figure 1.

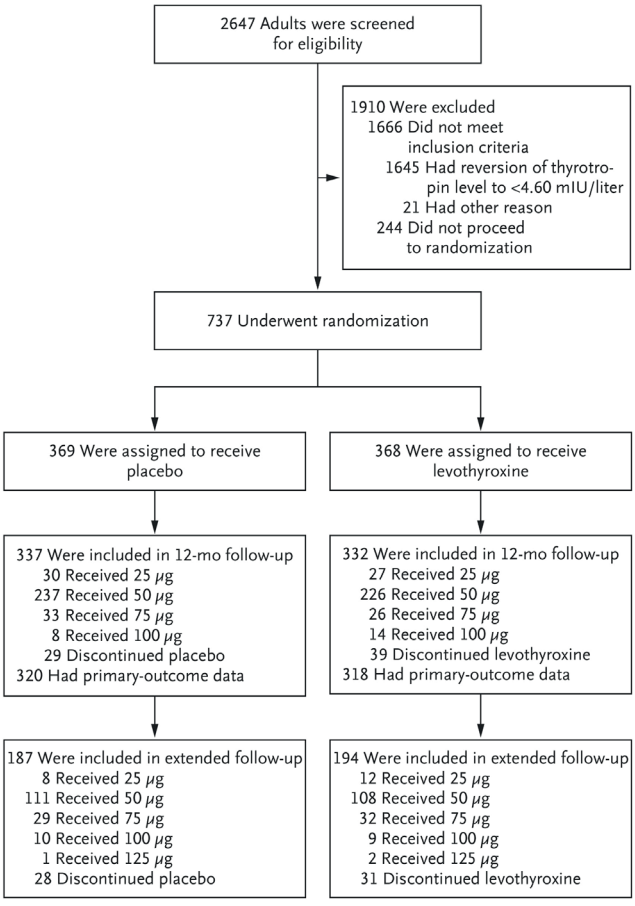


Figure 1. Randomization, Follow-up, and Dose Levels. Exclusions for other reasons included use of antithyroid medication (in 17 persons), recent thyroid surgery (in 1), recent acute coronary syndrome (in 1), current participation in another trial (in 1), and adrenal insufficiency (in 1). Two patients who were excluded because the thyrotropin level reverted to less than 4.60 mIU per litre also had an additional exclusion of galactose intolerance. Extended follow-up beyond 12 months was conducted in a subgroup of patients, with a median duration of follow-up from baseline of 24.2 months (interquartile range, 18.4 to 30.3) in the placebo group and 24.5 months (interquartile range, 18.4 to 30.5) in the levothyroxine group.

Table 1. Characteristics of the participants at baseline.

Characteristic[a]	Placebo Group (n = 369)	Levothyroxine Group (n = 368)
Age — yr		
Mean	74.8±6.8	74.0±5.8
Range	65.1–93.4	65.2–93.0
Female sex — no. (%)	198 (53.7)	198 (53.8)
White race — no. (%)[b]	362 (98.1)	362 (98.4)
Standard housing — no. (%)[c]	356 (96.5)	358 (97.3)
Previous medical conditions and clinical descriptors — no./total no.(%)		
Ischemic heart disease[d]	50/369 (13.6)	50/368 (13.6)
Atrial fibrillation	44/368 (12.0)	45/364 (12.4)
Hypertension	183/366 (50.0)	192/368 (52.2)
Diabetes mellitus	54/368 (14.7)	63/368 (17.1)
Osteoporosis	47/367 (12.8)	41/364 (11.3)
Current smoking	33/369 (8.9)	29/368 (7.9)
Median no. of concomitant medicines (IQR)	4 (2–6)	4 (2–6)
Median Mini–Mental State Examination score (IQR)[e]	29 (28–30)	29 (27–30)
Weight < 50 kg — no. (%)	5 (1.4)	5 (1.4)
Laboratory results		
Thyrotropin — mIU/litre	6.38±2.01	6.41±2.01
Median (IQR)	5.76 (5.10–6.94)	5.73 (5.12–6.83)
Range	4.60–17.60	4.60–17.60
Free thyroxine — pmol/litre[f]	13.3±1.9	13.4±2.1
Outcome measures[g]		
Hypothyroid Symptoms score	16.9±17.9	17.5±18.8
Tiredness score	25.5±20.3	25.9±20.6
EQ-5D descriptive index	0.847±0.171	0.846±0.187
EQ visual-analogue scale score	76.5±16.3	78.4±15.3
Hand-grip strength — kg	27.5±11.3	28.0±10.2
Letter–digit coding test score	25.2±8.3	24.9±7.4
Blood pressure — mm Hg		
Systolic	140.4±18.9	141.2±18.7
Diastolic	74.8±11.7	74.1±11.6
Body-mass index	27.7±4.6	28.1±5.3
Waist circumference — cm	97.5±12.8	98.5±13.6
Median Barthel Index (IQR)	20 (14–20)	20 (13–20)
Median Instrumental Activities of Daily Living score (IQR)	14 (7–14)	14 (7–14)

[a] Plus–minus values are means ±SD. There were no significant between-group differences in the baseline characteristics. IQR denotes interquartile range.

[b] Race was reported by the patient.

[c] Standard housing was defined as nonsheltered community accommodation. By contrast, sheltered housing is purpose-built grouped housing for older persons, often with an on-site manager or warden.

[d] Ischemic heart disease was defined as a history of angina pectoris or previous myocardial infarction.

[e] The Mini–Mental State Examination score is on a scale from 0 to 30, with higher scores indicating better cognitive function.

[f] To convert the values for free thyroxine to nanograms per decilitre, divide by 12.87.

[g] The Hypothyroid Symptoms score and the Tiredness score from the Thyroid-Related Quality of Life Patient-Reported Outcome (ThyPRO) questionnaire are each assessed on a scale from 0 to 100, with higher scores indicating more symptoms and tiredness, respectively. The minimum clinically important difference for each score has been estimated as 9 points. The EuroQoL [EQ] Group 5-Dimension Self-Report Questionnaire (EQ-5D) scores included both the EQ- 5D descriptive index (on a scale from −0.59 to 1.00) and the score on the EQ visual-analogue scale (on a scale from 0 to 100); higher scores on each scale indicate better quality of life. The score on the letter–digit coding test (a test of executive cognitive function) indicates the speed of processing according to the number of correct responses in matching nine letters with nine digits in 90 seconds (minimum score is 0, with higher scores indicating better executive cognitive function; there is no maximum score). The body-mass index is the weight in kilograms divided by the square of the height in meters. The Barthel Index uses a scale from 0 to 20 points, with higher numbers indicating better performance on activities of daily living. The Instrumental Activities of Daily Living scale has a maximum score of 14 (range, 0 to 14), with higher scores indicating better performance in activities of daily living.

Thyroid-function tests

The mean (±SD) thyrotropin level at baseline was 6.40±2.01 mIU per litre. The thyrotropin levels were reduced from baseline to a greater extent in the levothyroxine group than in the placebo group at all time points of review, with a mean between-group difference of 2.29 mIU per litre at 6 to 8 weeks after randomization ($p < 0.001$) (Supplemental table 2). At 12 months, the mean thyrotropin level was 5.48±2.48 mIU per litre in the placebo group, as compared with 3.63±2.11 mIU per litre in the levothyroxine group, resulting in a between-group difference of 1.92 mIU per litre ($p < 0.001$) (Table 2 and Figure 2). There was a significant interaction between the trial group and the office visit ($p=0.03$), with a reduction in the thyrotropin level being the greatest at 6 to 8 weeks.

Free thyroxine levels were not routinely measured, although the data were available in a subgroup of patients. The mean free thyroxine level was 2.3 pmol per litre (0.2 ng per decilitre) higher in the levothyroxine group than in the placebo group both at 6 to 8 weeks and at 12 months ($p < 0.001$ for both comparisons) (Supplemental table 3).

Table 2. Outcomes at 12 months and extended follow-up.[a]

Variable	Baseline		At 12 Mo				At Extended Follow-up Visit[b]			
	Placebo (N = 369)	Levothyroxine (N = 368)	Placebo (N = 320)	Levothyroxine (N = 318)	Difference (95% CI)	P Value	Placebo (N = 187)	Levothyroxine (N = 194)	Difference (95% CI)	P Value
Thyrotropin — mIU/litre	6.38±2.01	6.41±2.01	5.48±2.48	3.63±2.11	−1.92 (−2.24 to −1.59)	<0.001	5.28±2.50	3.47±2.08	−1.88 (−2.32 to −1.45)	<0.001
Median (IQR)	5.76 (5.10 to 6.94)	5.70 (5.12 to 6.83)	4.90 (3.91 to 6.46)	3.16 (2.45 to 4.22)	—	—	4.94 (3.78 to 6.26)	3.00 (2.26 to 4.16)	—	—
Primary outcomes[c]										
Hypothyroid Symptoms score	16.9±17.9	17.5±18.8	16.7±17.5	16.6±16.9	0.0 (−2.0 to 2.1)	0.99	15.2±15.9	17.9±9.1	1.0 (−1.9 to 3.9)	0.50
Tiredness score	25.5±20.3	25.9±20.6	28.6±19.5	28.7±20.2	0.4 (−2.1 to 2.9)	0.77	31.9±22.1	30.2±20.5	−3.5 (−7.0 to 0.0)	0.05
Secondary outcomes										
EQ-5D descriptive score	0.847±0.171	0.846±0.187	0.853±0.191	0.833±0.212	−0.025 (−0.050 to 0.000)	0.05	0.829±0.209	0.864±0.188	0.040 (0.005 to 0.075)	0.03
EQ VAS score	76.5±16.3	78.4±15.3	77.4±13.7	77.3±15.6	−1.3 (−3.2 to 0.6)	0.18	77.2±13.5	76.8±14.2	−0.8 (−3.2 to 1.7)	0.56
Hand-grip strength — kg	27.5±11.3	28.0±10.2	27.1±11.2	27.5±10.5	−0.1 (−0.9 to 0.7)	0.84	24.9±10.6	24.4±10.1	−0.6 (−1.7 to 0.6)	0.34
Blood pressure — mm Hg										
Systolic	140.4±18.9	141.2±18.7	138.4±17.8	138.3±18.7	0.1 (−2.1 to 2.4)	0.90	137.5±19.2	136.8±17.6	1.1 (−4.1 to 2.1)	0.51
Diastolic	74.8±11.7	74.1±11.6	73.5±11.1	72.8±11.4	−0.1 (−1.5 to 1.3)	0.93	72.3±11.4	72.0±11.5	0.5 (−1.4 to 2.4)	0.59
Body-mass index	27.7±4.6	28.1±5.3	27.7±4.6	27.9±5.1	0.0 (−0.2 to 0.2)	0.89	27.2±4.5	27.9±4.9	0.2 (−0.1 to 0.5)	0.30

Table 2. Outcomes at 12 months and extended follow-up.[a] (*continued*)

Variable	Baseline		At 12 Mo				At Extended Follow-up Visit[b]			
	Placebo (N = 369)	Levothyroxine (N = 368)	Placebo (N = 320)	Levothyroxine (N = 318)	Difference (95% CI)	P Value	Placebo (N = 187)	Levothyroxine (N = 194)	Difference (95% CI)	P Value
Waist circumference — cm	97.5±12.8	98.5±13.6	96.8±13.1	98.0±13.2	0.4 (−0.4 to 1.3)	0.34	96.0±13.8	97.6±13.4	0.3 (−0.9 to 1.5)	0.66
Adverse symptom assessment										
Hyperthyroid Symptoms score[d]	10.5±11.2	10.5±11.2	10.3±11.3	10.5±10.8	0.6 (−0.7 to 1.9)	0.35	9.8±11.0	11.1±11.7	0.7 (−1.2 to 2.5)	0.46

[a] All these analyses were conducted in the modified intention-to-treat population, which included all the participants who had undergone randomization (excluding those who had under- gone randomization in error) for whom data was available on the outcome of interest. For analyses at the 12-month visit to be valid, they must have been conducted at 12 months (within a ±31-day window) after randomization. Results at 12 months, at the extended follow-up visit, and between-group differences are adjusted for stratification variables (country, sex, and starting dose of levothyroxine) and baseline levels of the same variable with the use of linear regression. Between-group differences are the value in the levothyroxine group minus the value in the placebo group. Data for the extended follow-up visit were additionally adjusted for time to visit. Data were missing for the following outcomes: for the thyrotropin level at 12 months for 7 patients in the placebo group and 1 in the levothyroxine group and at extended follow-up for 7 in the placebo group and 6 in the levothyroxine group; for the EQ-5D score at extended follow-up for 1 in the levothyroxine group; for the EQ visual-analogue scale (VAS) score at 12 months for 1 in the placebo group and at extended follow-up for 1 in the levothyroxine group; for hand-grip strength at baseline for 11 in the placebo group and 10 in the levothyroxine group, at 12 months for 22 in the placebo group and 16 in the levothyroxine group, and at extended follow-up for 14 in the placebo group and 8 in the levothyroxine group; for blood pressure at baseline and at 12 months for 1 in the placebo group and at extended follow-up for 5 in each group; on body-mass index at baseline for 1 in each group, at 12 months for 2 in the placebo group and 1 in the levothyroxine group, and at extended follow-up for 2 in the placebo group and 4 in the levothyroxine group; and for waist circumference at baseline for 1 in each group, at 12 months for 1 in the placebo group and 2 in the levothyroxine group, and at extended follow-up for 2 in the placebo group and 4 in the levothyroxine group. CI denotes confidence interval.

[b] Extended follow-up beyond 12 months was performed in a subgroup of patients. The median duration of follow-up from baseline was 24.2 months (interquartile range, 18.4 to 30.3) in the placebo group and 24.5 months (interquartile range, 18.4 to 30.5) in the levothyroxine group.

[c] The two primary outcomes were the Hypothyroid Symptoms score and the Tiredness score from the ThyPRO questionnaire at 12 months (adjusted as stated above). The range of each scale is 0 to 100, with higher scores indicating more symptoms. The minimum clinically important difference for each score has been estimated as 9 points.

[d] The score on the Hyperthyroid Symptoms scale was recorded as a measure of possible adverse effects (on a scale from 0 to 100, with higher scores indicating more symptoms; minimum clinically important difference has been estimated as 9 points).

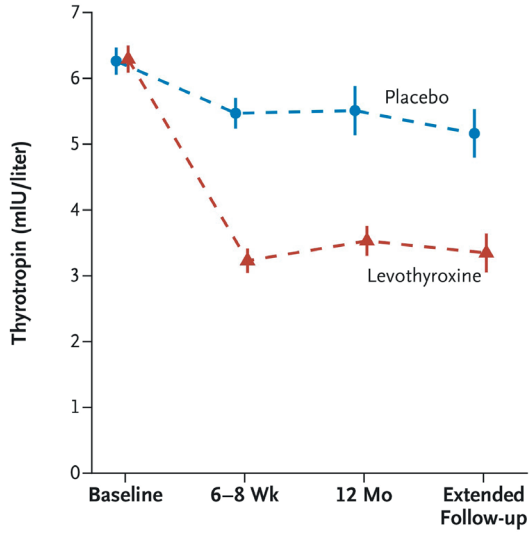


Figure 2. Thyrotropin levels in the Placebo Group and Levothyroxine Group. Shown are the results of a modified intention-to-treat analysis. Data are means, and error bars indicate 95% confidence intervals. Extended follow-up beyond 12 months was conducted in a subgroup of patients, with a median duration of follow-up from baseline of 24.2 months (interquartile range, 18.4 to 30.3) in the placebo group and 24.5 months (interquartile range, 18.4 to 30.5) in the levothyroxine group. $p < 0.001$ for between-group differences in the thyrotropin level at 6 to 8 weeks, 12 months, and extended follow-up. Analyses were adjusted for stratification variables (country, sex, and starting dose of levothyroxine) and baseline thyrotropin level with the use of linear regression; data for the extended follow-up visit were additionally adjusted for time to visit.

Thyroid-specific quality of life

The mean Hypothyroid Symptoms score at 12 months (with adjustment for baseline score) was 16.7±17.5 in the placebo group and 16.6±16.9 in the levothyroxine group (p=0.99). The mean Tiredness score was 28.6±19.5 in the placebo group and 28.7±20.2 in the levothyroxine group (p=0.77). We found no differences in the mean change at 1 year in the Hypothyroid Symptoms score (0.2±15.3 in the placebo group and 0.2±14.4 in the levothyroxine group) or the Tiredness score (3.2±17.7 and 3.8±18.4, respectively) (Table 2). There were no significant between-group differences in either of these measures at 6 to 8 weeks (Supplemental table 4). There was a small-magnitude between-group difference in the Tiredness score, with a lower value in the levothyroxine group than in the placebo group (difference, −3.49; p=0.05) at the extended follow-up review (Table 2). Prespecified analyses according to sex and baseline thyrotropin level did not reveal any subgroups of patients who benefited from treatment with levothyroxine. Per-protocol analyses and sensitivity analyses with the use of multiple imputation of missing values showed no significant differences between the levothyroxine group and the placebo group (Supplemental tables 4 and 5).

Other outcome measures

The EQ-5D descriptive index showed a small deterioration at 12 months (mean difference between the levothyroxine group and the placebo group, −0.025; p=0.05) but a minor improvement at extended follow-up (mean difference, 0.040; p=0.03); there were no significant between-group differences at 6 to 8 weeks. There were no significant between-group differences in the score on the EQ visual-analogue scale (Table 2, and Supplemental table 2).

No significant effects were seen in any of the other secondary-outcome measures, either in the modified intention-to-treat or per-protocol analyses or in the prespecified subgroups (Table 2, and Supplemental tables 4, 6, 7, and 8). Results regarding cardiovascular events and total and cardiovascular mortality are provided in Table 3 and in Supplemental figures 1 and 2.

Table 3. Clinical outcomes and adverse events.[a]

Variable	All Patients (N = 737)	Placebo Group (N = 369)	Levothyroxine Group (N = 368)	Hazard Ratio (95% CI)
Clinical outcome				
Fatal or nonfatal cardiovascular event — no. (%)	38 (5.2)	20 (5.4)	18 (4.9)	0.89 (0.47–1.69)
Cardiovascular death — no. (%)	3 (0.4)	1 (0.3)	2 (0.5)	—
Death from any cause — no. (%)	15 (2.0)	5 (1.4)	10 (2.7)	1.91 (0.65–5.60)
Serious adverse event				
No. of patients with ≥ 1 serious adverse event	181 (24.6)	103 (27.9)	78 (21.2)	0.94 (0.88–1.00)[b]
No. of events	343	201	142	—
Adverse event of special interest				
New-onset atrial fibrillation — no. (%)	24 (3.3)	13 (3.5)	11 (3.0)	0.80 (0.35–1.80)
Heart failure — no. (%)	9 (1.2)	6 (1.6)	3 (0.8)	—
Fracture — no. (%)	17 (2.3)	8 (2.2)	9 (2.4)	1.06 (0.41–2.76)
New diagnosis of osteoporosis — no. (%)	7 (0.9)	4 (1.1)	3 (0.8)	—
Withdrawal				
Permanent discontinuation of trial regimen — no.(%)	160 (21.7)	79 (21.4)	81 (22.0)	1.06 (0.78–1.44)
Withdrawal from follow-up — no. (%)	41 (5.6)	22 (6.0)	19 (5.2)	0.84 (0.46–1.56)

[a] This table includes serious adverse events and adverse events of special interest in the modified intention-to-treat population and data on withdrawals from trial regimen and follow-up. Hazard ratios were not calculated for cardiovascular death, heart failure, or new diagnosis of osteoporosis owing to the small number of events.

[b] p = 0.05. Hazard ratios for treatment were obtained from a Cox proportional-hazards regression model predicting survival from randomised trial group and stratification variables (country, sex, and dose at randomization).

Adverse effects and events

We found no significant difference in the Hyperthyroid Symptoms score (according to the ThyPRO assessment) with levothyroxine, as compared with placebo, at any time point (Table 2, and Supplemental table 2). The incidence of serious adverse events of special interest (atrial fibrillation, heart failure, fracture, or new diagnosis of osteoporosis) was similar in the two groups (Table 3). The number of patients with at least one serious adverse event was slightly higher in the placebo group than in the levothyroxine group (p=0.049), as was the total number of serious adverse events. However, we observed no pattern of event type that contributed to this difference. The proportions of patients who discontinued the trial regimen or who withdrew from follow-up were similar in the two groups (Table 3).

DISCUSSION

In this multicentre, double-blind, randomised, placebo-controlled, parallel-group trial involving older participants with subclinical hypothyroidism, treatment with levothyroxine was associated with a persistently lower serum thyrotropin level than was placebo (between-group difference, approximately 2 mIU per litre), with the maximum effects seen at time of first review (6 to 8 weeks). We found that levothyroxine had no consistent beneficial effect on thyroid-related symptoms. This finding was true in both older men and older women and for different thyrotropin levels at baseline. Our trial had good statistical power to detect a clinically meaningful effect on thyroid-related quality of life, with 95% confidence intervals that excluded a beneficial effect greater than 2.1 points (on a scale from 0 to 100) in either of the two primary outcomes. If a symptom benefit was to have occurred, it would have been expected to be seen at 12 months.

The subsequent small-magnitude between-group difference in tiredness with levothyroxine versus placebo in the subgroup of patients who had extended follow-up is likely to be a chance finding. In contrast, an observational study of the treatment of autoimmune hypothyroidism in middle-age participants (median baseline thyrotropin level, 8.1 mIU per litre) showed that the Tiredness score improved markedly (reduction of 12 points at 6 months) and that the Hypothyroid Symptoms score also was reduced (by 2 points).[22] A small reduction in tiredness has previously been shown in a short-term trial of levothyroxine for the treatment of subclinical hypothyroidism in 120 middle-age participants.[23] There are limited data from high-quality, randomised, controlled trials regarding the effects of levothyroxine replacement in older persons with subclinical hypothyroidism.[1] Studies have generally been small (≤ 120 participants) and underpowered, often focusing on younger participants and with a short duration of follow-up.[9,10]

Levothyroxine treatment yielded no significant beneficial effects on a range of secondary-outcome measures. We found a slight deterioration (of borderline statistical significance) in the EQ-5D descriptive index with levothyroxine versus placebo at 12 months but an improvement versus placebo in the subgroup of patients who completed extended follow-up (median, 24.5 months). The effects we observed were in opposite directions at these different time points and were of very small magnitude (−0.025 at 12 months and 0.040 at extended follow-up), and therefore these are likely to be random chance findings. The estimated minimally important difference in the EQ-5D descriptive index that has been reported for other conditions is summarised in a recent review as being between 0.037 and 0.069.[24] No effect of treatment was seen with regard to the EQ visual-analogue scale scores. Therefore, it appears that levothyroxine had no clinically significant effects on generic health-related quality of life.

Muscle function has been described as being adversely affected by underactive thyroid.[25] However, we found that hand-grip strength did not change from baseline significantly more with levothyroxine treatment than with placebo. Similarly, it has been suggested that the speed of information processing is slowed in persons with subclinical hypothyroidism.[4] However, we found no benefit with levothyroxine with regard to executive cognitive function as measured by the letter–digit coding test. There also was no effect of treatment on blood pressure, weight, waist circumference, body-mass index, or the Barthel Index or Instrumental Activities of Daily Living scores.

Participants were monitored closely for adverse effects from levothyroxine treatment. We found no increase in hyperthyroid symptoms after the initiation of treatment, and there was no significant excess of serious adverse events of special interest, including atrial fibrillation, heart failure, fracture, or new diagnosis of osteoporosis. We believe that the slight excess of patients who had serious adverse events in the placebo group is a chance finding; the events were spread among a range of body systems, and no particular pattern was observed. Observational studies also have not shown any association of treatment of subclinical hypothyroidism with an increased risk of adverse events.[26]

Many older persons with biochemical results that are consistent with subclinical hypothyroidism will have reversion to a euthyroid state if they are followed up without treatment. In total, approximately three out of five persons that we screened for entry into the trial on the basis of previously elevated thyrotropin levels had reversion to normal thyroid biochemical results and were therefore excluded from the trial. These data are consistent with several other observational and trial cohorts that showed a high proportion of participants with an elevated thyrotropin level having reversion to biochemical euthyroidism during follow-up.[4,27,28]

Our trial has certain strengths. The trial included a sufficient number of participants to provide good statistical power to show no benefits regarding symptoms. We used validated measures of thyroid-specific quality of life that have been shown to be sensitive to change,[14,17] as well as a range of secondary outcomes of clinical relevance. However, the trial also had certain limitations. First, we chose to set a thyrotropin target of 0.40 to 4.60 mIU per litre with levothyroxine treatment, which is an approach that reflects recent guidelines, particularly for older persons.[7] However, some authorities have recommended a lower thyrotropin target (e.g., 0.40 to 2.50 mIU per litre).[29] We cannot exclude the possibility that this more aggressive treatment approach might be beneficial. Second, since few participants had a baseline thyrotropin level of more than 10 mIU per litre, we cannot address whether there are benefits from treatment in this subgroup. Third, the symptom levels at trial entry were low, so we cannot exclude the possibility of benefit in persons with more marked symptoms. Fourth, we did not measure thyroid antibody levels. Antibody-positive patients are more likely than antibody-negative patients to have progressive hypothyroidism and therefore may be more likely to have a benefit from long-term levothyroxine treatment.[7] Finally, our trial was underpowered to detect any effect of levothyroxine on the incidence of cardiovascular events or mortality. Therefore, we cannot exclude the possibility that treatment with levothyroxine may provide cardiovascular protection or cause harm.

In conclusion, this trial indicated that treatment with levothyroxine in older persons with subclinical hypothyroidism provided no symptomatic benefits.

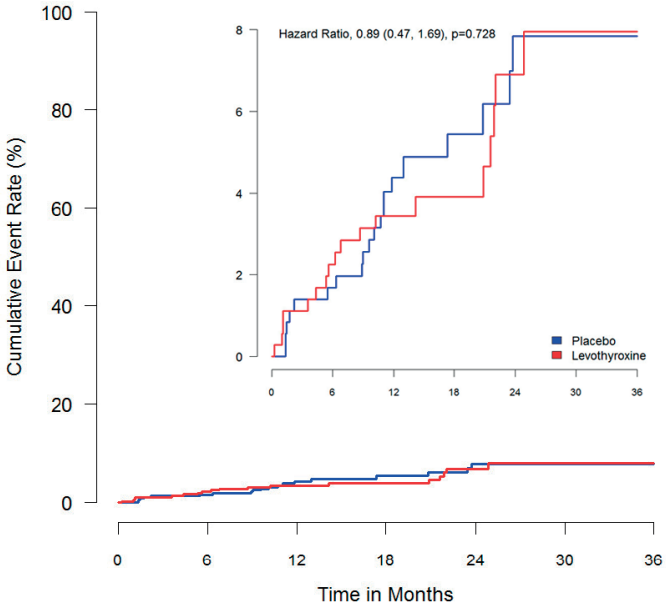
Online supplemental material

https://www.nejm.org/doi/full/10.1056/NEJMoa1603825

REFERENCES

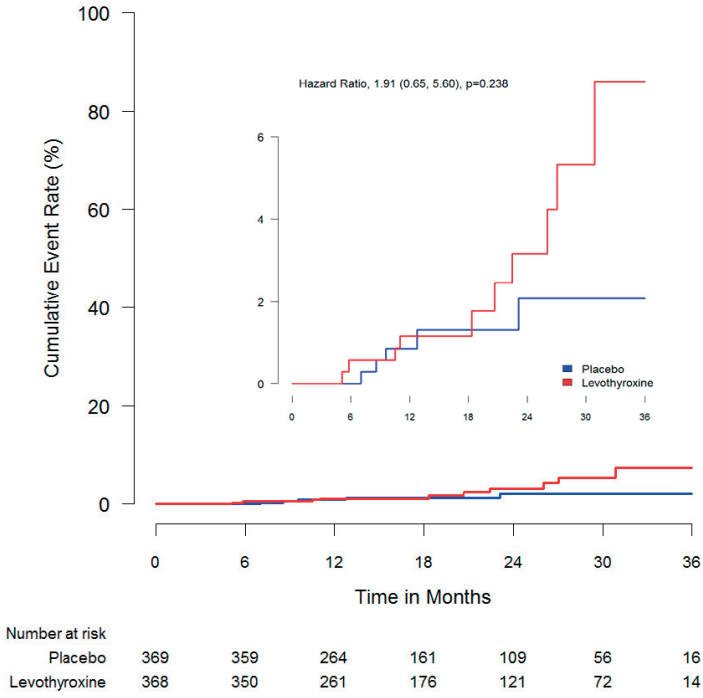
1. Rugge JB, Bougatsos C, Chou R. Screening and treatment of thyroid dysfunction: an evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med.* 2015;162(1):35-45.
2. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med.* 2000;160(4):526-534.
3. Monzani F, Di Bello V, Caraccio N, et al. Effect of levothyroxine on cardiac function and structure in subclinical hypothyroidism: a double blind, placebo-controlled study. *J Clin Endocrinol Metab.* 2001;86(3):1110-1115.
4. Parle J, Roberts L, Wilson S, et al. A randomized controlled trial of the effect of thyroxine replacement on cognitive function in community-living elderly subjects with subclinical hypothyroidism: the Birmingham Elderly Thyroid study. *J Clin Endocrinol Metab.* 2010;95(8):3623-3632.
5. Baumgartner C, den Elzen WP, Blum MR, et al. Variation in treatment strategies of Swiss general practitioners for subclinical hypothyroidism in older adults. *Swiss Med Wkly.* 2015;145:w14156.
6. Carle A, Pedersen IB, Knudsen N, Perrild H, Ovesen L, Laurberg P. Hypothyroid symptoms and the likelihood of overt thyroid failure: a population-based case-control study. *European journal of endocrinology / European Federation of Endocrine Societies.* 2014;171(5):593-602.
7. Javed Z, Sathyapalan T. Levothyroxine treatment of mild subclinical hypothyroidism: a review of potential risks and benefits. *Ther Adv Endocrinol Metab.* 2016;7(1):12-23.
8. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA.* 2010;304(12):1365-1374.
9. Mikhail GS, Alshammari SM, Alenezi MY, Mansour M, Khalil NA. Increased atherogenic low-density lipoprotein cholesterol in untreated subclinical hypothyroidism. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2008;14(5):570-575.
10. Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. *Cochrane Database Syst Rev.* 2007;2007(3):CD003419.
11. Stott DJ, Gussekloo J, Kearney PM, et al. Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST). *BMC Endocr Disord.* 2017;17(1):6.
12. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* 2013;310(20):2191-2194.
13. Medical Research Council. MRC guidelines for good clinical practice in clinical trials. In. London1998.
14. Watt T, Hegedus L, Groenvold M, et al. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(1):161-167.
15. Wong CK, Lang BH, Lam CL. A systematic review of quality of thyroid-specific health-related quality-of-life instruments recommends ThyPRO for patients with benign thyroid diseases. *J Clin Epidemiol.* 2016;78:63-72.
16. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199-208.
17. Watt T, Bjorner JB, Groenvold M, et al. Development of a Short Version of the Thyroid-Related Patient-Reported Outcome ThyPRO. *Thyroid.* 2015;25(10):1069-1079.
18. Gunther CM, Burger A, Rickert M, Crispin A, Schulz CU. Grip strength in healthy caucasian adults: reference values. *J Hand Surg Am.* 2008;33(4):558-565.

19. Houx PJ, Shepherd J, Blauw GJ, et al. Testing cognitive function in elderly populations: the PROSPER study. PROspective Study of Pravastatin in the Elderly at Risk. *J Neurol Neurosurg Psychiatry.* 2002;73(4):385-389.
20. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland state medical journal.* 1965;14:61-65.
21. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function. *JAMA.* 1963;185:914-919.
22. Winther KH, Cramon P, Watt T, et al. Disease-Specific as Well as Generic Quality of Life Is Widely Impacted in Autoimmune Hypothyroidism and Improves during the First Six Months of Levothyroxine Therapy. *PLoS One.* 2016;11(6):e0156925.
23. Razvi S, Ingoe L, Keeka G, Oates C, McMillan C, Weaver JU. The beneficial effect of L-thyroxine on cardiovascular risk factors, endothelial function, and quality of life in subclinical hypothyroidism: randomized, crossover trial. *J Clin Endocrinol Metab.* 2007;92(5):1715-1723.
24. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-Defined Estimates of the Minimally Important Difference for EQ-5D-5L Index Scores. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2017;20(4):644-650.
25. Reuters VS, Teixeira Pde F, Vigario PS, et al. Functional capacity and muscular abnormalities in subclinical hypothyroidism. *Am J Med Sci.* 2009;338(4):259-263.
26. Razvi S, Weaver JU, Butler TJ, Pearce SH. Levothyroxine treatment of subclinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. *Arch Intern Med.* 2012;172(10):811-817.
27. Diez JJ, Iglesias P, Burman KD. Spontaneous normalization of thyrotropin concentrations in patients with subclinical hypothyroidism. *J Clin Endocrinol Metab.* 2005;90(7):4124-4127.
28. Parle JV, Franklyn JA, Cross KW, Jones SC, Sheppard MC. Prevalence and follow-up of abnormal thyrotrophin (TSH) concentrations in the elderly in the United Kingdom. *Clin Endocrinol (Oxf).* 1991;34(1):77-83.
29. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. *Eur Thyroid J.* 2013;2(4):215-228.



Number at risk							
Placebo	369	344	243	151	101	52	14
Levothyroxine	368	342	249	165	111	66	14

Supplemental figure 1. Time course of incident fatal plus nonfatal cardiovascular events in placebo and levothyroxine groups (modified intention to treat population).



Supplemental figure 2. Time course of all-cause mortality in placebo and levothyroxine groups (intention to treat population).

Supplemental table 1. Baseline characteristics of study participants who provided the co- primary outcome measures (ThyPRO Hypothyroid Symptom and Fatigue scores at 12 months).

	Placebo (n=320)	Levothyroxine (n=318)
Demographics		
Age (years)	74.6 (6.8)	73.8 (5.7)
Mean, SD and range	[65.1-93.4]	[65.2-93.0]
Female sex	170 (53.1%)	166 (52.2%)
White race	315 (98.4%)	313 (98.4%)
Standard housing	307 (95.9%)	311 (97.8%)
Previous medical conditions / clinical descriptors		
Ischemic heart disease	43 (13.5%)	38 (11.9%)
Atrial fibrillation	38 (11.9%)	35 (11.1%)
Hypertension	156 (49.2%)	163 (51.3%)
Diabetes mellitus	45 (14.1%)	55 (17.3%)
Osteoporosis	44 (13.8%)	36 (11.5%)
Current smokers	24 (7.5%)	25 (7.9%)
Number of concomitant medicines [median and lower quartile, upper quartile]	4 (2, 6)	4 (2, 6)
Mini-mental state examination [median (lower quartile, upper quartile)]	29 (28, 30)	29 (27, 29)
Weight < 50Kg	3 (0.9%)	4 (1.3%)
Laboratory results		
TSH (mIU/L)		
Mean (SD)	6.3 (1.7)	6.4 (2.1)
Median (lower quartile, upper quartile)	5.8 (5.1, 6.8)	5.7 (5.1, 6.8)
fT4 (pmol/L)	13.4 (1.8)	13.4 (2.0)
Outcome measures		
ThyPRO Hypothyroid Symptoms (0-100)	16.5 (17.8)	16.4 (17.8)
ThyPRO Tiredness (0-100)	25.4 (20.2)	24.8 (19.6)
EuroQol-5D	0.857 (0.163)	0.863 (0.161)
EuroQol visual analogue scale	77.0 (16.2)	79.1 (15.4)
Handgrip strength (Kg)	27.9 (11.3)	28.6 (10.3)
Letter Digit Coding Test	25.3 (8.1)	25.2 (7.5)
Systolic Blood Pressure (mmHg)	141.4 (18.4)	141.2 (18.4)
Diastolic Blood Pressure (mmHg)	75.4 (11.7)	74.3 (11.6)
Body Mass Index (kg/m^2)	27.7 (4.5)	28.1 (5.2)
Waist circumference (cm)	97.5 (12.9)	98.4 (13.2)
Barthel index [median and range]	20 (14, 20)	20 (13, 20)
Instrumental Activities of Daily Living [median and range]	14 (7, 14)	14 (7, 14)

Results for continuous variables are expressed as mean (SD), categorical variables as number (percent), unless otherwise stated.

Standard housing was defined as non-sheltered community accommodation. Sheltered housing is purpose-built grouped housing for older people, often with an on-site scheme manager or warden. Ischemic heart disease was defined as history of angina pectoris and / or previous myocardial infarction. The Letter Digit Coding Test score (test of executive cognitive function) is number of correct responses in 90 seconds. The body mass index is the weight in kilograms divided by the square of the height in meters. The Barthel index uses the 20-point scale; the Instrumental Activities of Daily Living scale has a maximum score of 14; higher scores are associated with better performance.

No statistically significant between-group differences were seen in baseline characteristics of study participants who provided the co-primary outcome measures.

Abbreviations: fT4, free thyroxine; SD, Standard Deviation; TSH, thyroid-stimulating hormone

Supplemental table 2. Study outcomes recorded at 6-8 weeks (modified intention to treat analysis).

	Baseline		6-8 weeks		Between group difference at 6-8 weeks (95% CI);
	Placebo	**Levothyroxine**	**Placebo**	**Levothyroxine**	**Levothyroxine-placebo**
TSH (mIU/L)					
Mean (SD)	6.38 (2.01)	6.41 (2.01)	5.59 (2.24)	3.35(1.78)	-2.29 (-2.54, -2.03)[a]
Median (lower quartile, upper quartile)	5.76(5.10,6.94)	5.72 (5.10,6.83)	5.10 (4.08,6.48)	3.04 (2.20,4.02)	
	n=369	n=368	n=355	n=350	
Primary outcomes					
ThyPRO Hypothyroid Symptoms (0-100)	16.9 (17.9)	17.5 (18.8)	14.2 (17.3)	15.2 (17.2)	0.7 (-1.3, 2.7)
	n=369	n=368	n=357	n=351	
ThyPRO Tiredness (0-100)	25.5 (20.3)	25.9 (20.6)	24.1 (19.9)	25.8 (20.2)	1.6 (-0.5, 3.7)
	n=369	n=368	n=357	n=351	
Secondary outcomes					
EuroQol 5D	0.847 (0.171)	0.846 (0.187)	0.857(0.174)	0.848 (0.200)	-0.010 (-0.031, 0.010)
	n=369	n=368	n=356	n=351	
EuroQol visual anologue scale	76.5 (16.3)	78.40 (15.3)	77.9 (15.8)	77.9 (15.6)	-1.1 (-2.9, 0.6)
	n=369	n=368	n=356	n=352	
Adverse symptom assessment					
ThyPRO Hyperthyroid Symptoms (0-100)	10.5 (11.2)	10.5 (11.2)	8.8 (10.2)	8.9 (9.7)	0.3 (-0.8, 1.4)
	n=369	n=368	n=357	n=351	

Results are expressed as mean (SD) unless otherwise stated.

Results at 6-8 weeks and between-group differences are adjusted for stratification variables (country, sex and starting dose of levothyroxine) and baseline levels of the same variable using linear regression.

[a] $p < 0.001$

Abbreviations: CI, confidence interval; TSH, thyroid-stimulating hormone; ThyPRO, Thyroid-Related quality of life Patient-Reported Outcome questionnaire.

Supplemental table 3. Serum free thyroxine at study baseline, 6-8 weeks and 12 months (modified intention to treat analysis).

	Baseline		6-8 weeks		Between group difference at 6-8 weeks (95% CI)	12 months		Between group difference at 12months (95% CI)
	Placebo	Levothyroxine	Placebo	Levothyroxine	Levothyroxine-placebo difference	Placebo	Levothyroxine	Levothyroxine-placebo difference
Free thyroxine (pmol/L)	13.3 (1.9)	13.4 (2.1)	12.6 (2.4)	15.1 (2.9)	2.3 (1.8, 2.8)[a]	12.4 (1.8)	14.7 (2.8)	2.3 (1.8, 2.7)[a]
	n=369	n=368	n=110	n=108		n=89	n=98	

Free thyroxine was routinely measured at study baseline, but in trial it was not routinely measured; the data presented at 6-8 weeks and 12 months are from a convenience sample.

[a] $p < 0.001$

Supplemental table 4. Study outcomes at 12 months and extended follow-up (per protocol analysis).

	Baseline		12 months		Between group difference at 12 months (95% CI)	Extended follow-up visit		Between group difference at extended follow- up visit (95% CI)
	Placebo	Levothyroxine	Placebo	Levothyroxine	Levothyroxine-placebo difference	Placebo	Levothyroxine	Levothyroxine-placebo difference
TSH (mIU/L)								
Mean (SD)	6.38 (2.01)	6.41 (2.01)	5.54 (2.49)	3.37 (1.86)	-2.22 (-2.55,-1.90) **	5.39 (2.53)	3.14 (1.69)	-2.38 (-2.80, -1.95) **
Median (lower quartile, upper quartile)	5.76 (5.10,6.94)	5.72 (5.10,6.83)	4.91 (3.96,6.55)	3.01 (2.38,3.93)		5.00 (3.79,6.24)	2.90 (2.17,3.72)	
	n=369	n=368	n=291	n=277		n=154	n=161	
Primary outcomes								
ThyPRO Hypothyroid Symptoms (0-100)	16.9 (17.9)	17.5 (18.8)	16.2 (17.1)	15.7 (16.6)	0.0 (-2.1, 2.2)	14.9 (16.9)	18.3 (18.9)	2.6 (-0.5, 5.7)
	n=369	n=368	n=295	n=278		n=158	n=162	
ThyPRO Tiredness (0-100)	25.5 (20.3)	25.9 (20.6)	28.1 (18.9)	27.9 (19.7)	0.4 (-2.2,3.0)	30.9 (22.3)	29.7 (20.2)	-2.1 (-5.7,1.6)
	n=369	n=368	n=295	n=278		n=158	n=162	
Secondary outcomes								
EuroQol 5D	0.847 (0.171)	0.846 (0.187)	0.860 (0.178)	0.842 (0.198)	-0.025 (-0.050,0.000)*	0.842 (0.188)	0.867 (0.189)	-0.027 (-0.008,0.063)
	n=369	n=368	n=295	n=278		n=158	n=162	
EuroQol visual analogue scale	76.5 (16.3)	78.4 (15.3)	77.5 (13.5)	78.4 (14.75)	-0.4 (-2.3, 1.5)	78.4 (12.9)	76.9 (14.3)	-1.5 (-4.1, 1.2)
	n=369	n=368	n=294	n=278		n=158	n=162	

Supplemental table 4. Study outcomes at 12 months and extended follow-up (per protocol analysis). (*continued*)

	Baseline		12 months		Between group difference at 12 months (95% CI)	Extended follow-up visit		Between group difference at extended follow- up visit (95% CI)
Handgrip strength (kg)	27.5 (11.3)	28.0 (10.2)	27.2 (11.1)	27.8 (10.5)	0.0 (-0.9, 0.8)	24.9 (10.4)	24.8 (10.1)	-0.6 (-1.9, 0.6)
	n=358	n=358	n=274	n=264		n=149	n=158	
Systolic blood pressure (mmHg)	140.4 (18.9)	141.2 (18.7)	138.4 (17.4)	138.0 (18.7)	0.1 (-2.2, 2.5)	135.8 (17.9)	136.4 (17.9)	1.0 (-2.2, 4.2)
	n=368	n=368	n=294	n=278		n=155	n=160	
Diastolic blood pressure (mmHg)	74.8 (11.7)	74.1 (11.6)	73.6 (10.9)	72.8 (11.4)	0.3 (-1.2, 1.8)	72.1 (10.4)	71.8 (11.8)	1.3 (-0.7, 3.3)
	n=368	n=368	n=294	n=278		n=155	n=160	
Body mass index (kg/m^2)	27.7 (4.6)	28.1 (5.3)	27.8 (4.7)	28.0 (5.0)	0.0 (-0.2, 0.2)	27.3 (4.6)	28.1 (4.8)	0.1 (-0.2, 0.4)
	n=368	n=367	n=293	n=277		n=158	n=162	
Waist circumference (cm)	97.5 (12.8)	98.5 (13.6)	97.0 (13.2)	98.0 (12.6)	0.3 (-0.6,1.2)	95.7 (14.3)	97.9 (13.0)	0.2 (-1.1, 1.4)
	n=368	n=367	n=294	n=276		n=158	n=161	
Adverse symptom assessment								
ThyPRO Hyperthyroid Symptoms (0-100)	10.5 (11.2)	10.5 (11.2)	10.1 (11.2)	10.1 (10.6)	0.8 (-0.5,2.1)	9.8 (11.3)	10.7 (11.5)	0.5 (-1.5, 2.5)
	n=369	n=368	n=295	n=278		n=158	n=162	

* p=0.05,** p< 0.001. Results are expressed as mean (SD) unless otherwise stated.

Results at 12 months, at the extended follow-up visit and between-group differences are adjusted for stratification variables (country, sex and starting dose of levothyroxine) and baseline levels of the same variable using linear regression; data for the extended follow-up visit are additionally adjusted for time to visit. The body mass index is the weight in kilograms divided by the square of the height in metres.

Abbreviations: CI, confidence interval; IQR, interquartile range; TSH, thyroid-stimulating hormone; ThyPRO, Thyroid-Related quality of life Patient-Reported Outcome questionnaire.

Supplemental table 5. Sensitivity analyses for missing data for study outcomes at 12 months: treatment effect estimated using multiple imputation of missing values.

	Treatment Effect (Levothyroxine - Placebo)	95% confidence interval
Primary Outcomes		
ThyPRO Hypothyroid Symptoms (0-100)	-0.3	-2.2, 1.5
ThyPRO Tiredness (0-100)	0.8	-1.5, 3.0
Secondary Outcomes		
EuroQol 5D	-0.26[a]	-0.05, -0.00
EuroQol visual analogue scale	-1.1	-2.8, 0.6
Handgrip strength (kg)	-0.1	-0.8, 0.6
Systolic blood pressure (mmHg)	0.5	-1.5, 2.5
Diastolic blood pressure (mmHg)	0.2	-1.1, 1.4
Body mass index (kg/m^2)	0.1	-0.1, 0.2
Waist circumference (cm)	0.5	-0.3, 1.2
Adverse symptom assessment		
ThyPRO Hyperthyroid Symptoms (0-100)	0.3	-0.9, 1.4

Missing values are predicted from age, gender, baseline value and value at 6-8 week visit if available, in 10 imputations. The treatment effect is estimated in a linear mixed effects regression model predicting change from baseline to follow-up visit with the following covariates: randomised treatment, baseline value of outcome variable and stratification variables (site, gender, dose at randomisation).

[a] p = 0.03

Supplemental table 6. Executive cognitive function (letter digit coding test), activities of daily living, and comprehensive thyroid-related quality of life at final review in placebo and levothyroxine groups (modified intention to treat analysis).

	Baseline		Final follow-up visit		Between group difference at final follow-up visit (95% CI)
	Placebo	**Levothyroxine**	**Placebo**	**Levothyroxine**	**Levothyroxine-placebo difference**
Letter-digit coding test	25.2 (8.3)	28.0 (10.2)	27.1 (11.2)	27.5 (10.5)	-0.1 (-0.9, 0.7)
	n=366	n=358	n=298	n=302	
Barthel index [median and range]	20 (14, 20)	20 (13, 20)	20 (12, 20)	20 (9, 20)	-0.1 (-0.3, 0.1)
	n=369	n=367	n=325	n=321	
Instrumental Activities of Daily Living [median and range]	14 (7, 14)	14 (7, 14)	14 (5, 14)	14 (3, 14)	-0.1 (-0.3, 0.1)
	n=368	n=368	n=325	n=322	
ThyPRO-39 (0-100)	-	-	15.4 (11.3)	15.0 (12.0)	-0.5 (-2.2, 1.3)
			n=325	n=323	

Final follow-up visit was the last review visit; this included participants who stopped the study at 12 months, and participants with extended follow-up (beyond 12 months); final review in the placebo group was at a median of 17.7 months (lower quartile 12.0, upper quartile 24.6) and in the levothyroxine group at a median of 17.9 months (12.0, 17.0).

The Letter Digit Coding Test score (test of executive cognitive function; number of correct responses in 90 seconds), Barthel Index (20-point scale) and Instrumental Activities of Daily Living (14-point scale) were recorded at baseline and at the final visit. ThyPRO-39 was recorded only at the final visit.

There were no significant between-group differences in these outcome measures at final follow-up visit.

Abbreviations: ThyPRO-39, Thyroid-Related quality of life Patient-Reported Outcome 39-point questionnaire (comprehensive assessment of thyroid-related quality of life).

Supplemental table 7. Primary outcomes for subgroups of sex at 12-month follow-up (modified intention to treat analysis).

	Baseline		12 months		Between group difference at 12 months (95% CI)
	Placebo	Levothyroxine	Placebo	Levothyroxine	Levothyroxine-placebo difference
Male					
TSH (mIU/L)					
Mean (SD)	6.34 (2.02)	6.37 (1.87)	5.58 (2.55)	3.76 (1.97)	-1.83 (-2.41,-1.25)[a]
Median (lower quartile, upper quartile)	5.72 (5.12,6.75)	5.70 (5.10,6.91)	4.95 (3.96,6.63)	3.38 (2.60,4.21)	
	n=155	n=157	n=155	n=157	
ThyPRO Hypothyroid Symptoms (0-100)	14.3 (15.5)	12.6 (14.6)	13.5 (15.8)	14.1 (15.8)	1.6 (-1.4, 4.6)
	n=159	n=157	n=159	n=157	
ThyPRO Tiredness (0-100)	21.9 (19.2)	20.9 (16.5)	25.8 (18.8)	26.8 (19.0)	1.6 (-2.0, 5.2)
	n=159	n=157	n=159	n=157	
Female					
TSH (mIU/L)					
Mean (SD)	6.35 (1.81)	6.41 (2.16)	5.67 (4.06)	3.55 (2.22)	-2.14 (-2.69, -1.58)[a]
Median (lower quartile, upper quartile)	5.80 (5.10,6.99)	5.74 (5.11,6.82)	4.82 (3.90,6.42)	3.06 (2.23,4.25)	
	n=173	n=174	n=173	n=174	
ThyPRO Hypothyroid Symptoms (0-100)	18.7 (19.5)	20.9 (20.5)	20.3 (18.6)	19.6 (18.5)	-2.0 (-4.8, 0.8)
	n=178	n=175	n=178	n=175	
ThyPRO Tiredness (0-100)	29.2 (20.7)	29.0 (22.3)	31.6 (20.2)	31.0 (21.6)	-0.4 (-3.8, 3.0)
	n=178	n=175	n=178	n=175	

Results at 12 months are adjusted for stratification variables (country, sex and starting dose of levothyroxine) and baseline levels of the same variable using linear regression.

[a] $p < 0.001$

Supplemental table 8. Primary outcomes for subgroups of baseline TSH at 12-month follow-up (modified intention to treat analysis).

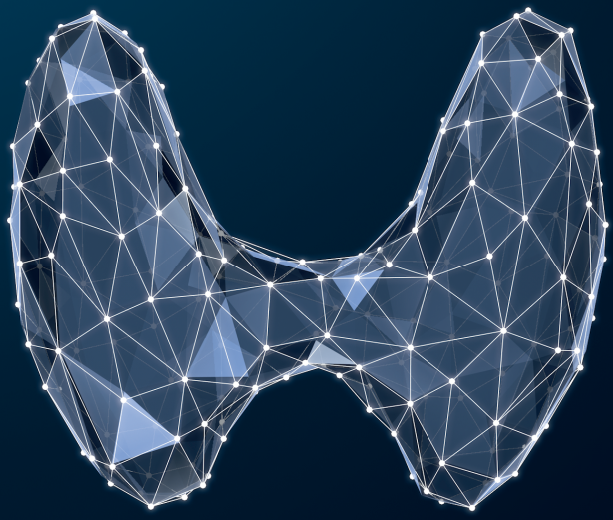
	Baseline		12 months		Between group difference at 12 months (95% CI)
	Placebo	**Levothyroxine**	**Placebo**	**Levothyroxine**	**Levothyroxine-placebo difference**
< 7 mIU/L					
TSH (mIU/L)					
Mean (SD)	5.52 (0.64)	5.55 (0.66)	5.0 (3.3)	3.4 (1.7)	-1.61 (-2.07, -1.16)[a]
Median (lower quartile, upper quartile)	5.40 (4.98,5.99)	5.40 (5.00,6.02)	4.58 (3.69,5.67)	3.12 (2.40,3.99)	
	n=250	n=256	n=250	n=256	
ThyPRO Hypothyroid Symptoms (0-100)	17.0 (17.9)	17.1 (17.7)	18.1 (18.0)	17.4 (17.2)	-0.8 (-3.1, 1.6)
	n=255	n=257	n=255	n=257	
ThyPRO Tiredness (0-100)	25.9 (20.4)	25.2 (20.1)	29.9 (19.8)	30.3 (20.9)	0.9 (-1.9, 3.7)
	n=255	n=257	n=255	n=257	
7-9.99 mIU/L					
TSH (mIU/L)					
Mean (SD)	8.08 (0.84)	8.18 (0.89)	6.95 (2.18)	3.99 (2.26)	-3.01 (-3.95, -2.07)[a]
Median (lower quartile, upper quartile)	8.02 (7.38,8.62)	7.9 (7.46,8.70)	6.76 (5.42,8.00)	3.56 (2.50,4.53)	
	n=62	n=57	n=62	n=57	
ThyPRO Hypothyroid Symptoms (0-100)	14.4 (16.2)	18.1 (21.5)	12.9 (13.7)	16.4 (18.8)	1.2 (-3.6, 6.0)
	n=65	n=57	n=65	n=57	
ThyPRO Tiredness (0-100)	25.5 (21.4)	25.1 (19.3)	24.8 (19.3)	24.4 (16.3)	-0.8 (-6.5, 5.0)
	n=65	n=57	n=65	n=57	
≥10 mIU/L					
TSH (mIU/L)					
Mean (SD)	12.48 (2.35)	12.78 (2.43)	10.06 (5.15)	6.10 (3.94)	-4.19 (-5.95, -2.43)[a]
Median (lower quartile, upper quartile)	11.60 (10.74,13.31)	11.95 (10.95,14.10)	10.90 (7.73,13.21)	4.62 (3.72,7.42)	

Supplemental table 8. Primary outcomes for subgroups of baseline TSH at 12-month follow-up (modified intention to treat analysis). (*continued*)

	Baseline		12 months		Between group difference at 12 months (95% CI)
	n=16	n=18	n=16	n=18	
ThyPRO Hypothyroid Symptoms (0-100)	20.2 (22.5)	12.2 (18.2)	19.1 (23.4)	14.2 (16.4)	0.6 (-8.3, 9.6)
	n=17	n=18	n=17	n=18	
ThyPRO Tiredness (0-100)	25.0 (15.8)	25.7 (24.8)	28.8 (20.3)	25.4 (24.3)	-2.8 (-13.6, 8.0)
	n=17	n=18	n=17	n=18	

Results at 12 months are adjusted for stratification variables (country, sex and starting dose of levothyroxine) and baseline levels of the same variable using linear regression.

[a] $p < 0.001$



Chapter 6

Study protocol: a randomised controlled trial on the clinical effects of levothyroxine treatment for subclinical hypothyroidism in people aged 80 years and over

Robert S Du Puy
Iris Postmus
David J Stott
Manuel R Blum
Rosalinde KE Poortvliet
Wendy PJ den Elzen
Robin P Peeters
Barbara C van Munster
Bruce HR Wolffenbuttel
Rudi GJ Westendorp
Patricia M Kearney
Ian Ford
Sharon Kean
C Martina Messow
Torquil Watt
J Wouter Jukema
Olaf M Dekkers
Jan WA Smit
Nicolas Rodondi
Jacobijn Gussekloo
Simon P Mooijaart

BMC Endocrine Disorders, 2018;18(1):67

DOI: 10.1186/s12902-018-0285-8

ABSTRACT

Background

Subclinical hypothyroidism is common in older people and its contribution to health and disease needs to be elucidated further. Observational and clinical trial data on the clinical effects of subclinical hypothyroidism in persons aged 80 years and over is inconclusive, with some studies suggesting harm and some suggesting benefits, translating into equipoise whether levothyroxine therapy provides clinical benefits. This manuscript describes the study protocol for the Institute for Evidence-Based Medicine in Old Age (IEMO) 80-plus thyroid trial to generate the necessary evidence base.

Methods

The IEMO 80-plus thyroid trial was explicitly designed as an ancillary experiment to the Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism randomised placebo controlled Trial (TRUST) with a near identical protocol and shared research infrastructure. Outcomes will be presented separately for the IEMO and TRUST 80-plus groups, as well as a pre-planned combined analysis of the 145 participants included in the IEMO trial and the 146 participants from the TRUST thyroid trial aged 80 years and over.

The IEMO 80-plus thyroid trial is a multi-centre randomised double-blind placebo-controlled parallel group trial of levothyroxine treatment in community-dwelling participants aged 80 years and over with persistent subclinical hypothyroidism (TSH $\geq$ 4.6 and $\leq$ 19.9 mIU/L and fT4 within laboratory reference ranges). Participants are randomised to levothyroxine 25 or 50 micrograms daily or matching placebo with dose titrations according to TSH levels, for a minimum follow-up of one and a maximum of three years.

Primary study endpoints: hypothyroid physical symptoms and tiredness on the thyroid-related quality of life patient-reported outcome (ThyPRO) at one year. Secondary endpoints: generic quality of life, executive cognitive function, handgrip strength, functional ability, blood pressure, weight, body mass index, and mortality. Adverse events will be recorded with specific interest on cardiovascular endpoints such as atrial fibrillation and heart failure.

Discussion

The combined analysis of participants in the IEMO 80-plus thyroid trial with the participants aged over 80 in the TRUST trial will provide the largest experimental evidence base on multimodal effects of levothyroxine treatment in 80-plus persons to date.

Trial registration

Nederlands (Dutch) Trial Register: NTR3851 (12–02-2013), EudraCT: 2012–004160-22 (17–02-2013), ABR-41259.058.13 (12–02-2013).

BACKGROUND

Subclinical hypothyroidism (SCH) is a common aberrant biochemical finding defined as an elevated serum thyroid-stimulating hormone (TSH) and normal circulating thyroid hormone level.[1] SCH is associated with multiple health problems in old age ranging from mild non-specific symptoms such as tiredness and emotional susceptibility to coronary heart disease and decreased physical and cognitive functioning.[2]

As 8–18% of those over 65 years are affected and inference from both observational and experimental studies maintain the clinical equipoise whether the merits of levothyroxine treatment outweigh the risks,[3] the Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism randomised placebo controlled Trial (TRUST) [4] was designed to resolve this clinical uncertainty. The outcomes of the TRUST trial provided robust information that for community-dwelling persons of 65 years of age and older with SCH, levothyroxine treatment provides no apparent benefits.[4]

There are ample data to suggest that thyroid function is mediated by age and that the effects of SCH may be profoundly different in octogenarians and older.[3] Older persons generally require different dosages of levothyroxine to achieve euthyroidism than younger counterparts possibly due to changes in body weight, composition or hormonal status [5] and are at higher risk of adverse effects of overtreatment including cardiovascular events, arrhythmias and fractures.[6] In a large-scale, observational follow-up study among 599 community-dwelling participants aged 85 years and over, increasing levels of TSH were associated with prolonged life span.[7] This association, however, could not be confirmed in a later Individual Patient Data meta-analysis investigating mortality information in 4,344 participants with SCH aged 80 years and over.[8] In addition, members of families with exceptional longevity are characterised by slightly higher TSH and slightly lower circulating thyroid hormone levels when compared with the general population.[9]

To help resolve this clinical uncertainty of levothyroxine replacement treatment for SCH in older persons, we have performed an randomised controlled trial including participants over 80 years old in the presence of comorbid conditions; the Institute for Evidence-Based Medicine in Old Age (IEMO) 80-plus thyroid trial. The TRUST trial was not designed specifically to investigate the effects in 80-plus participants and was consequently inadequately powered for a subgroup analysis in participants aged 80 and over. The IEMO 80-plus thyroid trial was designed jointly with the TRUST trial as an ancillary trial using the same trial infrastructure and protocol to allow a pre-planned, joint analysis of all participants aged 80 and over. This combined endeavour will provide experimental evidence on potential multimodal effects of levothyroxine treatment from the largest sample of 80-plus persons with SCH to date.

Among the specific study objectives are:

1. Does levothyroxine treatment for SCH provide benefits for 80-plus persons with SCH?
2. Are benefits seen across a wide range of outcomes, including health-related quality of life, muscle function, cognition and prevention of cardiovascular disease?
3. Are benefits seen in specific subgroups of people with SCH, including women, and those with mild degrees of SCH (TSH 4.6–10 mIU/L)?
4. Are any benefits offset by adverse effects, such as atrial fibrillation or heart failure?

METHODS AND DESIGN

The IEMO 80-plus thyroid trial was designed as an ancillary randomised double-blind placebo-controlled parallel group trial of levothyroxine for persons over 80 years with subclinical hypothyroidism. From the outset the study was designed jointly and in parallel with the TRUST trial (details provided elsewhere [10]) and both trials share a near identical design and infrastructure including study protocols, standard operating procedures, independent data monitoring and endpoint committees, databases, statisticians and study nurses.

Initially, the IEMO 80-plus thyroid trial aimed to include 450 participants. Additionally, a preplanned combined analysis with the data from all 80-plus participants from the TRUST trial, resulting in a total of 900 participants in the final pooled analyses, was conceived to maximise statistical power. During the inclusion phase, it became apparent that the proposed target of 450 80-plus participants was unfeasible within the allotted study period (mirroring the experiences of the TRUST trial [10]) and revised power calculations were proposed with the new projected target of 145 IEMO 80-plus trial participants (see Sample size calculation).

Originally the trial was executed in 4 regions of the Netherlands (Leiden University Medical Center, Erasmus University Medical Center, University Medical Center Groningen and the University of Amsterdam). During the inclusion period, in an attempt to maximise the inclusion rate, organisational changes were accepted allowing for inclusion of participants from all locations within the Netherlands, coordinated by the Leiden University Medical Center. Additionally, because the trial infrastructure was already in place for the TRUST trial, additional participants were recruited from the University Hospital Bern in Switzerland.

Study population

One hundred forty-five community-dwelling participants ≥ 80 years with SCH are recruited. Similar to TRUST, participants are identified from clinical and primary care laboratory databases from all patients having biochemical features consistent with SCH. SCH is defined as persistently elevated TSH levels (≥ 4.6 and ≤ 19.9 mIU/L), measured on a minimum of two

occasions at least 3 months and no more than 3 years apart prior to enrolment and free thyroxine (fT4) within the laboratory reference range. All participants gave written individual informed consent to participate.

Exclusion criteria

- Participants currently on levothyroxine, antithyroid medication (including carbimazole, methimazole, propylthiouracil and potassium perchlorate), amiodarone or lithium.
- Recent thyroid surgery or radio-iodine therapy (within 12 months).
- Grade IV NYHA heart failure.
- Prior clinical diagnosis of dementia.
- Recent hospitalisation for major illness (within 4 weeks).
- Recent acute coronary syndrome, including myocardial infarction or unstable angina (within 4 weeks).
- Acute myocarditis or acute pancarditis
- Untreated adrenal insufficiency or adrenal disorder
- Terminal illness.
- Patients known to have rare hereditary problem of galactose intolerance.
- Participants who are participating in ongoing RCTs of therapeutic interventions (including clinical trials of investigational medicinal products [CTIMPs])
- Plan to move out of the region in which the trial is being conducted within the next 2 years.

Intervention

The investigational medicinal products are levothyroxine sodium (T4) as 25 or 50 microgram tablets for oral administration and a matching placebo. All tablets are white and round in shape with the strength imprinted, identically packaged in blisters and packed in plain cardboard cartons to maintain study blinding. Participants are advised to take the suggested dose of study medication once daily half an hour before breakfast.

The intervention group will start with levothyroxine 50 micrograms daily (25 micrograms in participants with < 50 kg body weight or with a history of coronary heart disease) and the control group with matching placebo for six to eight weeks.

After 6–8 weeks a venous blood sample is taken for TSH assessment. Based on the TSH results, the data centre advises the new dose of study medication or placebo to the clinical investigators.

- If TSH < 0.4 mIU/L: the treatment dose is reduced to 25 micrograms levothyroxine in those starting on 50 micrograms; reduced to 0 in those starting on 25 micrograms – effected by giving placebo matching the 25 micrograms dose. These participants will have a further

TSH check after 6–8 weeks. If TSH remains < 0.4 mIU/L participant will be withdrawn from randomised treatment and referred to usual care.
- If TSH ≥ 0.4 and < 4.6 mIU/L: no change to the treatment dose.
- If TSH remains elevated (≥ 4.6 mIU/L): 25 micrograms of levothyroxine will be added. Giving a total daily dose of 75 micrograms levothyroxine for those starting on 50 micrograms, or a total daily dose of 50 micrograms levothyroxine for those starting on 25 micrograms.

A maximum of two levothyroxine up-titrations at the start of the trial and one up-titration at 12- and 24-month (± 1 month) intervals with repeated TSH measurements after 6–8 weeks ensure adequate levothyroxine treatment while avoiding potential over-replacement. The maximum possible dose of levothyroxine is 150 micrograms.

A mock titration adopting an adaptive schedule is performed in the placebo group by the data centre. A similar proportion of placebo patients will have up and down titrations of study medication as the intervention group to ensure the number of tablets and assessments is similar in both groups.

Because all thyroid function measurements are available only to the data centre, the clinical investigators remain fully blinded to the treatment allocation process during the trial.

Accountability logs recording the quantities of study medication dispensed to and returned from study participants, batch numbers and expiry dates are available for all study drug movements.

Criteria for discontinuing or modifying allocated study medication:

- If overt biochemical hypothyroidism is identified (TSH > 20 mIU/L and/or fT4 below the reference range) a second TSH with fT4 within 2 weeks is requested. Upon confirmation of biochemical hypothyroidism, the participant will be withdrawn from the study treatment and referred to the General Practitioner (GP) for usual care.
- If overt biochemical hyperthyroidism is identified (TSH < 0.4 mIU/L) in the placebo group, or consecutively in the treatment group despite down titrations, the participant will be withdrawn from the study treatment and referred to the GP for usual care.
- If for clinical reasons (e.g. major illness) a proposed change in study medication or placebo is deemed inappropriate the algorithm is overridden by the local principal coordinator and no change in study medication takes place.

Randomisation

Participants are randomised to either the levothyroxine or placebo treatment arm (ratio 1:1) using the randomly permuted block method, stratified by site, sex and starting dose. The data

centre (Robertson Centre for Biostatistics, University of Glasgow, Scotland) independently provides the randomisation schedule. Mawdsley Brooks & Co. implements the schedule through identical packaging of levothyroxine and matching placebo tablets.

Patient allocation is conducted via the dedicated trial web portal by the study nurses. When a participant is eligible based on entering the eligibility criteria in an electronic case report form (eCRF) supervised by a medically certified Principal Investigator, a central computer will trigger the decision.

Blinding

Participants are blinded to treatment allocation by using matching tablets and packaging for levothyroxine and placebo. All study personnel remain blinded for the duration of the trial through remote analysis of laboratory results of TSH in the data centre, ensuring the trial stays double blinded. GPs will remain blinded to treatment allocation and TSH tests unless otherwise required in the event of an emergency medical situation. An Interactive Voice Response System at the data centre allows for individual emergency allocation information to be released to an unblinded study physician through 24-h telephone access. All participants will learn the treatment allocation within 15 working days of receiving the final visit and completing all the data to aid in any further treatment decisions with the GP.

All laboratory tests for TSH and fT4 are performed at the local GP and clinical laboratories. The results in the treatment phase are uploaded to the independent data centre which in turn advises the study site on dose titration through the dedicated trial web portal. The study team remains unaware of the results of the thyroid function testing. Additionally, all cooperating GPs were asked to refrain from additional thyroid function measurements to ensure adequate blinding.

Data collection

Data collection will be performed by study research nurses at baseline and predetermined follow-up visits at the participant's home or place of residence. All participants are followed up for a minimum of 1 year with a likely average of 2 years. Participants are reviewed face-to-face by the study nurses at recruitment, study baseline, 6–8 weeks, 12 months, 24 months, 36 months and at the final visit. This personal approach ensures data quality and promotes participant retention and complete follow-up. In addition, interim telephone contact or visits (depending on the desire of the patient) are made by study nurses at 6, 18 and 30 months (depending on total duration of follow-up), including recording of possible cardiovascular and serious adverse events (SAEs). For a timeline of assessments and visits see Table 1.

Table 1. Detailed schedule of assessments.

Months of follow up	**0 visit**	**6-8 wks**	**6m**	**12m**	**18m**	**24m**	**30m**	**36m**	**Final[a] visit**
		visit	**call/visit**	**visit**	**call/visit**	**visit**	**call/visit**	**visit**	
Participant characteristics & medical history	x								
Weight, height, waist circumference and BMI	x			x					x
Concomitant medication	x	x		x		x		x	x
Home support	x								x
Safety and monitoring									
Morbidity, mortality, hospitalisation and GP contacts		x	x	x	x	x	x	x	x
Serious Adverse Events		x	x	x	x	x	x	x	x
Single-lead ECG (for AF)	x			x		x		x	x
Drug adherence		x	x	x	x	x	x	x	x
Outcomes									
Thyroid related quality of life (ThyPRO)	x	x		x					x
Generic quality of life (EQ-5D-3L)	x	x		x					x
Cognitive function									
MMSE	x								
Letter Digit Coding Test	x								x
Functional ability									
ADL (BI), IADL (OARS), falls questionnaire	x								x
Handgrip strength & 6-meter gait speed	x			x					x
Blood pressure	x			x					x
Fatal and non-fatal cardiovascular events		x	x	x	x	x	x	x	x
Arthritis complaints									x
Treatment Satisfaction (TSQM vII)									x
Laboratory analysis									
Thyroid function	x	x		x		x		x	x
Haemoglobin	x			x					
Blood samples for biobank	x			x					

[a] the final visit assessments may substitute for any assessment time between 12 and a maximum of 42 months

All study nurses are trained simultaneously on the data to be assessed. All measuring equipment is calibrated before the start and annually thereafter to safeguard reliability and validity. The Data centre will develop and manage a dedicated, anonymised trial web portal, including the electronic case report forms in Dutch and Swiss Standard German. This portal is based on the dedicated trial web portal from the TRUST trial to maximise the homogeneity of data and to allow for pre-planned pooled analysis of the results. Personal information used for trial logistics is collected and stored in a separate electronic study database in accordance with legal and ethical requirements.

Data validation checks give study personnel immediate feedback on missing or out of range values. Logic checks reduce the possibilities of entering invalid data. Database validation checks are run routinely and are tracked and escalated as appropriate. Data will be locked at the end of the study according to preregistered lockdown procedures by the data centre. The data centre will provide the independent data monitoring committee (IDMC) and the authorities with (annual) safety reports on the Data.

Outcomes

At 6–8 weeks we expect most patients allocated to levothyroxine to be biochemically euthyroid, and at this time point short-term improvements (such as in thyroid-related quality of life) will be assessed. By 1 year the medium-term effects of levothyroxine treatment should emerge (such as muscle function). The long-term effects of treatment of SCH will be determined by assessment over the full course of the study, with a mean of 2 years treatment duration.

In the screening phase, results from TSH and fT4 tests, exclusion criteria, informed consent for the screening phase of the study, informed consent for the trial phase of the study are obtained by the study nurses.

During the baseline phase of the study the following data are recorded:

- Participant characteristics: age, sex, ethnicity, information on alcohol and tobacco use, height.
- Any clinical changes that would violate the inclusion or exclusion criteria
- Concomitant drugs used: prescribed medication, over-the-counter non-steroidal anti-inflammatory drugs and aspirin
- History of disease: Cardiovascular disease including history of ischaemic heart disease (angina pectoris or previous myocardial infarction), cerebrovascular disease (ischaemic stroke, transient ischaemic attack) or peripheral vascular disease (intermittent claudication), or any revascularisation procedure for ischaemic vascular disease. History of atrial fibrillation, epilepsy, hypertension, diabetes mellitus or osteoporosis.

- Single lead ECG: to check for atrial fibrillation.
- Cognitive function: Mini-mental state examination (MMSE [11]) score as an indicator of general cognitive function. This will not be used as an outcome measure due to insensitivity to change during the trial.
- Home support services: (e.g. home help, meals-on-wheels, district nursing) and home circumstances (e.g. living alone, co-habiting, standard or sheltered housing, or entry to care home)

Primary study endpoints

The main study primary endpoints are mean change from baseline scores in thyroid-related quality of life and symptom burden assessed using the hypothyroid symptoms scale score and tiredness symptoms scale score on the thyroid-related quality of life patient-reported outcome (ThyPRO) [12] at 12 months after recruitment. The primary analyses will be done in the 80 years and over group (IEMO and TRUST participants). The results will be compared through subgroup analysis with those in the 79 years and under group (TRUST participants) as a secondary analysis. The ThyPRO is an 85-item patient-reported outcome measure, evaluating symptoms, well-being and function on 85 items summarised in 14 scales, ranging 0–100, with higher scores representing more symptoms or impact of disease. For this study three scales with 19 items are evaluated: Tiredness, Hypothyroid physical symptoms and Hyperthyroid physical symptoms.

Secondary study endpoints

- Generic quality of life: EuroQOL EQ-5D-3 L [13] at baseline, 6–8 weeks, 12 months and final follow up.
- Thyroid-related quality of life ThyPRO [12] at baseline, 6–8 weeks and at final follow-up.
- Thyroid-related quality of life: ThyPRO-39 [14] recorded at final follow-up (additional 28 questions).
- Executive cognitive function: Letter Digit Coding Test [LDCT] [15] at baseline and final follow-up.
- Handgrip strength: Jamar hand dynamometer (best of 3 measures in dominant hand) at baseline, 12 months and final follow up.
- Functional ability: Activities of Daily Living (Barthel Index [BI] [16,17]), Instrumental Activities of Daily Living (Older Americans Resources and Services [OARS] [18]), 6-m gait speed [19], independent living status and falls questionnaire at baseline and final follow up.
- Blood pressure: systolic and diastolic measured at baseline, 12 months and final follow up
- Height, weight, waist circumference and body mass index: recorded at baseline, 12 months and final follow up

- Mortality: all-cause and cardiovascular are requested through national mortality registries
- Fatal and non-fatal cardiovascular events: including acute myocardial infarction, stroke, amputations for peripheral vascular disease and revascularisations for atherosclerotic vascular disease (including for acute coronary syndrome and heart failure hospitalisations).

Additional measurements

- Treatment satisfaction with trial medication: Treatment Satisfaction Questionnaire for Medication vII (TSQM [20]) and desire of post-trial medication continuation recorded at final follow up.
- Arthritis: data regarding joints, skeletal functioning and arthritis are recorded through an arthritis questionnaire at final follow up.
- Haemoglobin: measured on a full blood count at baseline and 12 months.

See Table 1 for detailed schedule of assessments.

Safety

Full details of all Serious Adverse Events (SAEs), Adverse Events (AEs) of special interest (atrial fibrillation, heart failure, fractures, new diagnosis of osteoporosis), study treatment withdrawals and ThyPRO hyperthyroid symptoms are recorded at all visits and telephone contacts. Participants and GPs have 24-h access to an emergency trial phone number operated by a certified physician for the reporting of SAEs.

Biobank

Blood samples for the IEMO biobank are collected at baseline (40 ml venous blood) and at 12 months (10 ml venous blood). The following 19 aliquots (0.75 ml each) are stored per participant at baseline: 3 EDTA plasma, 1 whole blood, 2 citrated plasma, 1 NaF plasma, 1 buffy coat, 3 heparin plasma, and 8 serum aliquots. The 12 months bloods are stored in four serum 0.75 ml aliquots per participant.

Analyses in the IEMO biobank will be performed in combination with the TRUST biobank. Both biobanks are organised by the same biobank committee. The IEMO biobank will be stored at the Department of Clinical Chemistry of Leiden University Medical Center (LUMC), the Netherlands. The biobank consists of all plasma, serum, and DNA material of all randomised IEMO participants that provide consent for storing biobank material. The Department of Clinical Chemistry of the LUMC is fully accredited (EN ISO 15189:2012) by the Dutch Accreditation Council. The Biobank adheres to all necessary quality assurance standards and legal guidelines.

Sample size calculation

The total number of participants in all published trials on SCH before 2017 is 450 across 12 studies, including only a small number of older people and very heterogeneous endpoints across studies. We aim to study endpoints that are of particular relevance for the oldest old, including endpoints in those with considerable comorbidity.

Originally, the IEMO 80-plus thyroid trial had set out to analyse 450 participants with SCH aged 80 years and over. Additionally, a pre-planned pooled analysis of 900 participants was agreed upon, of which 450 were recruited directly through this study and a subset over the age of 80 years from the TRUST trial would add another 450 participants, to further increase the statistical power to detect significant changes in this subgroup. The power calculations were based on two main study endpoints:

1. Fatal and non-fatal cardiovascular events.
2. Change in thyroid-related quality of life (ThyPRO Tiredness and Hypothyroid physical symptoms).

Due to several limiting factors including delays in starting the studies, caused by difficulties procuring study medication and matching placebos, it proved impossible to reach this number, similar to the experiences in the TRUST trial [10]. Therefore, in 2015, revised power calculations were proposed (study protocol amendment 8, 04/06/2015) and accepted by the funding agent, sponsor, medical ethical committee (15/07/15) and competent authority (03/07/2015). These revisions detailed the change of primary study endpoint cardiovascular events into a secondary study outcome, accepting the possibility of being underpowered to answer this secondary endpoint. This allowed the power calculations to be revised according to the remaining primary outcome thyroid-related quality of life.

The resulting revised sample-size calculation is based on the pre-planned pooled analysis of one of the co-primary endpoints of thyroid-related quality of life (ThyPRO Hypothyroid physical symptoms and Tiredness scale score). According to previous studies applying the ThyPRO, a study should be adequately powered for at least a difference of 9 points to be clinically meaningful. Using an expected standard deviation of the difference of 26 [21] and a power of 80%, 132 participants are required per trial group adding to a total of 264 participants to be included in the combined 80-plus analyses. For all secondary continuous endpoints this sample size is deemed large enough to provide statistically robust results. For the secondary endpoints on cardiovascular events and mortality the possibility of being underpowered is accepted.

Over a recruitment period of almost 3 years the TRUST trial recruited 737 participants to the trial of which 146 participants were aged 80 and over. Assuming 10% loss to follow-up in

both trials a projected 145 additional participants will be recruited in the IEMO trial. The follow up phase of the trial is expected to be complete in May 2018 with one additional month of SAE recording.

Data analysis

The data centre (Robertson Centre for Biostatistics, Glasgow, ISO 9001/2008 certified) is responsible for writing, implementing and revising a statistical analysis plan that is agreed upon before locking the study database and will have full access to the final study database for the planned analyses. A copy of the statistical analysis plan is appended to this manuscript as Additional file 1. All analyses are based on a modified intention-to-treat principle and the primary time-point for analysis is after 12 months of treatment. The main analyses will be based on the combined IEMO and TRUST 80-plus participants (n = 291).

Analyses will be presented separately for:

- the IEMO 80-plus participant cohort (n = 145)
- the TRUST 80-plus participant subset (n = 146)
- the combined IEMO and TRUST 80-plus participants compared with the TRUST 80-minus participants (n = 291 vs n = 591)

Summary information for all participants and between the treatment groups will be made available. Similar to the TRUST trial [10], continuous variables measured at baseline and follow-up will be analysed at each time point comparing treatment groups adjusting for stratification variables and baseline levels of the same variable using analysis of covariance. Additionally, repeated measures regression analysis will be performed with regards to the primary time-point and final assessment for each participant. For calculating ThyPRO scores, raw total scores containing valid missing items will be scaled to maintain the maximum possible score. Clinical outcome data will be analysed using time-to-first-event Cox proportional hazards regression analysis in models that contain the randomised treatment allocation and stratification variables as covariates. Treatment effect will be analysed using the Wald-test and corresponding point estimates and 95% confidence intervals for the hazard ratio for treatment will be estimated. The assumption of proportionality of hazards will be tested.

Analysis of the primary outcomes will be performed in the modified intention to treat (ITT) population, based on participants with data available on the outcome of interest. The ITT population will be used for analyses on efficacy and safety. In addition, analyses using mixed effects models and multiple imputations will be used for sensitivity analysis. The per protocol population will also be used for all primary and secondary outcomes as exploratory analyses.

Owing to the intended similarities in study design between the IEMO 80-plus thyroid trial and the TRUST trial, the data allow for a pooled subgroup analysis of the TRUST and IEMO 80-plus participants compared with the TRUST 80-minus participants. Outcome differences between these groups will highlight the additional clinical merits or adverse effects of levothyroxine replacement therapy for older participants aged 80 and over.

Other pre-planned subgroup analyses include: baseline TSH in two groups (< 10/ ≥ 10 mIU/L) or in three groups (< 7/7–9.99/ ≥ 10 mIU/L), sex (male/female). However, we accept that our study will be underpowered for some of the smaller subgroups, such as male participants, TSH above 10.0 and below 19.9 mIU/L. We should however have sufficient statistical power in the combined analysis to detect beneficial effects in the larger or dominant subgroups, such as female and TSH in the range above 4.6 and below 10.0 mIU/L.

Monitoring and committees

To secure the highest quality of participant care and safety, the careful titration algorithm avoids the possibility of prolonged periods of levothyroxine over-replacement. Similarly, the system guards against participants developing overt hypothyroidism that might require open-label levothyroxine use.

All SAEs, AEs and AEs of special interest are recorded, notified, assessed, reported, analysed and managed in accordance with the Medicines for Human Use (Clinical Trials) Regulations 2004 and the study protocol. All events are followed up until resolution or stabilization occurs, and are assessed for seriousness, expectedness and causality by the chief investigator. Serious adverse events are reported to the sponsor by thorough recording in the eCRF and to the local accredited Medical Ethics Committee and competent authority. Annually and at the end of the trial 100% study monitoring visits are conducted by independent clinical research associates, in accordance with the Netherlands Federation of University Medical Centres' report 'Kwaliteitsborging van Mensgebonden onderzoek'. All important decisions made leading to protocol modifications are communicated to all relevant parties, including the trial registry, ethical committees and competent authorities.

All main decisions for the study were made by the steering group. Its members are: Dr. Simon P Mooijaart, Dr. Jacobijn Gussekloo, Dr. Olaf M Dekkers, Dr. Jan Smit, Dr. J Wouter Jukema, Dr. Anton. JM de Craen (Leiden, the Netherlands, Deceased).

Each national site was supervised by a local organising committee and Principal Investigator. For the Netherlands the organising committee was: Dr. Simon P Mooijaart (PI), Dr. Rosalinde KE Poortvliet, Dr. Iris Postmus, Robert S Du Puy, MSc, Professor Robin. P Peeters, Professor

Bruce. HR Wolffenbuttel and Dr. Barbara. C van Munster. For Switzerland the organising committee was: Professor N Rodondi (PI) and Dr. Manuel Blum,

An Independent Data Monitoring Committee (IDMC) assesses safety data in order to protect the ethical and safety interests of the participants recruited into the study, while safeguarding, as far as possible, the scientific validity of the study. The IDMC reviews annual safety and efficacy data and may request additional data if considered necessary. The IDMC meets at least once a year and is composed of medical experts and a biostatistician without any involvement in the study as investigators or as study participant care physicians. The committee is empowered to make a recommendation on early stopping when there is overwhelming evidence of benefit for the primary outcome or when it considers there is adequate evidence of harm. The IDMC members are: Professor Gary Ford (Chair; Chief Executive Officer of the Oxford Academic Health Science Network, Oxford), Professor Thompson G Robinson (University Hospitals of Leicester NHS Trust, Department of Cardiovascular Sciences, Leicester Royal Infirmary, Leicester), Professor Colin Dayan (Institute of Molecular and Experimental Medicine, Cardiff University School of Medicine, Heath Park, Cardiff), Professor Kathleen Bennett (Department of Pharmacology and Therapeutics, Trinity Centre for Health Sciences, St James's Hospital, Dublin).

A study Endpoint Committee, blinded to treatment allocation, provides independent and unbiased review of clinical endpoint events which occur during the study, ensures unified and unambiguous events evaluation practices across the study and compensates for regional diversity in medical practice at the site of endpoint evaluation and classification. All causes of death, stroke, myocardial infarction and heart failure hospitalisations are potential endpoints to be reviewed on the data supplied through the eCRF and if necessary, acquired source documentation. The Endpoint Committee members are: Professor Peter Langhorne (Chair; Professor of Stroke Care, Institute of Cardiovascular and Medical Sciences, University of Glasgow), Professor J Wouter Jukema (Vice-chair; Professor of Cardiology, Leiden University Medical Center, The Netherlands), Dr. Tinh-Hai Collet (Department of Ambulatory Care and Community Medicine, University of Lausanne, Switzerland), Professor Olaf M Dekkers Leiden University Medical Center, The Netherlands) and Dr. Anne Marie O'Flynn (Department of Epidemiology and Public Health, UCC, Ireland).

A TRUST/IEMO Biobank committee supervises the storage and analysis of the biobank samples. Members are: Professor Patricia M Kearney, Dr. H Anette van Dorland (Bern, Switzerland), Dr. Wendy PJ den Elzen.

Each national study site is supervised by a local sponsor, responsible for the oversight of the clinical trial and supplying proper insurance to cover any liabilities during and after the trial

arising from trial conduct and participation. The sponsors are not involved in the preparation, or approval of any scientific outputs.

Dissemination

This study is well suited to promote effective dissemination of the results and implications. Arrangements regarding sharing of data and joint publication are laid down in a Memorandum of Understanding between the TRUST trial and IEMO 80-plus thyroid trial project group. Due to its role as a knowledge centre with an education and research program in the field of ageing, vitality and geriatric medicine, the Leiden Academy on Vitality and Ageing is well placed to play a coordinating role in the dissemination activities. Schildklier Organisatie Nederland, the patient advocacy group, will closely collaborate with the study team to help align the study outputs with the patients and public need.

The Institute for Evidence-based Medicine in Old Age (the Netherlands) is ideally placed to ensure that the results of the study are considered by relevant professionals, and will be included in the leading clinical guidelines. In cooperation with the Cochrane collaboration the results of the trial will be offered for the update of the Cochrane systematic review of treatment of subclinical hypothyroidism, allowing for independent scientific interpretation, placing results in context and maximising understanding of the implication of the trial.

To comply with the general social responsibility associated with clinical research, the trial results will be proactively disseminated to the general public and key public health stakeholders through established media networks.

DISCUSSION

In the latest Cochrane review of levothyroxine replacement therapy for SCH (12 studies, only 491 participants in total) most studies excluded those who suffered from multimorbidity, none of the studies reported on oldest old separately and two trials excluded those over the age of 80 years.[22] Robust evidence for the potential clinical merits or adverse effects in 80-plus persons with SCH is greatly needed to help guide clinical practice.

The IEMO 80-plus thyroid trial is a representative randomised controlled trial on levothyroxine treatment for SCH, with representative 80-plus persons and a wide range of characteristics and morbidities, studying end-points that are relevant to the older population and clinical practice. The combined analysis of participants in the IEMO 80-plus thyroid trial with those aged over 80 who participated in the TRUST trial will provide the largest experimental evi-

dence base on the multimodal effects of levothyroxine treatment in 80-plus persons to date. Trial results are expected to be publicly disseminated in the fall of 2018.

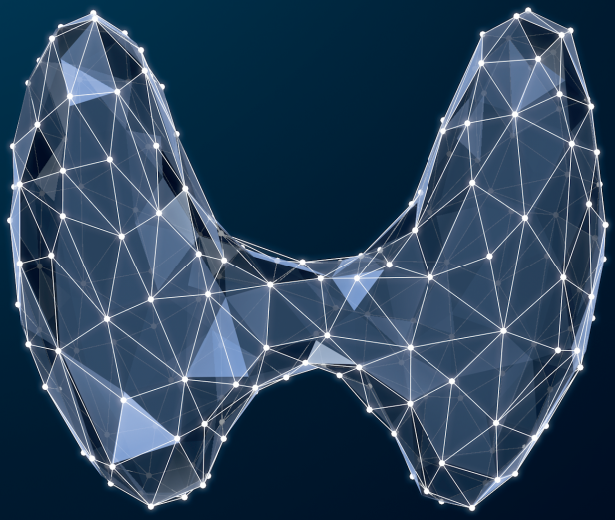
Online supplemental material

https://bmcendocrdisord.biomedcentral.com/articles/10.1186/s12902-018-0285-8

REFERENCES

1. Jones DD, May KE, Geraci SA. Subclinical thyroid disease. *Am J Med.* 2010;123(6):502-504.
2. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2012;18(6):988-1028.
3. Biondi B, Cooper DS. The clinical significance of subclinical thyroid dysfunction. *Endocrine reviews.* 2008;29(1):76-131.
4. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.
5. Devdhar M, Drooger R, Pehlivanova M, Singh G, Jonklaas J. Levothyroxine replacement doses are affected by gender and weight, but not age. *Thyroid.* 2011;21(8):821-827.
6. Flynn RW, Bonellie SR, Jung RT, MacDonald TM, Morris AD, Leese GP. Serum thyroid-stimulating hormone concentration and morbidity from cardiovascular disease and fractures in patients on long-term thyroxine therapy. *J Clin Endocrinol Metab.* 2010;95(1):186-193.
7. Gussekloo J, van Exel E, de Craen AJ, Meinders AE, Frolich M, Westendorp RG. Thyroid status, disability and cognitive function, and survival in old age. *JAMA.* 2004;292(21):2591-2599.
8. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA.* 2010;304(12):1365-1374.
9. Jansen SW, Akintola AA, Roelfsema F, et al. Human longevity is characterised by high thyroid stimulating hormone secretion without altered energy metabolism. *Sci Rep.* 2015;5:11525.
10. Stott DJ, Gussekloo J, Kearney PM, et al. Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST). *BMC Endocr Disord.* 2017;17(1):6.
11. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". *Journal of Psychiatric Research.* 1975;12(3):189-198.
12. Watt T, Hegedus L, Groenvold M, et al. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(1):161-167.
13. Rabin R, Gudex C, Selai C, Herdman M. From translation to version management: a history and review of methods for the cultural adaptation of the EuroQol five-dimensional questionnaire. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2014;17(1):70-76.
14. Watt T, Bjorner JB, Groenvold M, et al. Development of a Short Version of the Thyroid-Related Patient-Reported Outcome ThyPRO. *Thyroid.* 2015;25(10):1069-1079.
15. Smith A. The Symbol Digit Modalities Test. A neuropsychological test for economic screening of learning and other cerebral disorders. *Learning Disorders.* 1968(3):82-91.
16. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland state medical journal.* 1965;14:61-65.
17. Quinn TJ, Langhorne P, Stott DJ. Barthel index for stroke trials: development, properties, and application. *Stroke.* 2011;42(4):1146-1151.
18. George LK, Fillenbaum GG. OARS methodology. A decade of experience in geriatric assessment. *J Am Geriatr Soc.* 1985;33(9):607-615.
19. Bohannon RW. Measurement of gait speed of older adults is feasible and informative in a home-care setting. *Journal of geriatric physical therapy (2001).* 2009;32(1):22-23.

20. Atkinson MJ, Kumar R, Cappelleri JC, Hass SL. Hierarchical construct validity of the treatment satisfaction questionnaire for medication (TSQM version II) among outpatient pharmacy consumers. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2005;8 Suppl 1:S9-S24.
21. Winther KH, Cramon P, Watt T, et al. Disease-Specific as Well as Generic Quality of Life Is Widely Impacted in Autoimmune Hypothyroidism and Improves during the First Six Months of Levothyroxine Therapy. *PLoS One.* 2016;11(6):e0156925.
22. Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. *Cochrane Database Syst Rev.* 2007;2007(3):CD003419.



Chapter 7

Association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism

Simon P Mooijaart
Robert S Du Puy
David J Stott
Patricia M Kearney
Nicolas Rodondi
Rudi GJ Westendorp
Wendy PJ den Elzen
Iris Postmus
Rosalinde KE Poortvliet
Diana van Heemst
Barbara C van Munster
Robin P Peeters
Ian Ford
Sharon Kean
C Martina Messow
Manuel R Blum
Tinh-Hai Collet
Torquil Watt
Olaf M Dekkers
J Wouter Jukema
Johannes WA Smit
Peter Langhorne
Jacobijn Gussekloo

JAMA, 2019;322(20):1-11

DOI: 10.1001/jama.2019.17274

Treatment of older adults with subclinical hypothyroidism – reply

Simon P Mooijaart, Robert S Du Puy

JAMA, 2020;323(11):1097-1098

DOI: 10.1001/jama.2020.0948

ABSTRACT

Importance

It is unclear whether levothyroxine treatment provides clinically important benefits in adults aged 80 years and older with subclinical hypothyroidism.

Objective

To determine the association of levothyroxine treatment for subclinical hypothyroidism with thyroid-related quality of life in adults aged 80 years and older.

Design, Setting, and Participants

Prospectively planned combined analysis of data involving community-dwelling adults aged 80 years and older with subclinical hypothyroidism. Data from a randomised clinical trial were combined with a subgroup of participants aged 80 years and older from a second clinical trial. The trials were conducted between April 2013 and May 2018. Final follow-up was May 4, 2018.

Exposures

Participants were randomly assigned to receive levothyroxine (n = 112; 52 participants from the first trial and 60 from the second trial) or placebo (n = 139; 53 participants from the first trial and 86 from the second trial).

Main Outcomes and Measures

Co-primary outcomes were Thyroid-Related Quality of Life Patient-Reported Outcome (ThyPRO) questionnaire scores for the domains of hypothyroid symptoms and tiredness at 1 year (range, 0-100; higher scores indicate worse quality of life; minimal clinically important difference, 9).

Results

Of 251 participants (mean age, 85 years; 118 [47%] women), 105 were included from the first clinical trial and 146 were included from the second clinical trial. A total of 212 participants (84%) completed the study. The hypothyroid symptoms score decreased from 21.7 at baseline to 19.3 at 12 months in the levothyroxine group vs from 19.8 at baseline to 17.4 at 12 months in the placebo group (adjusted between-group difference, 1.3 [95% CI, −2.7 to 5.2]; P = .53). The tiredness score increased from 25.5 at baseline to 28.2 at 12 months in the levothyroxine group vs from 25.1 at baseline to 28.7 at 12 months in the placebo group (adjusted between-group difference, −0.1 [95% CI, −4.5 to 4.3]; P = .96). At least 1 adverse event occurred in 33 participants (29.5%) in the levothyroxine group (the most common adverse event was cerebrovascular accident, which occurred in 3 participants [2.2%]) and 40

participants (28.8%) in the placebo group (the most common adverse event was pneumonia, which occurred in 4 [3.6%] participants).

Conclusions and Relevance

In this prospectively planned analysis of data from 2 clinical trials involving adults aged 80 years and older with subclinical hypothyroidism, treatment with levothyroxine, compared with placebo, was not significantly associated with improvement in hypothyroid symptoms or fatigue. These findings do not support routine use of levothyroxine for treatment of subclinical hypothyroidism in adults aged 80 years and older.

Trial Registration

ClinicalTrials.gov Identifier: NCT01660126; Netherlands Trial Register: NTR3851

Key Points

Question

Among adults aged 80 years and older with subclinical hypothyroidism, what is the association between treatment with levothyroxine and thyroid-related symptoms?

Findings

In this pooled analysis of data from 2 randomised clinical trials that included 251 participants aged 80 years and older, treatment with levothyroxine, compared with placebo, was not significantly associated with improvement in thyroid-related patient-reported quality of life outcome scores (range, 0-100; higher scores indicate worse quality of life; minimal clinically important difference, 9) for hypothyroid symptoms (adjusted between-group difference, 1.3) or tiredness (adjusted between-group difference, 0.1).

Meaning

These findings do not support routine treatment with levothyroxine for subclinical hypothyroidism in adults aged 80 years and older.

INTRODUCTION

The prevalence of subclinical hypothyroidism increases with age.[1] Subclinical hypothyroidism is defined by elevated levels of thyrotropin (often referred to as thyroid-stimulating hormone [TSH]) simultaneously with free thyroxine (FT4) within the normal range. Some patients with subclinical hypothyroidism report symptoms such as constipation, mental slowness, fatigue, or depressive symptoms.[2,3] Subclinical hypothyroidism has also been associated with an increased risk of cardiovascular disease.[4]

In a 2017 clinical randomised trial of 737 participants aged 65 years and older, treatment with levothyroxine demonstrated no benefit on the primary outcome of thyroid-specific quality of life.[5] However, individuals aged 80 years and older with subclinical hypothyroidism have been underrepresented in clinical trials[6,7] and outcomes such as quality of life have not been reported for this age group. Because the prevalence of comorbidities and frailty increase with age, it is possible that benefits and harms from managing subclinical hypothyroidism may differ in adults aged 80 years and older compared with younger age groups.[8] The lack of evidence for older patients may have contributed to significant treatment variation by primary care clinicians.[9]

This study combined data from the Institute for Evidence-Based Medicine in Old Age (IEMO) 80-plus thyroid trial[10] and the subgroup of participants in the Thyroid Hormone Replacement for Untreated Older Adults With Subclinical Hypothyroidism Trial (TRUST) aged 80 years and older.[5] The 2 trials examined the effects of managing subclinical hypothyroidism with levothyroxine on quality of life in adults aged 80 years and older.

METHODS

Both clinical trials were approved by the Central Committee on Research Involving Human Subjects in the Netherlands and by the Bern and Lausanne ethics committees and Swissmedic, the Swiss authority on drugs, in Switzerland. One trial was also approved by the Multicentre Research Ethics Committee and the Medicines and Healthcare products Regulatory Agency in the United Kingdom, with cosponsors the National Health Services Greater Glasgow and Clyde and the University of Glasgow, and by the Clinical Research Ethics Committee, Cork, and the Health Products Regulatory Authority (formerly known as the Irish Medicines Board) in Ireland. Written informed consent was obtained from all participants. The study protocol and statistical analysis plans for each trial have been published.[10,11] The protocol and statistical analysis plan for the combined analyses are available in Supplement 1 and Supplement 2.

Design

The included studies were randomised, double-blind, placebo-controlled parallel-group trials investigating the effects of levothyroxine treatment for persons with subclinical hypothyroidism aged 80 years and older and aged 65 years and older.[10,11] They were conducted with similar study designs and included a prospectively planned combined analysis of all participants aged 80 years and older. The cohorts are presented and analysed as a single study group throughout this report.

Study Population

One trial recruited community-dwelling participants aged 80 and older years between May 2014 and May 2017, from sites in the Netherlands and Switzerland,[10] with a final date of follow-up of May 4, 2018. The other trial recruited community-dwelling participants aged 65 years and older from sites in the Netherlands, Switzerland, Ireland, and the United Kingdom between April 2013 and May 2015, with a final date of follow-up of October 31, 2016 [11]; only participants aged 80 years and older from this trial were included in this report. Participants were followed up for a minimum of 12 months and a maximum of 36 months. Information on participant race was collected to evaluate for racial differences in outcomes. Participants selected their race from 8 prespecified options reflecting the most common racial groups in the countries of the study sites.

Inclusion Criteria

Eligibility criteria for both trials have been published.[10,11] Individuals with persistent subclinical hypothyroidism aged 80 years and older, defined as elevated thyrotropin levels (4.6-19.9 mIU/L), measured on at least 2 occasions between 3 months and 3 years apart, who had FT4 levels within laboratory reference ranges were eligible. Eligible persons were identified from lists of patients with laboratory test results from hospitals and primary care practices. Exclusion criteria included use of levothyroxine, antithyroid medication, amiodarone, or lithium; recent thyroid surgery or radioiodine therapy; New York Heart Association class IV heart failure; clinical diagnosis of dementia; recent hospitalization for major illness; recent acute coronary syndrome, acute myocarditis, or pancarditis; and terminal illness.

Randomization and Blinding

Participants were randomised in a 1:1 ratio using a computer-based program to receive levothyroxine or placebo using randomly permuted blocks in a block size of 4, stratified by site, sex, and starting dose. Randomization was performed separately for each trial. The independent data centre (Robertson Centre for Biostatistics, University of Glasgow, United Kingdom) provided the randomization schedule; the 2 interventions were identically packaged by Mawdsley Brooks & Co (United Kingdom). Participants, general practitioners, and

study personnel were blinded to treatment allocation and thyroid function test results throughout the study.

Intervention and Control

The study medication consisted of levothyroxine sodium tablets and matching placebo tablets taken orally once daily. The levothyroxine group started with 50 µg daily (or 25 µg for participants with body weight < 50 kg or a history of coronary heart disease) and the placebo group started with a matching placebo for 6 to 8 weeks. The dose of levothyroxine was adjusted in 25-µg increments based on thyrotropin levels measured 6 to 8 weeks after starting the intervention, 6 to 8 weeks after each dose adjustment, and at 12- and 24-month follow-up with the goal of attaining a thyrotropin level within the reference range (0.4-4.6 mIU/L) in the levothyroxine group. An identical schedule for adjusting the dose of the placebo was used to achieve an equal number of titrations between the groups to maintain blinding. Laboratory test results were uploaded in the computer system by laboratory personnel not involved in the study, and study medication was prescribed in a blinded fashion using an automated system, according to the algorithm. The participants, investigators, and treating physicians were unaware of the results of thyrotropin measurements throughout the course of the trial and remained blinded for treatment allocation.

Study Outcomes

Primary outcomes and prespecified secondary outcomes were measured at the 12-month follow-up and at the end of the study, defined as the last study visit for each participant. The co-primary study outcomes were the change from baseline to the 12 month follow-up in the hypothyroid symptoms score (4 items) and tiredness score (7 items) from the Thyroid-Related Quality of Life Patient-Reported Outcome (ThyPRO) [12] questionnaire (range, 0-100; higher scores indicate more symptoms; minimal clinically important difference [MCID], 9).[13]

Prespecified secondary outcomes differed between the protocol (Supplement 1), which was finalised in June 2016, and the statistical analysis plan (Supplement 2), which was finalised in May 2018 before the results from the trial of patients aged 80 years and older were known to the investigators. The secondary outcomes were prespecified in the statistical analysis plan (Supplement 2) and include change from baseline to 12 months in thyrotropin levels; general quality of life measured using the EuroQol-5D index (range, −0.59 to 1.00; higher scores indicate a better quality of life; MCID, 0.037-0.069) [14] and the EuroQoL Visual Analogue Scale (range, 0-100; higher scores indicate better quality of life; MICD, 8) [15]; handgrip strength using the Jamar isometric dynamometer (best of 3 measurements in the dominant hand; MCID, 5.0-6.5 kg [16]); weight and body mass index (MCID, 5%-10% change for both [17,18]); waist circumference (MCID, 4 cm [19]); systolic and diastolic blood pressure (MCID, 5 mm Hg [20]), FT4 test results; and incidence of falls.

The statistical analysis plan (Supplement 2) also prespecified end-of-study outcomes, which were measured at the final follow-up visit attended by each participant. The prespecified end-of-study outcomes were all of the secondary outcomes listed above in addition to change from baseline in the hypothyroid symptoms and tiredness scores (assessed with the ThyPRO), Barthel Index [21] activities of daily living score (range, 0-20; higher scores indicate greater ability to perform activities of daily living; MCID, 1.85 [22,23]), Older Americans Resources and Services [24] instrumental activities of daily living score (range, 0-14; higher scores indicate better performance in instrumental activities of daily living; MCID, 1 [25]), executive cognitive function measured with the Letter Digit Coding Test [26] (the number of correct substitutions of digits and letters in a 90-second period; minimum score, 0; higher scores indicate better cognitive function; MCID, 4 or 10% [27]), living situation (living independently or not independently and living alone or not living alone), gait speed (MICD for change, 0.1-0.2 m/s [28]), and ThyPRO-39 questionnaire score. However, the following prespecified outcomes were listed in the statistical analysis plan but were not reported here: FT4 test results, ThyPRO-39 score, gait speed, and falls. Information on gait speed and falls was only available for 1 of the trials.

Adverse Events

The statistical analysis plan (Supplement 2) prespecified the following outcomes, consisting of adverse events, at the end of the study: combined fatal and nonfatal cardiovascular events (defined as acute myocardial infarction, stroke, amputations for peripheral vascular disease, revascularizations for atherosclerotic vascular disease, acute coronary syndrome, and heart failure hospitalizations), fatal cardiovascular events, and mortality. The following safety outcomes were also prespecified in the statistical analysis plan (Supplement 2): hypothyroidism, atrial fibrillation, heart failure, and fractures. Adverse events were reviewed by a blinded adjudication group of 5 investigators. The hyperthyroid symptom score from the ThyPRO questionnaire [12] was used to assess for overtreatment or the development of hyperthyroidism.

Sample Size

The sample size was calculated for the primary outcomes of mean change from baseline in the hypothyroid symptoms and tiredness scores at 12 months (pooled data). All sample size calculations were based on statistical power of 0.80 and a 2-sided α of .05. The 2-sided significance level for each primary outcome was .025. The study was considered positive if this level of significance was achieved for either outcome.

Assuming an SD of 26 for the score change for each primary outcome,[29] 264 participants were required (132 per group) to detect the MCID of 9 on the hypothyroid symptoms and tiredness scores. Assuming a dropout rate of 10% in the first 12 months, the target sample

size was 291 participants in the pooled analysis. Originally, the proposed sample size was 900 participants for the primary outcomes of fatal and nonfatal cardiovascular events and mean change in hypothyroid symptoms and tiredness scores on the ThyPRO questionnaire. Due to recruitment difficulty, protocols were amended and fatal or nonfatal cardiovascular events became prespecified secondary outcomes in the protocols and adverse events in the statistical analysis plan (Supplement 2).[10,11] Investigators were blinded to this process.

Statistical Methods

Primary outcome and adverse event analyses were performed for participants with data available at the 12-month follow-up. In a preplanned secondary analysis, analyses were repeated for the outcome at the final visit instead of 12 months. Continuous variables measured at baseline and during follow-up were analysed at each time point, comparing change from baseline between the placebo and levothyroxine groups, using mixed-effects models, adjusting for stratification variables and baseline values and including study as a random effect. Outcomes measured at more than 1 follow-up were also analysed using linear mixed-effects regression analysis, including data at all time points up to 12 months and repeated analyses including data at all available time points. Additional sensitivity analyses were performed using mixed-effect models and multiple imputations.[30] Ten imputed data sets were generated, imputing all missing values of the outcome variable at 12-month follow-up from age, sex, baseline thyrotropin level, baseline measurement of the outcome variable and, if available, measurement of the outcome variable at 6- to 8-week follow-up visits.

Total ThyPRO scores per domain containing missing items were scaled to maintain the maximum possible score for analyses. If more than 50% of the items for a score were missing, the score was considered to be missing. Categorical outcomes were analysed using Cox proportional hazards regression analysis in models that contained the randomised treatment allocation and stratification variables as covariates. Adjusted differences were analysed using the Wald test. Corresponding point estimates and 95% CIs for the hazard ratio for treatment were estimated. The assumption of proportionality of hazards was checked, and the assumption was met, using diagnostic plots of the log(−log(survival)) vs log(survival) and of the Schoenfield residuals over time. The 2-sided significance level was .025 for each primary outcome (hypothyroid symptoms and tiredness scores) and .05 for secondary outcomes. Because of the potential for type I error due to multiple comparisons, secondary outcomes should be interpreted as exploratory. All statistical analyses were performed in R, version 3.2.4 (R Development Core Team).

RESULTS

Of 251 randomised participants aged 80 years and older in the combined analyses, 112 (52 from the first trial and 60 from the second) were randomised to receive levothyroxine and 139 (53 from the first trial and 86 from the second) were randomised to receive the placebo. Results are presented for the combined group of participants in the 2 trials and for each trial separately (Supplemental tables 1, 3, 4, and 5 and Supplemental figures 1 and 2). In one trial, 342 adults aged 80 years and older were screened for eligibility. In the other trial, 2,647 participants aged 65 years and older were screened for eligibility (Figure 1; Supplemental figure 1). The mean (SD) age of participants was 84.6 (3.6) years and 118 (47%) were women. A higher percentage of participants in the placebo group had a history of ischemic heart disease compared with the levothyroxine group (27.3% vs 20.5%) (Table 1). Mean (SD)

Table 1. Baseline Characteristics of Participants

Characteristic	No. (%)	
	Levothyroxine (n = 112)	Placebo (n = 139)
Age, mean (SD) [range], y	84.0 (3.3) [80.0-97.1]	85.0 (3.7) [80.1-96.7]
Sex		
Men	60 (53.6)	73 (52.5)
Women	52 (46.4)	66 (47.5)
Self-identified race[a]		
White	112 (100.0)	137 (98.6)
Asian	0	1 (0.7)
Black	0	0
Other	0	1 (0.7)
Living independently	106 (94.6)	129 (92.8)
Previous medical conditions or clinical descriptors		
Hypertension	60 (53.6)	63 (45.7)
Ischemic heart disease	23 (20.5)	38 (27.3)
Atrial fibrillation	24 (21.6)	24 (17.3)
Osteoporosis	16 (14.5)	23 (16.5)
Diabetes mellitus	14 (12.5)	17 (12.3)
Current smoker	6 (5.4)	6 (4.3)
Concomitant medicines	106 (94.6)	134 (96.4)
No. of concomitant medications, median (IQR)[b]	5.0 (3.0-7.0)	4.0 (2.5-7.0)
Mini-Mental State Examination score, median (IQR)[c]	28.0 (27.0-29.5)	28.0 (27.0-29.0)
Weight < 50 kg	1 (0.9)	1 (0.7)
Laboratory results		
Thyrotropin level, mIU/L[d]		
Mean (SD)	6.4 (1.8)	6.3 (1.9)
Median (IQR) [range]	5.8 (5.1-7.2) [4.6-12.5]	5.7 (5.2-6.6) [4.6-17.6]
Free thyroxine, pmol/L[e]	13.8 (2.1)	13.8 (2.2)

Table 1. Baseline Characteristics of Participants (*continued*)

Characteristic	No. (%)	
	Levothyroxine (n = 112)	Placebo (n = 139)
Outcome measures		
ThyPRO hypothyroid symptoms score[f]		
Mean (SD)	20.8 (19.1)	19.7 (19.3)
Median (IQR)	18.8 (6.2, 31.2)	12.5 (6.2, 31.2)
ThyPRO tiredness score[f]		
Mean (SD)	25.3 (21.0)	25.7 (19.6)
Median (IQR)	21.4 (9.8-35.7)	18.8 (10.7-35.7)
EuroQol-5D score[g]	0.774 (0.222)	0.808 (0.215)
EuroQol VAS score[h]	74.8 (14.8)	73.9 (14.4)
Handgrip strength, kg[i]	24.8 (9.0)	24.3 (9.9)
Letter-digit coding test[j]	21.6 (7.1)	21.3 (7.9)
Blood pressure, mean (SD), mmHg		
Systolic	144.2 (20.4)	145.4 (20.7)
Diastolic	70.8 (11.8)	72.0 (11.9)
Body mass index	27.8 (4.5)	27.4 (3.9)
Waist circumference, cm	99.2 (12.1)	98.0 (11.0)
Barthel index, median (IQR)[k]	20.0 (19.0-20.0)	20.0 (19.0-20.0)
Instrumental activities of daily living score, median(IQR)	14.0 (13.0-14.0)	14.0 (13.0-14.0)

[a] Participants selected their race from prespecified options reflecting the most common race groups in the study site countries.

[b] The number of distinct ATC codes, excluding emollients and protectives, antiseptics and disinfectants, topical products for joint and muscular pain, nasal preparations, and local eye and ear medications.

[c] Range, 0-30; 29-30 indicates no cognitive impairment; 25-28, mild cognitive impairment; 19-24, moderate cognitive impairment; less than 19, cognitive impairment.

[d] Inclusion criteria range, 4.6-19.99 mIU/L.

[e] Inclusion criteria ranges were laboratory and method specific (all reference ranges within 9 and 2 pmol/L), however, repeated within-participant measurements were solely performed using the same method for that individual to exclude any bias due to different assays.

[f] Range 0-100; higher scores indicate more symptoms. The mean scores in the general population are 14 for hypothyroid symptoms and 35 for tiredness.

[g] Range 0.50-1.00; higher scores indicate a better quality of life.

[h] Range 0-100; higher scores indicate better health.

[i] Higher scores indicate better muscle strength; median for women aged at least 80 years, 15.4 kg; median for men older than 80 years, 24.5 kg.

[j] The number of digits coded within 90 seconds; higher scores indicate better executive cognitive function.

[k] Range 0-20; 20 indicates fully independent in activities of daily living and mobility; 15-19, moderately to fully independent; 10-14, needing help but capable of own activities; 5-9, severely dependent; 0-4, totally dependent.

[l] Range 0-14, higher scores indicate better performance in instrumental activities of daily living.

Abbreviations: IQR, interquartile range; ThyPRO, Thyroid-Related Quality of Life Patient-Reported Outcome; VAS, visual analogue scale.

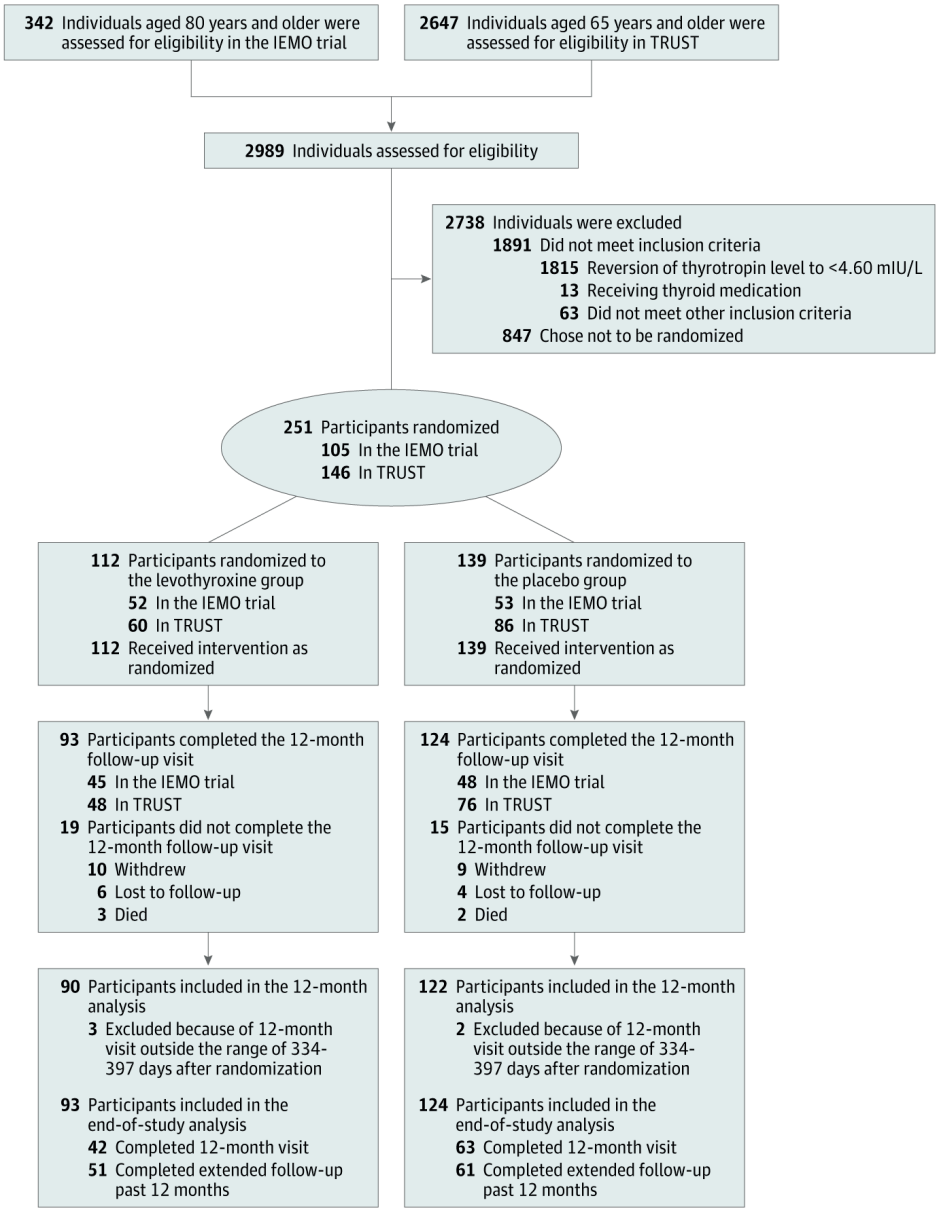


Figure 1. Recruitment, randomization, and patient flow of the participants in a study of the association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism

The difference in the number of participants randomised was due to a chance occurrence and, in part, to the failure to stratify randomization by age in the Thyroid Hormone Replacement for Untreated Older Adults With Subclinical Hypothyroidism Trial (TRUST). The 12-month window was defined as between 334 and 397 days after randomization. IEMO indicates Institute for Evidence-Based Medicine in Old Age.

thyrotropin at baseline was 6.3 (1.9) mIU/L in the placebo group vs 6.4 (1.8) mIU/L in the levothyroxine group. No other clinically relevant differences between treatment groups at baseline were observed (Supplemental table 1).

Dropout rates were similar between the groups (10 participants [8.9%] in the levothyroxine group vs 9 [6.5%] in the placebo group). After 12 months, 122 (88%) participants in the placebo group and 90 (80%) in the levothyroxine group had follow-up laboratory measures available (Supplemental table 2). Mean (SD) thyrotropin levels decreased from 6.20 (1.48) mIU/L at baseline to 5.49 (2.21) mIU/L at 12 months in the placebo group and from 6.50 (1.80) mIU/L at baseline to 3.69 (1.81) mIU/L at 12 months in the levothyroxine group (estimated mean between-group difference, −1.9 mIU/L [95% CI, −2.49 to −1.45]; $p < 0.001$; Figure 2) in regression models. Thyrotropin levels were significantly different between the placebo and levothyroxine groups at all time points during follow-up (all $p < 0.001$), including when the 2 studies were analysed separately (all $p < 0.01$) (Supplemental figure 2).

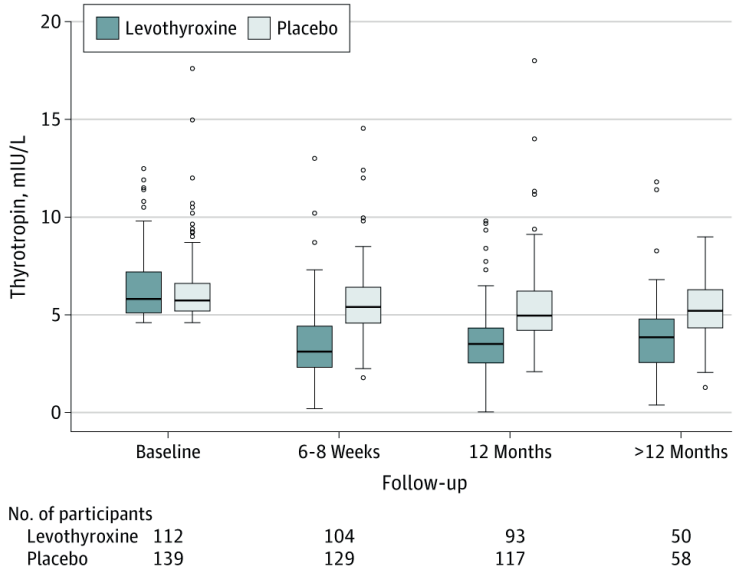


Figure 2. Thyrotropin Levels of Participants in a Study of the Association Between Levothyroxine Treatment and Thyroid-Related Symptoms Among Adults Aged 80 Years and Older With Subclinical Hypothyroidism

There was no significant association of levothyroxine treatment with change in the hypothyroid symptoms score at the 12-month follow-up (21.7 at baseline and 19.3 at 12 months in the levothyroxine group vs 19.8 at baseline and 17.4 at 12 months in the placebo group; adjusted between-group difference, 1.27 [95% CI, −2.69 to 5.23]; $P = .53$). There was also no significant association in the tiredness score (25.5 at baseline and 28.2 at 12 months in the

Table 2. Outcomes at 12 months for participants in a study of the association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism

Outcomes	Mean (SD)				Adjusted Difference (95% CI)[a]	p Value
	Baseline		12 Months			
	Levothyroxine	Placebo	Levothyroxine	Placebo		
Co-primary outcomes	(n = 90)	(n = 122)	(n = 90)	(n = 122)		
ThyPRO hypothyroid symptoms score[b]	21.7 (19.5)	19.8 (19.6)	19.3 (18.2)	17.4 (18.1)	1.27 (−2.69 to 5.23)	0.53
ThyPRO tiredness score[b]	25.2 (21.5)	25.1 (19.5)	28.2 (20.0)	28.7 (19.9)	−0.10 (−4.51 to 4.31)	0.96
Prespecified secondary outcomes						
Thyrotropin, mIU/L	6.50 (1.80)	6.20 (1.48)	3.69 (1.81)	5.49 (2.21)	−1.97 (−2.49 to −1.45)	<0.001
	(n = 90)	(n = 116)	(n = 90)	(n = 116)		
EuroQol-5D[c]	0.785 (0.199)	0.811 (0.210)	0.754 (0.268)	0.785 (0.244)	−0.012 (−0.063 to 0.039)	0.64
	(n = 90)	(n = 122)	(n = 90)	(n = 122)		
EuroQol VAS[d]	75.27 (14.59)	73.98 (14.26)	74.16 (13.67)	73.67 (13.58)	−0.42 (−3.57 to 2.72)	0.79
	(n = 90)	(n = 122)	(n = 90)	(n = 122)		
Handgrip strength, kg	25.4 (9.4)	24.7 (10.2)	23.4 (9.7)	23.0 (9.2)	−0.27 (−1.79 to 1.25)	0.73
	(n = 81)	(n = 114)	(n = 81)	(n = 114)		
Blood pressure, mmHg	(n = 90)	(n = 122)	(n = 90)	(n = 122)		
Systolic	144.4 (19.4)	146.2 (20.7)	141.3 (19.0)	142.6 (20.7)	−0.42 (−5.23 to 4.39)	0.86
Diastolic	71.3 (11.8)	72.3 (12.3)	68.7 (11.9)	69.6 (12.5)	−0.31 (−3.06 to 2.44)	0.83
Weight	76.5 (13.1)	75.1 (12.3)	76.2 (12.7)	73.9 (12.3)	0.97 (0.11 to 1.82)	0.03
	(n = 90)	(n = 121)	(n = 90)	(n = 121)		
Body mass index	27.7 (4.4)	27.5 (3.9)	27.6 (4.4)	27.1 (3.9)	0.38 (0.08 to 0.68)	0.01
	(n = 90)	(n = 121)	(n = 90)	(n = 121)		
Waist circumference, cm	99.0 (11.6)	98.0 (10.9)	99.0 (11.4)	96.5 (11.7)	1.52 (0.09 to 2.95)	0.04
	(n = 90)	(n = 121)	(n = 90)	(n = 121)		

[a] Adjusted difference was estimated in linear regression models predicting change from baseline to 12-month visit (95% CI) with study site, sex, and randomization dose as stratification variables and study as random effect.

[b] Range 0-100; higher scores indicate more symptoms. The mean scores in the general population are 14 for hypothyroid symptoms and 35 for tiredness.

[c] Range 0.50-1.00; higher scores indicate a better quality of life.

[d] Range 0-100; higher scores indicate better health.

Abbreviations: ThyPRO, Thyroid-Related Quality of Life Patient-Reported Outcome; VAS, visual analogue scale.

levothyroxine group vs 25.1 at baseline and 28.7 at 12 months in the placebo group; adjusted between-group difference, −0.10 [95% CI, −4.51 to 4.31]; P = .96) (Table 2).

Levothyroxine treatment was not significantly associated with changes from baseline to 12 months in general quality of life measured with the EuroQol-5D index (change from 0.785 to 0.754 in the levothyroxine group vs 0.811 to 0.785 in the placebo group; adjusted difference, −0.012 [95% CI, −0.063 to 0.039]) or physical function as measured by handgrip strength at 12 months (change from 25.4 kg to 23.4 kg in the levothyroxine group vs 24.7 kg to 23.0 kg in the placebo group; adjusted between-group difference, −0.27 kg [95% CI, −1.79 to 1.25]) (Table 2).

Treatment with levothyroxine was associated with a statistically significant increase in body mass index (between-group difference, 0.38 [95% CI, 0.08-0.68]; P = .01) and in waist circumference (between-group difference, 1.52 cm [95% CI, 0.09-2.95]; P = .04) compared with placebo (Table 2).

Data for 17 participants (12%) in the placebo group and 22 (20%) in the levothyroxine group who had incomplete follow-up data were imputed. Sensitivity analyses using repeated measures mixed-effect regression models and imputed data for outcomes at 12 months showed similar results (Supplemental table 4).

For the end-of-study outcomes, the mean follow-up was 17 months. Treatment with levothyroxine was not significantly associated with thyroid-specific or overall quality of life based on the end-of-study outcomes (Table 3). There was no significant association of levothyroxine treatment with activities of daily living measured using the Barthel Index (from 19.3 at baseline to 19.0 at the end of the study in the levothyroxine group vs from 19.4 to 19.1 in the placebo group; adjusted mean difference, 0.09 [95% CI, −0.33 to 0.52]) or in executive cognitive function measured with the Letter Digit Coding Test (mean between-group difference, 1.24 [95% CI, −0.30 to 2.78]) at the end of the study.

During a mean follow-up of 17 months (median follow-up, 13 months), 9 participants (3.6%) died (1 cardiovascular death). Levothyroxine was not associated with increased rates of fatal or nonfatal cardiovascular events (unadjusted hazard ratio, 0.61 [95% CI, 0.24-1.50]; event rate per 100 person-years of 4.2 in the levothyroxine group and 7.64 in the placebo group) or overall mortality (unadjusted hazard ratio, 1.39 [95% CI, 0.37-5.19]; event rate of 2.99 in the levothyroxine group and 2.02 in the placebo group) (Table 4). A total of 73 participants (29%; 33 [29.5%] in the levothyroxine group and 40 [28.8%] in the placebo group) experienced 1 or more serious adverse events. Adverse events included new-onset atrial fibrillation (10 participants [4.5%]), heart failure (9 [4.1%]), and fractures (9 [4.1%]). Hypothyroidism did not occur in any of the participants. In the levothyroxine group, the most common adverse

Table 3. End-of-study outcomes for participants in a study of the association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism.

Prespecified Secondary Outcome	Mean (SD)				Adjusted Difference (95% CI)[a]	p Value
	Baseline		End of Study			
	Levothyroxine	Placebo	Levothyroxine	Placebo		
ThyPRO hypothyroid symptoms score[b]	22.2 (19.4)	19.9 (19.5)	19.9 (19.6)	18.2 (19.9)	−0.48 (−6.88 to 5.91)	0.88
	(n = 93)	(n = 124)	(n = 93)	(n = 124)		
ThyPRO Tiredness score[b]	24.9 (21.4)	25.1 (19.4)	29.5 (18.2)	30.2 (23.3)	−0.77 (−7.07 to 5.54)	0.81
	(n = 93)	(n = 124)	(n = 93)	(n = 124)		
EuroQol-5D[b]	0.777 (0.212)	0.812 (0.209)	0.763 (0.228)	0.806 (0.213)	−0.024 (−0.097 to 0.049)	0.52
	(n = 93)	(n = 124)	(n = 93)	(n = 124)		
EuroQol VAS[b]	74.88 (14.68)	73.94 (14.20)	74.10 (11.97)	74.28 (14.32)	−1.60 (−6.16 to 2.96)	0.49
	(n = 93)	(n = 123)	(n = 93)	(n = 123)		
Handgrip strength, kg[b]	24.9 (9.3)	24.1 (10.4)	19.4 (8.0)	21.8 (10.6)	−1.43 (−4.17 to 1.31)	0.31
	(n = 89)	(n = 109)	(n = 89)	(n = 109)		
Letter-digit coding test[b]	22.0 (7.1)	21.9 (7.7)	21.6 (9.2)	20.5 (7.1)	1.24 (−0.30 to 2.78)	0.11
	(n = 85)	(n = 109)	(n = 85)	(n = 109)		
Barthel index[b]	19.3 (1.5)	19.4 (1.2)	19.0 (1.9)	19.1 (2.1)	0.09 (−0.33 to 0.52)	0.66
	(n = 91)	(n = 117)	(n = 91)	(n = 117)		
Instrumental activities of daily living[b]	13.3 (1.3)	13.2 (1.5)	12.4 (2.4)	12.7 (2.3)	−0.40 (−0.92 to 0.13)	0.14
	(n = 91)	(n = 117)	(n = 91)	(n = 117)		

[a] Adjusted difference was estimated in linear mixed-effects regression models predicting change from baseline to end of the study with study site, sex, and intervention as stratification variables and study as random effect.

[b] For scale definitions see Table 1 footnotes.

Abbreviations: ThyPRO, Thyroid-Related Quality of Life Patient-Reported Outcome; VAS, visual analogue scale.

Table 4. Clinical and adverse events in in a study of the association between levothyroxine treatment and thyroid-related symptoms among adults aged 80 years and older with subclinical hypothyroidism.

Outcomes	No. (%)		Event Rate per 100 Person-Years		Estimated Risk Difference (95% CI)[b]
	Levothyroxine (n = 112)	Placebo (n = 139)	Levothyroxine (n = 112)	Placebo (n = 139)	
Prespecified secondary outcomes					
Clinical outcomes					
Fatal or nonfatal cardiovascular event	7 (6.3)	14 (10.1)	4.52	7.64	Hazard ratio, 0.61 (0.24 to 1.50)
Death from any cause	5 (4.5)	4 (2.9)	2.99	2.02	Hazard ratio, 1.39 (0.37 to 5.19)
Adverse events					
Cardiovascular death[c]	0	1 (0.7)	0.00	0.51	
Serious adverse events					
Events	53	61			
Participants with >1 serious adverse event[d]	33 (29.5)	40 (28.8)	−0.01 (−0.04 to 0.01)		
Adverse event of special interest					
New-onset atrial fibrillation	4 (3.6)	6 (4.3)	2.57	3.19	0.00 (−0.02 to 0.03)
Heart failure	3 (2.7)	6 (4.3)	1.90	3.21	0.01 (−0.03 to 0.05)
Fracture	4 (3.6)	5 (3.6)	2.53	2.68	0.00 (−0.04 to 0.03)
Hypothyroidism[e]	0 (0.0)	0 (0.0)			
Withdrawal					
Permanent discontinuation of trial regimen[f]	38 (33.9)	43 (30.9)	27.34	24.84	−0.04 (−0.15 to 0.05)
Withdrawal from follow-up[g]	10 (8.9)	9 (6.5)	5.98	4.56	−0.03 (−0.09 to 0.02)
ThyPRO hyperthyroid symptoms scores[h]	10.9 (11.3)	9.1 (10.8)	9.6 (9.4)	9.3 (10.4)	Adjusted difference, −0.50 (−2.62 to 1.63)[e]

[a] Adverse events were recorded and reported until the end of the study. Preplanned secondary outcomes effects were estimated using Cox proportional hazard regression models adjusted for sex, dose at randomization, study site and study. Adverse events event rates were estimated using Cox proportional hazard regression models with, where possible, adjustment for study site, study, dose at randomization, sex, and age, presented as risk differences and 95% CIs, obtained through bootstrap resampling in 1000 iterations

[b] Unless otherwise noted.

[c] For outcomes with too few events to run regression models, event rates and a log-rank test p value were reported. Serious adverse events were all undesired medical events involving a participant, which are not necessarily associated with the treatment, that are fatal, threaten the life of the participant, make hospital admission or an extension of the admission necessary, cause persistent or significant invalidity or work disability, manifest themselves in a congenital abnormality or malformation, or could, according to the researchers, have developed to a serious undesired medical event but were prevented because of premature interference.

[d] Analysis adjusted for study site, sex, dose at randomization, and age.

[e] Defined as thyrotropin level of 20 mIU/L or higher during trial laboratory measurements.

[f] Analysis adjusted for study site, sex, and dose at randomization.

[g] Adjusted difference was estimated in linear regression models predicting change from baseline to 12-month visit (95% CI) with study site, sex, and randomization dose as stratification variables and study as random effect.

[h] For scale definitions see Table 1 footnotes.

events were stroke (3 participants [2.2%]), anaemia (2 [1.4%]), and pneumonia (2 [1.4%]). In the placebo group, the most common adverse events were pneumonia (4 participants [3.6%]), cardiac failure (2 [1.8%]), and respiratory failure (2 [1.8%]). At the end of the study, treatment discontinuation occurred in 81 participants (32%) during follow-up, whereas total study withdrawal occurred in 19 (8%) participants. There was no significant difference in treatment discontinuation (38 participants [31%] in the levothyroxine group and 43 [34%] in the placebo) or withdrawal (2 participants [3.8%] in the levothyroxine group and 1 [1.9%] in the placebo group) between the groups.

DISCUSSION

In this prospective analysis that combined data from 2 trials of community-dwelling adults aged 80 years and older with subclinical hypothyroidism, levothyroxine treatment, compared with placebo, was not associated with improvement in hypothyroid symptoms or fatigue. There was no association of levothyroxine treatment with risk of adverse events or secondary outcomes, except for with body mass index and waist circumference; however, the magnitude of these associations was small and likely due to chance given the large number of comparisons. There were no differences in dropout rates between treatment groups, suggesting that levothyroxine treatment was not associated with adverse effects.

These findings do not support routine use of levothyroxine for managing subclinical hypothyroidism in adults aged 80 years and older. Consistent with results reported in the current trial, European and US guidelines do not recommend routine treatment for individuals aged 80 years and older with subclinical hypothyroidism.[31,32] However, in both guidelines, treatment is recommended for individuals with thyrotropin levels of 10 mIU/L or higher (age is not mentioned), and the European guideline suggests ongoing monitoring of thyroid function in patients older than 80 years with subclinical hypothyroidism. Participants in the current study had only mildly elevated thyrotropin levels and a low symptom burden at baseline, consistent with findings in a population study of older individuals with subclinical hypothyroidism.[1] The findings of the present study are thus relevant to the large group of adults aged 80 years and older with few symptoms in whom elevated thyrotropin levels are identified during a routine evaluation.

A Cochrane review of 11 double-blinded randomised clinical trials that examined the effects of thyroid hormone treatment of subclinical hypothyroidism on various outcomes reported no association of thyroid hormone treatment with clinically relevant outcomes.[7] However, these 11 randomised clinical trials of 350 participants included few older participants and had outcomes that were diverse with respect to length of follow-up and type of outcome

studied, such as lipids, mood, and heart function.[7] Results of the current study were consistent with the overall results of the included trial of adults aged 65 years and older, in which levothyroxine treatment was found to have no beneficial effect on thyroid-related symptoms, generic quality of life, cognitive or physical function, and activities of daily living scores, and there was no increase in adverse events.[5]

Limitations

This study has several limitations. First, there were no preplanned subgroup analyses of participants with high symptom burden or higher elevated thyrotropin level at baseline. Results may not apply to these participants. Second, antithyroid antibody status, which may identify individuals who have an increased risk of progression to overt hypothyroidism, was not available. Third, the study population was homogeneous with respect to race. Fourth, there were participants who discontinued treatment (32%), which may have biased results. However, numbers of and reasons for discontinuation were similar between treatment groups.

Conclusions

In this prospectively planned analysis of data from 2 clinical trials involving adults aged 80 years and older with subclinical hypothyroidism, treatment with levothyroxine, compared with placebo, was not significantly associated with improvement in hypothyroid symptoms or fatigue. These findings do not support routine use of levothyroxine for treatment of subclinical hypothyroidism in adults aged 80 years and older.

Online supplemental material

https://jamanetwork.com/journals/jama/fullarticle/2753909

REFERENCES

1. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab.* 2002;87(2):489-499.
2. Biondi B, Palmieri EA, Fazio S, et al. Endogenous subclinical hyperthyroidism affects quality of life and cardiac morphology and function in young and middle-aged patients. *J Clin Endocrinol Metab.* 2000;85(12):4701-4705.
3. Sgarbi JA, Villaca FG, Garbeline B, Villar HE, Romaldini JH. The effects of early antithyroid therapy for endogenous subclinical hyperthyroidism in clinical and heart abnormalities. *J Clin Endocrinol Metab.* 2003;88(4):1672-1677.
4. Rodondi N, Aujesky D, Vittinghoff E, Cornuz J, Bauer DC. Subclinical hypothyroidism and the risk of coronary heart disease: a meta-analysis. *Am J Med.* 2006;119(7):541-551.
5. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.
6. Feller M, Snel M, Moutzouri E, et al. Association of Thyroid Hormone Therapy With Quality of Life and Thyroid-Related Symptoms in Patients With Subclinical Hypothyroidism: A Systematic Review and Meta-analysis. *JAMA.* 2018;320(13):1349-1359.
7. Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. *Cochrane Database Syst Rev.* 2007(3):CD003419.
8. Mooijaart SP, Broekhuizen K, Trompet S, et al. Evidence-based medicine in older patients: how can we do better? *The Netherlands journal of medicine.* 2015;73(5):211-218.
9. den Elzen WP, Lefebre-van de Fliert AA, Virgini V, et al. International variation in GP treatment strategies for subclinical hypothyroidism in older adults: a case-based survey. *Br J Gen Pract.* 2015;65(631):e121-132.
10. Du Puy RS, Postmus I, Stott DJ, et al. Study protocol: a randomised controlled trial on the clinical effects of levothyroxine treatment for subclinical hypothyroidism in people aged 80 years and over. *BMC Endocr Disord.* 2018;18(1):67.
11. Stott DJ, Gussekloo J, Kearney PM, et al. Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST). *BMC Endocr Disord.* 2017;17(1):6.
12. Watt T, Hegedus L, Groenvold M, et al. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(1):161-167.
13. Watt T. Abstract 2054. International Society of Quality of Life Research 19 Annual Conference; October 24-27, 2012; Budapest, Hungary.
14. McClure NS, Sayah FA, Xie F, Luo N, Johnson JA. Instrument-Defined Estimates of the Minimally Important Difference for EQ-5D-5L Index Scores. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2017;20(4):644-650.
15. Zanini A, Aiello M, Adamo D, et al. Estimation of minimal clinically important difference in EQ-5D visual analog scale score after pulmonary rehabilitation in subjects with COPD. *Respiratory care.* 2015;60(1):88-95.
16. Bohannon RW. Minimal clinically important difference for grip strength: a systematic review. *Journal of physical therapy science.* 2019;31(1):75-78.
17. Chen JY, Xu S, Pang HN, et al. Change in Body Mass Index After Total Knee Arthroplasty and Its Influence on Functional Outcome. *The Journal of arthroplasty.* 2018;33(3):718-722.

18. Paans N, Stevens M, Wagenmakers R, et al. Changes in body weight after total hip arthroplasty: short-term and long-term effects. *Physical therapy.* 2012;92(5):680-687.
19. Health NIo. *The Practical Guide to the Identification, Evaluation and Treatment of Overweight and Obesity in Adults.* Vol NIH Publication number 00-4084. Bethesda, MD2000.
20. Muntner P, Shimbo D, Tonelli M, Reynolds K, Arnett DK, Oparil S. The relationship between visit-to-visit variability in systolic blood pressure and all-cause mortality in the general population: findings from NHANES III, 1988 to 1994. *Hypertension (Dallas, Tex : 1979).* 2011;57(2):160-166.
21. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Maryland state medical journal.* 1965;14:61-65.
22. Hsieh YW, Wang CH, Wu SC, Chen PC, Sheu CF, Hsieh CL. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. *Neurorehabilitation and neural repair.* 2007;21(3):233-238.
23. Post MW, van Asbeck FW, van Dijk AJ, Schrijvers AJ. [Dutch interview version of the Barthel Index evaluated in patients with spinal cord injuries]. *Nederlands tijdschrift voor geneeskunde.* 1995;139(27):1376-1380.
24. George LK, Fillenbaum GG. OARS methodology. A decade of experience in geriatric assessment. *J Am Geriatr Soc.* 1985;33(9):607-615.
25. Abdulaziz K, Brehaut J, Taljaard M, et al. National Survey of Emergency Physicians to Define Functional Decline in Elderly Patients with Minor Trauma. *Cjem.* 2015;17(6):639-647.
26. Houx PJ, Shepherd J, Blauw GJ, et al. Testing cognitive function in elderly populations: the PROSPER study. PROspective Study of Pravastatin in the Elderly at Risk. *J Neurol Neurosurg Psychiatry.* 2002;73(4):385-389.
27. Benedict RH, DeLuca J, Phillips G, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England).* 2017;23(5):721-733.
28. Bohannon RW, Glenney SS. Minimal clinically important difference for change in comfortable gait speed of adults with pathology: a systematic review. *Journal of evaluation in clinical practice.* 2014;20(4):295-300.
29. Winther KH, Cramon P, Watt T, et al. Disease-Specific as Well as Generic Quality of Life Is Widely Impacted in Autoimmune Hypothyroidism and Improves during the First Six Months of Levothyroxine Therapy. *PLoS One.* 2016;11(6):e0156925.
30. Carpenter J, Kenward M. *Missing data in randomised controlled trials: a practical guide.* Birmingham: Health Technology Assessment Methodology Programme; 2007.
31. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2012;18(6):988-1028.
32. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. *Eur Thyroid J.* 2013;2(4):215-228.

Supplemental table S1. Characteristics of participants aged 80 years and over at baseline for the IEMO 80-plus Thyroid trial and TRUST Thyroid trial participants separately.

Characteristic	Levothyroxine (N=112)		Placebo (N=139)	
	IEMO 80-plus (N=52)	TRUST 80-plus (N=60)	IEMO 80-plus (N=53)	TRUST 80-plus (N=86)
Age (yrs)				
Mean (SD)	84.5 (3.6)	83.6 (3.1)	85.8 (3.9)	84.5 (3.6)
Range	80.0 – 97.1	80.2 – 93.0	80.4 – 96.7	80.1 – 93.4
Sex				
Male	26 (50.0%)	34 (56.7%)	27 (50.9%)	46 (53.5%)
Female	26 (50.0%)	26 (43.3%)	26 (49.1%)	40 (46.5%)
Self-identified race				
White	52 (100.0%)	60 (100.0%)	52 (98.1%)	85 (98.8%)
Asian	0	0	1 (1.9%)	0
Black	0	0	0	0
Other[a]	0	0	0	1 (1.2%)
Living independently	49 (94.2%)	57 (95.0%)	50 (94.3%)	79 (91.9%)
Previous medical conditions and clinical descriptors				
Hypertension	27 (51.9%)	33 (55.0%)	24 (45.3%)	39 (45.9%)
Ischemic heart disease	13 (25.0%)	10 (16.7%)	12 (22.6%)	26 (30.2%)
Atrial fibrillation	13 (25.0%)	11 (18.6%)	9 (17.0%)	15 (17.4%)
Osteoporosis	9 (17.3%)	7 (12.1%)	11 (20.8%)	12 (14.0%)
Diabetes mellitus	7 (13.5%)	7 (11.7%)	3 (5.7%)	14 (16.5%)
Current smoking	3 (5.8%)	3 (5.0%)	2 (3.8%)	4 (4.7%)
Concomitant medicines	49 (94.2%)	57 (95.0%)	50 (94.3%)	84 (97.7%)
No. of concomitant medications (IQR)[b]	5.0 (3.0, 7.0)	5.0 (3.0, 7.0)	4.0 (2.0, 6.0)	4.5 (3.0, 7.0)
Mini-mental state examination score (IQR)[c]	29.0 (27.0, 30.0)	28.0 (27.0, 29.0)	28.0 (27.0, 29.0)	29.0 (27.0, 29.0)
Weight < 50kg	1 (1.9%)	0 (0.0%)	0 (0.0%)	1 (1.2%)
Laboratory results				
Thyrotropin (mIU/L)[d]				
Mean (SD)	6.4 (1.7)	6.5 (1.9)	6.2 (1.6)	6.3 (2.0)
Median (IQR)	5.7 (5.0, 7.2)	5.9 (5.1, 7.1)	5.7 (5.1, 7.0)	5.7 (5.3, 6.5)
Range	4.6 – 11.5	4.6 – 12.5	4.6 – 12.0	4.6 – 17.6
fT4 (pmol/litre)[e]	14.2 (2.0)	13.6 (2.1)	14.0 (2.5)	13.7 (2.0)
Outcome measures				
ThyPro - Hypothyroid Symptoms score[f]				
Mean (SD)	21.6 (19.2)	20.1 (19.2)	20.5 (21.3)	19.3 (18.1)
Median (IQR)	18.8 (6.2, 37.5)	18.8 (6.2, 26.6)	12.5 (6.2, 31.2)	18.8 (6.2, 31.2)

Supplemental table S1. Characteristics of participants aged 80 years and over at baseline for the IEMO 80-plus Thyroid trial and TRUST Thyroid trial participants separately. (*continued*)

Characteristic	Levothyroxine (N=112)		Placebo (N=139)	
	IEMO 80-plus (N=52)	TRUST 80-plus (N=60)	IEMO 80-plus (N=53)	TRUST 80-plus (N=86)
ThyPro - Tiredness score[f]				
Mean (SD)	24.1 (20.8)	26.4 (21.2)	24.5 (18.5)	26.4 (20.2)
Median (IQR)	20.1 (7.1, 35.7)	21.4 (12.1, 39.3)	17.9 (10.7, 35.7)	21.4 (11.2, 35.7)
Euroqol-5D[g]	0.772 (0.220)	0.776 (0.226)	0.765 (0.275)	0.835 (0.163)
Euroqol VAS[h]	75.3 (14.3)	74.4 (15.4)	73.1 (11.2)	74.3 (16.1)
Hand-grip strength (kg)[i]	25.4 (10.8)	24.3 (7.2)	26.1 (9.3)	23.3 (10.2)
Letter-digit coding test[j]	20.3 (6.9)	22.6 (7.2)	20.6 (7.4)	21.7 (8.2)
Blood pressure				
Systolic (mmHg)	145.8 (21.8)	142.8 (19.1)	153.3 (22.0)	140.5 (18.4)
Diastolic (mmHg)	72.2 (11.2)	69.5 (12.2)	73.6 (13.2)	71.0 (11.0)
Body mass index (kg/m^2)	27.7 (4.7)	27.8 (4.3)	27.5 (4.1)	27.4 (3.8)
Waist circumference (cm)	98.3 (11.8)	99.9 (12.5)	97.7 (11.9)	98.2 (10.5)
Barthel Index (IQR)[k]	20.0 (19.0, 20.0)	20.0 (19.0, 20.0)	20.0 (19.0, 20.0)	20.0 (19.0, 20.0)
Instrumental Activities of Daily Living (IQR)[l]	14.0 (13.0, 14.0)	14.0 (12.8, 14.0)	14.0 (12.0, 14.0)	14.0 (13.0, 14.0)

Results were reported as mean (standard deviation, SD) or median (inter quartile range, IQR) for continuous variables, and as n (%) for proportions.

[a] Participants selected race from 'white', 'asian', 'black' and 5 other pre-specified options reflecting most common race groups in the study site countries

[b] Number of concomitant medications used, both prescribed or over-the-counter available, counted as number of distinct ATC codes excluding emollients and protectives, antiseptics and disinfectants, topical products for joint and muscular pain, nasal preparations, ophthalmologicals or otologicals.

[c] range 0-30, higher scores indicate less cognitive impairment

[d] Inclusion criterium reference range 4.6 - 19.99 mIU/L

[e] Inclusion criterium reference ranges were lab and method specific (all reference ranges within 9 and 26 pmol/L), however repeated within-participant measurements were solely performed using the same method for that individual to exclude any bias due to different assays. SI conversion of fT4 pmol/l to ng/dl, divide by 12.87.

[f] Range 0-100, higher scores represent more symptoms or severity of disease

[g] Range 0.50-1.00, higher scores indicate a better quality of life

[h] Range 0-100, higher scores indicate better health

[i] Higher scores indicate better muscle strength

[j] Number of digits coded within 90 seconds. No upper limit, higher scores indicate better executive cognitive function

[k] Range 0-20, higher scores indicate higher independence in activities of daily living

[l] Range 5-30, higher scores indicate more impairment in instrumental activities of daily living

Abbreviations: fT4, free thyroxine; VAS, Visual Analog Scale.

Supplemental table S2. Characteristics of participants aged 80 years and over at baseline that were included in the primary efficacy analysis.

Characteristic	Total (N=212)	
	Levothyroxine (n=90)	Placebo (n=122)
Age (yrs)		
Mean (SD)	84.1 (3.4)	84.9 (3.7)
Range	80.0 – 97.1	80.1 – 96.7
Sex		
Male	49 (54.4%)	63 (51.6%)
Female	41 (45.6%)	59 (48.4%)
Self-identified race		
White	90 (100.0%)	120 (98.4%)
Asian	0 (0.0%)	1 (0.8%)
Black	0 (0.0%)	0 (0.0%)
Other[a]	0 (0.0%)	1 (0.8%)
Living independently	84.1 (3.4)	84.9 (3.7)
Previous medical conditions and clinical descriptors		
Hypertension	48 (53.3%)	55 (45.5%)
Ischemic heart disease	20 (22.2%)	33 (27.0%)
Atrial fibrillation	15 (16.7%)	22 (18.0%)
Osteoporosis	14 (15.9%)	20 (16.4%)
Diabetes mellitus	11 (12.2%)	15 (12.4%)
Current smoking	5 (5.6%)	5 (4.1%)
Concomitant medicines	84 (93.3%)	117 (93.3%)
No. of concomitant medications (IQR)[b]	5.0 (3.0, 7.0)	4.0 (3.0, 6.8)
Mini-mental state examination score (IQR)[c]	29.0 (27.0, 30.0)	28.0 (27.0, 29.0)
Weight < 50kg	1 (1.1%)	0 (0.0%)
Laboratory results		
Thyrotropin (mIU/L)[d]		
Mean (SD)	6.5 (1.8)	6.2 (1.5)
Median (IQR)	5.8 (5.1, 7.4)	5.7 (5.2, 6.6)
Range	4.6 – 11.9	4.6 – 12.0
fT4 (pmol/litre)[e]	13.8 (2.0)	13.8 (2.1)
Outcome measures		
ThyPro - Hypothyroid Symptoms score[f]		
Mean (SD)	21.7 (19.5)	19.8 (19.6)
Median (IQR)	18.8 (6.2, 35.9)	12.5 (6.2, 31.2)
ThyPro - Tiredness score[f]		
Mean (SD)	25.2 (21.5)	25.1 (19.5)
Median (IQR)	20.1 (7.1, 35.7)	17.9 (10.7, 35.7)

Supplemental table S2. Characteristics of participants aged 80 years and over at baseline that were included in the primary efficacy analysis. (*continued*)

Characteristic	Total (N=212)	
	Levothyroxine (n=90)	Placebo (n=122)
Euroqol-5D[g]	0.8 (0.2)	0.8 (0.2)
Euroqol VAS[h]	75.3 (14.6)	74.0 (14.3)
Hand-grip strength (kg)[i]	25.1 (9.4)	24.8 (10.0)
Letter-digit coding test[j]	21.9 (7.1)	21.8 (7.7)
Blood pressure		
Systolic (mmHg)	25.2 (21.5)	25.1 (19.5)
Diastolic (mmHg)	20.1 (7.1, 35.7)	17.9 (10.7, 35.7)
Body mass index (kg/m^2)	0.8 (0.2)	0.8 (0.2)
Waist circumference (cm)	75.3 (14.6)	74.0 (14.3)
Barthel Index (IQR)[k]	25.1 (9.4)	24.8 (10.0)
Instrumental Activities of Daily Living (IQR)[l]	21.9 (7.1)	21.8 (7.7)

Results were reported as mean (standard deviation, SD) or median (inter quartile range, IQR) for continuous variables, and as n (%) for proportions. Participants selected race from 'white', 'asian', 'black' and 5 other prespecified options reflecting most common race groups in the study site countries

[a] Number of concomitant medications used, both prescribed or over-the-counter available, counted as number of distinct ATC codes excluding emollients and protectives, antiseptics and disinfectants, topical products for joint and muscular pain, nasal preparations, ophthalmologicals or otologicals.

[b] Range 0-30, higher scores indicate less cognitive impairment

[c] Inclusion criterium reference range 4.6 - 19.99 mIU/L

[d] Inclusion criterium reference ranges were lab and method specific (all reference ranges within 9 and 26 pmol/L), however repeated within-participant measurements were solely performed using the same method for that individual to exclude any bias due to different assays. SI conversion of fT4 pmol/L to ng/dl, divide by 12.87.

[e] Range 0-100, higher scores represent more symptoms or severity of disease

[f] Range 0.50-1.00, higher scores indicate a better quality of life. Average in general population for Hypothyroidism Symptoms score is 14 and for Tiredness score 35.

[g] Range 0-100, higher scores indicate better health, [l] Higher scores indicate better muscle strength

[h] Number of digits coded within 90 seconds. No upper limit, higher scores indicate better executive cognitive function

[i] Range 0-20, higher scores indicate higher independence in activities of daily living

[j] Range 5-30, higher scores indicate more impairment in instrumental activities of daily living. Five means functionally completely independent.

Abbreviations: fT4, free thyroxine; VAS, Visual Analogue Scale.

Supplemental table S3. Outcomes at 12 months and end of study for participants of the IEMO 80-plus Thyroid trial and TRUST Thyroid trial, with subclinical hypothyroidism aged 80 years and over, separately

	Baseline					
	IEMO 80-plus		TRUST 80-plus		IEMO 80-plus	
	Levo-thyroxine	Placebo	Levo-thyroxine	Placebo	Levo-thyroxine	Placebo
Variables	**(N=45)**	**(N=46)**	**(N=45)**	**(N=70)**	**(N=45)**	**(N=46)**
Co-primary outcomes						
ThyPro Hypothyroid Symptoms score[b]	23.1 (19.8)	22.1 (21.7)	20.4 (19.4)	18.2 (18.1)	17.2 (16.7)	16.1 (18.4)
ThyPro Tiredness score[b]	23.7 (21.4)	23.3 (18.9)	26.8 (21.7)	26.3 (19.9)	25.9 (20.7)	26.4 (22.1)
Pre-specified secondary outcomes						
Thyrotropin (mIU/L)	6.45 (1.75)	6.28 (1.66)	6.55 (1.87)	6.15 (1.36)	3.76 (1.67)	5.88 (2.66)
median (IQR)	5.70 (5.20, 7.40)	5.66 (5.08, 7.00)	5.82 (5.11, 7.22)	5.74 (5.35, 6.49)	3.40 (2.90, 4.30)	4.98 (4.52, 6.34)
Euroqol-5D[b]	0.774 (0.201)	0.768 (0.261)	0.796 (0.198)	0.838 (0.165)	0.729 (0.296)	0.709 (0.317)
Euroqol VAS[b]	76.22 (12.89)	73.33 (11.15)	74.31 (16.20)	74.39 (16.01)	74.96 (10.81)	73.60 (12.78)
Hand-grip strength (kg)[b]	24.8 (11.0)	26.1 (9.8)	26.2 (7.3)	23.9 (10.4)	22.7 (11.0)	23.2 (9.0)
Blood pressure						
Systolic (mmHg)	146.0 (21.0)	154.7 (22.2)	142.8 (17.7)	140.7 (17.7)	142.6 (20.0)	146.2 (23.3)
Diastolic (mmHg)	72.3 (11.2)	73.8 (13.5)	70.2 (12.4)	71.4 (11.5)	69.1 (11.8)	70.6 (14.6)
Body mass index (kg/m^2)	27.5 (4.9)	27.4 (4.1)	27.8 (3.8)	27.5 (3.8)	27.5 (4.9)	26.8 (3.9)
Waist circumference (cm)	99.0 (11.6)	98.0 (10.9)	100.1 (11.0)	98.5 (10.5)	99.0 (11.4)	96.5 (11.7)
					At end of study	
					Levo-thyroxine	Placebo
Letter-digit coding test[b,c]	20.6 (6.7)	21.5 (7.0)	23.6 (7.3)	22.2 (8.1)	19.8 (7.2)	19.6 (6.6)
Barthel index[b,c]	19.5 (1.3)	19.2 (1.5)	19.2 (1.6)	19.6 (0.9)	19.2 (1.5)	18.6 (2.8)
Instrumental Activities of Daily Living[b,c]	13.4 (0.8)	12.8 (2.0)	13.2 (1.6)	13.5 (1.0)	12.7 (1.7)	12.2 (3.0)

[a] Adjusted difference was estimated in linear regression models predicting change from baseline to 12-month visit (95% CI) with study site, sex and randomization dose as stratification variables and study as random effect.

[b] For scale definitions see Supplemental table 1 footnotes.

[c] Repeat Barthel index, Letter-digit coding test and Instrumental Activities of Daily Living measurements were performed at final visit (i.e. at 12, 24 or 36 months depending on follow-up time), these measurements were included in the estimations.

Abbreviations: VAS, Visual Analog Scale. Scores were presented as means (standard deviation, SD).

At 12 months					
IEMO 80-plus		TRUST 80-plus			
Adjusted difference[a] (95% CI)	p-value	Levo-thyroxine (N=45)	Placebo (N=70)	Adjusted Difference[a] (95% CI)	p-value
0.69 (-5.43, 6.80)	0.82	21.4 (19.6)	18.3 (18.0)	1.27 (-3.95, 6.49)	0.63
-0.78 (-8.26, 6.70)	0.84	30.6 (19.2)	30.2 (18.3)	-0.66 (-6.20, 4.89)	0.82
-2.23 (-3.07, -1.40)	<0.001	3.61 (1.96)	5.23 (1.84)	-1.75 (-2.45, -1.05)	<0.001
-	-	3.51 (2.26, 4.40)	4.99 (4.03, 6.18)	-	-
0.019 (-0.082, 0.120)	0.71	0.780 (0.237)	0.834 (0.167)	-0.026 (-0.074, 0.022)	0.28
0.03 (-4.20, 4.27)	0.99	73.36 (16.12)	73.72 (14.16)	-0.68 (-5.49, 4.13)	0.78
0.40 (-2.16, 2.97)	0.76	24.1 (7.9)	22.9 (9.4)	-0.45 (-2.32, 1.41)	0.63
0.04 (-8.38, 8.46)	0.99	139.9 (18.1)	140.3 (18.7)	-1.39 (-7.42, 4.65)	0.65
-0.66 (-5.13, 3.81)	0.77	68.3 (12.1)	68.9 (11.0)	-0.41 (-4.06, 3.24)	0.82
0.66 (0.21, 1.12)	0.005	27.7 (3.8)	27.3 (3.9)	0.15 (-0.27, 0.58)	0.47
1.52 (0.09, 2.96)	0.04	100.7 (10.7)	97.6 (12.0)	1.42 (-0.75, 3.60)	0.2
At end of study			At end of study		
Adjusted difference[a]	p-value	Levo-thyroxine	Placebo	Adjusted - difference[a]	p-value
1.05 (-0.52, 2.62)	0.19	23.5 (10.8)	20.9 (7.5)	1.64 (-1.00, 4.29)	0.22
0.31 (-0.38, 0.99)	0.38	18.9 (2.2)	19.3 (1.4)	-0.17 (-0.73, 0.39)	0.54
-0.04 (-0.88, 0.80)	0.92	12.1 (2.9)	13.0 (1.7)	-0.64 (-1.37, 0.09)	0.08

Supplemental table S4. Sensitivity analysis of outcomes at 12 months for participants with subclinical hypothyroidism aged 80 years and over.

Variables	Imputed data		Adjusted difference[a]	p-value
	Levothyroxine	Placebo		
Co-primary outcomes				
ThyPRO Hypothyroid Symptoms score[b]	22 (19.6%)	17 (12.2%)	1.79 (-1.60, 5.18)	0.30
ThyPRO Tiredness score[b]	22 (19.6%)	17 (12.2%)	0.84 (-2.99, 4.68)	0.66
Pre-specified secondary outcomes				
Thyrotropin (mIU/L)	22 (19.6%)	23 (16.5%)	-1.74 (-2.18, -1.30)	< 0.001
EuroQol-5D[b]	22 (19.6%)	17 (12.2%)	-0.01 (-0.05, 0.03)	0.67
EuroQol VAS[b]	22 (19.6%)	17 (12.2%)	-0.39 (-3.09, 2.30)	0.77
Hand-grip strength (kg)[b]	23 (22.1%)	20 (14.9%)	0.07 (-1.33, 1.19)	0.92
Blood pressure				
Systolic (mmHg)	22 (19.6%)	17 (12.2%)	-0.37 (-4.46, 3.73)	0.86
Diastolic (mmHg)	22 (19.6%)	17 (12.2%)	-0.17 (-2.52, 2.17)	0.89
Weight	21 (18.9%)	18 (12.9%)	0.85 (0.12, 1.57)	0.022
Body mass index (kg/m^2)	21 (18.9%)	18 (12.9%)	0.33 (0.07, 0.59)	0.012
Waist circumference (cm)	22 (19.6%)	18 (12.9%)	1.29 (0.08, 2.50)	0.037

[a] Adjusted difference was estimated in repeated linear mixed regression models predicting change from baseline to 12-month visit (95% CI) with study site, sex and randomization dose as stratification variables and study as random effect using data imputed from ten imputed data sets, imputing all missing values of the outcome variable at 12-month follow-up from age, sex, baseline thyrotropin, baseline measurement of the outcome variable and, if available, measurement of the outcome variable at 6-8-week follow-up visit.

[b] For scale definitions see Supplemental table 1 footnotes.

Abbreviations: VAS, Visual Analog Scale. Scores were presented as means (SD).

Supplemental table S5. Clinical and adverse events for the IEMO 80-plus Thyroid trial and TRUST Thyroid trial separately

Variables	IEMO 80-plus					TRUST 80-plus				
	Number of events (%)		Event rates (/100 person years)		p-value	Number of events (%)		Event rates (/100 person years)		p-value
	Levothyroxine (n=52)	Placebo (n=53)	Levothyroxine (n=52)	Placebo (n=53)		Levothyroxine (n=60)	Placebo (n=86)	Levothyroxine (n=60)	Placebo (n=86)	
Pre-specified secondary outcomes										
Clinical outcomes										
Fatal or nonfatal cardiovascular event	1 (1.9%)	5 (9.4%)	1.5	7.53	0.11	6 (10.0%)	9 (10.5%)	6.82	7.71	0.92
Death from any cause	2 (3.8%)	4 (7.5%)	2.85	5.86	0.42	3 (5.0%)	0 (0.0%)	3.08	0	0.06
Adverse events										
Cardiovascular death	0 (0.0%)	1 (1.9%)	0	1.46	0.32	0	0		-	
Serious adverse events										
Events	15	17				38	44			
Patients with >1 serious adverse event[a]	12 (23.1%)	13 (24.5%)	0.86 (0.34, 2.22)		0.76	21 (35.0%)	27 (31.4%)		1.29 (0.60, 2.76)	0.51
Adverse event of special interest										
New-onset atrial fibrillation	0 (0.0%)	2 (3.8%)	0	3.07	0.12	4 (6.7%)	4 (4.7%)	4.54	3.26	0.64
Heart failure	1 (1.9%)	3 (5.7%)	1.5	4.53	0.32	2 (3.3%)	3 (3.5%)	2.19	2.48	0.98
Fracture	1 (1.9%)	3 (5.7%)	1.48	4.65	0.3	3 (5.0%)	2 (2.3%)	3.3	1.64	0.54
Hypothyroidism[b]	0 (0.0%)	0 (0.0%)								
Withdrawal										
Permanent discontinuation of trial regimen	16 (30.8%)	18 (34.0%)	28.27	30.9	0.86	22 (36.7%)	25 (29.1%)	26.68	21.77	0.43
Withdrawal from follow-up[c]	2 (3.8%)	1 (1.9%)	2.85	1.47	0.56	8 (13.3%)	8 (9.3%)	8.23	6.19	0.43

Supplemental table S5. Clinical and adverse events for the IEMO 80-plus Thyroid trial and TRUST Thyroid trial separately (*continued*)

Variables	IEMO 80-plus					TRUST 80-plus				
	Number of events (%)		Event rates (/100 person years)		p-value	Number of events (%)		Event rates (/100 person years)		p-value
	Levothyroxine (n=52)	Placebo (n=53)	Levothyroxine (n=52)	Placebo (n=53)		Levothyroxine (n=60)	Placebo (n=86)	Levothyroxine (n=60)	Placebo (n=86)	
	Baseline (mean, SD)		At 12 months (mean, SD)			Baseline (mean, SD)		At 12 months (mean, SD)		
	Levothyroxine	Placebo	Levothyroxine	Placebo	p-value	Levothyroxine	Placebo	Levothyroxine	Placebo	p-value
ThyPro Hyperthyroid Symptoms scores[d]	11.4 (13.8)	7.2 (8.6)	7.8 (9.8)	9.1 (9.9)	0.06	10.4 (8.2)	10.4 (11.9)	11.4 (8.8)	9.5 (10.7)	0.31

Pre-planned secondary outcomes effects were estimated using cox-proportional hazard regression models adjusted for sex, dose at randomization and study site. Adverse events were estimated using
cox-proportional hazard regression models with, where possible, adjustment for study site, dose at randomization, sex and age, presented as event rates and p-values.

[a] For outcomes with too few events to run regression models, odds ratio (95% CI) and p-value were reported. Serious adverse events were all undesired medical events involving a participant, which are not necessarily associated with the

[b] treatment, that are: fatal, threaten the life of the subject, make hospital admission or an extension of the admission necessary, cause persistent or significant invalidity or work disability, manifest themselves in a congenital abnormality or

[c] malformation or could, according to the researchers, have developed to a serious undesired medical event, but were however prevented due to premature interference.

[d] Defined as thyrotropin level of 20 mIU/L or higher during trial laboratory measurements

[e] Analysis adjusted for study site, sex, and dose at randomization.

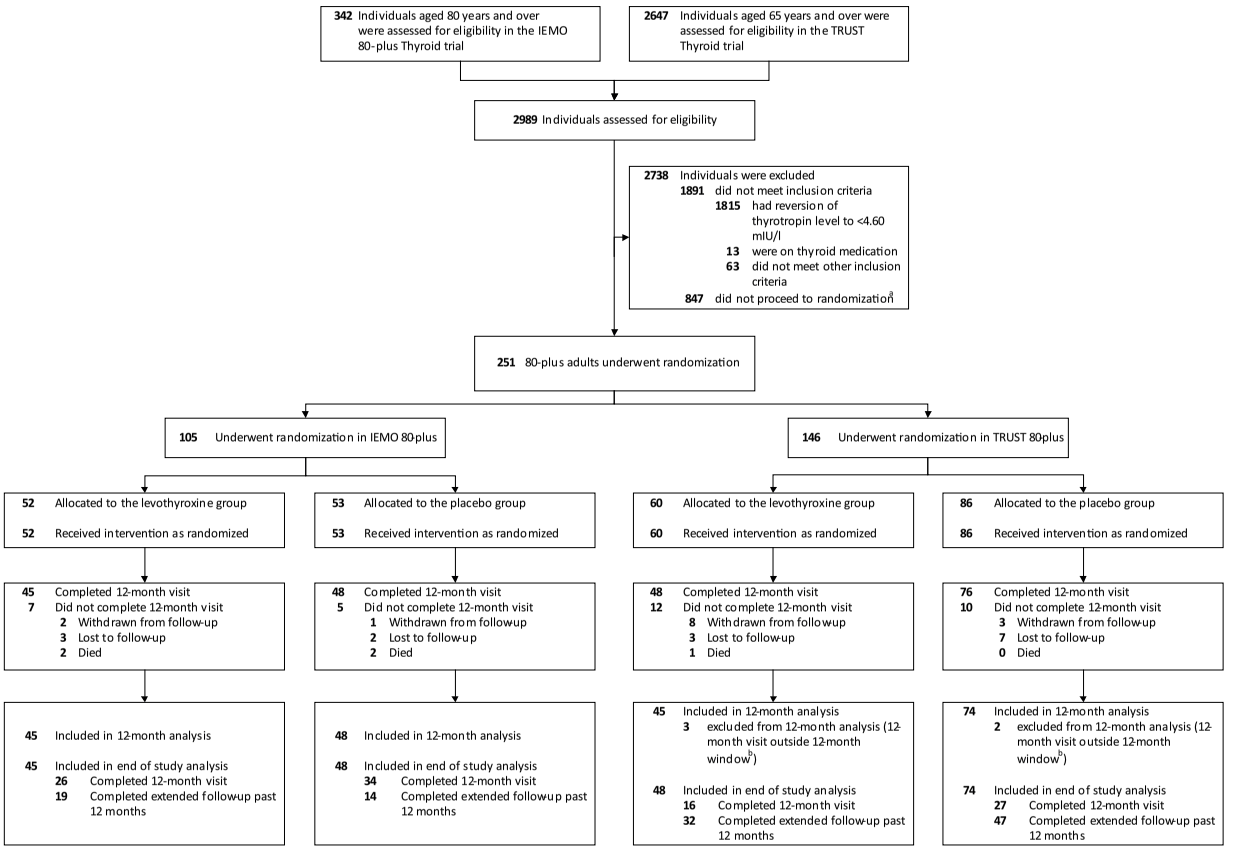
[f] For scale definitions see Supplemental table 1 footnotes. Scores were presented as means (standard deviation, SD)

IN REPLY

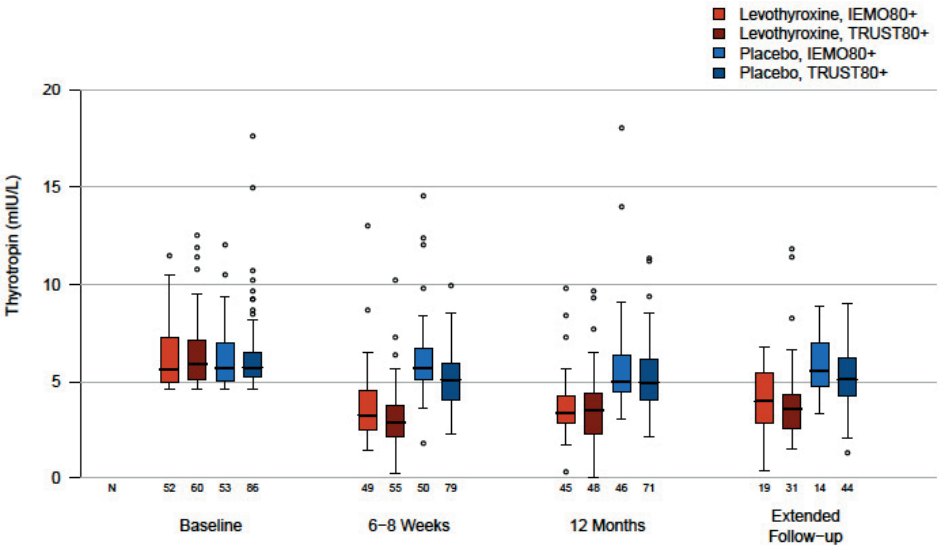
The population distribution of thyrotropin across the human life span has been reported in multiple studies with ambiguous results: both increase and decrease with increasing age have been reported. These inconsistencies may reflect differences in iodine intake across cohorts or age categories and underscore that population distributions alone are unfit to determine reference ranges. Other observational studies of subclinical hypothyroidism, including the NHANES study mentioned by Dr Hershman as well as an individual patient data meta-analysis by Rodondi et al,[1] showed no association of subclinical hypothyroidism with mortality and cardiovascular events in individuals aged 80 years and older. However, because these data have all been observational, they are not suitable to definitely establish or rule out causality. Therefore, until now, existing studies have not justified age-specific reference ranges and guidelines currently do not support them.[2,3]

In our study of 2 randomised clinical trials,[4] we included 251 individuals aged 80 years and older with persistently elevated thyrotropin levels higher than 4.5 mIU/L, according to generally accepted clinical guidelines. We observed no beneficial effects of supplementation, which is in agreement with the overall Thyroid Hormone Replacement for Untreated Older Adults With Subclinical Hypothyroidism Trial (TRUST) of 737 individuals aged 65 years and older,[5] of whom our study included only the 146 individuals aged 80 years and older. Furthermore, in subgroup analyses of TRUST, such as comparing individuals with a thyrotropin level below or above 7 mIU/L, the estimates were similar. These findings provide the highest level of evidence to date regarding thyroid hormone supplementation in older patients with subclinical hypothyroidism.

The definition of reference range and the limit for treatment are separate. We agree that there seems to be no treatment indication for patients aged 80 years and older with elevated thyrotropin levels and normal free thyroxine levels. In an Editorial accompanying our study, Cappola [6] suggested that the upper limit of the reference range for thyrotropin in individuals aged 80 years and older should be changed to 7 mIU/L. Although this is an arbitrary value, it removes the majority of older individuals with elevated thyrotropin from having an indication for supplementation, justified by our experimental evidence.



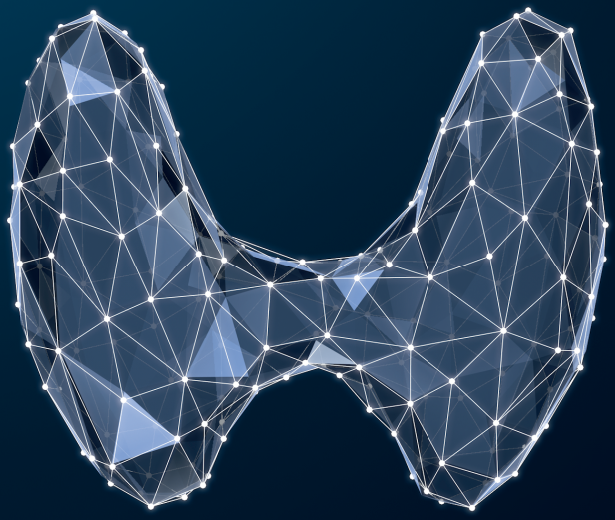
Supplemental figure 1. Recruitment, randomization, and patient flow of the participants age 80 years and over in the TRUST and IEMO thyroid trials separately. [a] These participants completed the screening phase and met the inclusion criteria, yet did not proceed to randomization. For example, participant opted not to participate due to developments in the private sphere or after consultation with children. [b] 12-month window was between 334 and 397 days after randomization



Supplemental figure 2. Thyrotropin levels in the levothyroxine and placebo group during the studies. Abbreviations: N, number of participants with available thyrotropin measurements. Tukey boxplots depicting medians, quartiles, the lowest and upper datum still within 1.5 IQR of the corresponding quartile and outliers.

REFERENCES

1. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA.* 2010;304(12):1365-1374.
2. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2012;18(6):988-1028.
3. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. *Eur Thyroid J.* 2013;2(4):215-228.
4. Mooijaart SP, Du Puy RS, Stott DJ, et al. Association Between Levothyroxine Treatment and Thyroid-Related Symptoms Among Adults Aged 80 Years and Older With Subclinical Hypothyroidism. *JAMA.* 2019;322(20):1-11.
5. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.
6. Cappola AR. The Thyrotropin Reference Range Should Be Changed in Older Patients. *JAMA.* 2019.



Chapter 8

The effect of levothyroxine treatment on haemoglobin levels in older adults with subclinical hypothyroidism: pooled individual results from two randomized controlled trials

Robert S Du Puy
Rosalinde KE Poortvliet
Simon P Mooijaart
David J Stott
Terry Quinn
Naveed Sattar
Rudi GJ Westendorp
Patricia M Kearney
Vera JC McCarthy
Stephen Byrne
Nicolas Rodondi
Oliver Baretella
Tinh-Hai Collet
Diana van Heemst
Olaf M Dekkers
J Wouter Jukema
Jan WA Smit
Jacobijn Gussekloo
Wendy PJ den Elzen

Submitted

ABSTRACT

Background

Subclinical thyroid dysfunction and anaemia are common disorders, and both have increasing prevalence numbers with advancing age. The aim of this study was to assess whether levothyroxine treatment leads to a rise in haemoglobin levels in older persons with subclinical hypothyroidism.

Methods

In a pre-planned combined analysis of two randomised controlled trials of community-dwelling persons with subclinical hypothyroidism, we included participants aged 65 years and older and randomly assigned them to levothyroxine or placebo treatment. The dose of levothyroxine was periodically titrated aiming at thyroid stimulating hormone (TSH) level within the reference range and mock titrations in the placebo group. The outcome of the present analysis was the change in haemoglobin level after 12 months.

Results

Analyses included 669 participants (placebo n=337, levothyroxine n=332) with a median age of 75 years (range 65 to 97) and mean baseline haemoglobin of 13.8 g/dL (standard deviation 1.3). Although levothyroxine treatment resulted in a reduction in TSH from baseline after 12 months of follow-up compared to placebo (treatment effect -1.98 mIU/L, 95%CI -2.30 to -1.66), the change in haemoglobin level was not different between the levothyroxine and the placebo groups (-0.03 g/dL [95%CI -0.16 to 0.11). Similar results were found in stratified analyses including sex, age or TSH levels. No difference in change of haemoglobin levels after 12 months was identified in 69 participants with anaemia at baseline (-0.33 g/dL [95% CI -0.87 to 0.21]).

Conclusions

In persons aged 65 years and older with subclinical hypothyroidism, treatment with levothyroxine does not lead to a rise in haemoglobin levels, regardless of the presence of anaemia.

Trial registration

TRUST: Clinicaltrials.gov: NCT01660126, https://clinicaltrials.gov/ct2/show/NCT01660126

IEMO: Netherlands Trial Register: NTR3851, https://www.trialregister.nl/trial/3681

INTRODUCTION

Subclinical thyroid dysfunction and anaemia are common disorders, and both have increasing prevalence numbers with advancing age.[1-4] The symptoms of both subclinical hypothyroidism and anaemia are frequently non-specific and of similar nature (e.g. fatigue, malaise, shortness of breath and exercise intolerance).

A number of studies have suggested a potential causal relationship between thyroid dysfunction and haematopoiesis.[5-8] Lower haemoglobin levels have been observed in persons with (subclinical) hypothyroidism compared to their euthyroid counterparts and small cohort studies have shown increases in haemoglobin and erythropoietin levels after treatment of overt hypothyroidism.[5,9-18]

We reported earlier the results of an individual participant data meta-analysis using data from 16 independent observational cohort studies including more than 42,000 participants. In participants with both overt and subclinical hypothyroidism, we demonstrated a higher anaemia prevalence.[19] In addition, reduced thyroid function at baseline was associated with an increased risk of anaemia during follow-up. To date, however, there is a lack of experimental evidence of the effect of levothyroxine treatment on haemoglobin levels in older persons with subclinical hypothyroidism.

The aim of this study was to assess whether levothyroxine treatment for subclinical hypothyroidism leads to a rise in haemoglobin levels in older persons. We performed a pre-planned secondary outcome analysis of two randomised controlled trials on levothyroxine treatment for subclinical hypothyroidism in older persons.

MATERIALS AND METHODS

The current pre-planned combined analysis uses data from the TRUST trial [20,21] and the IEMO 80-plus Thyroid trial.[22] Both trials were randomised, double-blind placebo-controlled parallel group trials investigating the potential multi-modal effects of levothyroxine treatment for persons with subclinical hypothyroidism aged 65 years and older, and aged 80 years and older, respectively. The identical design, study processes and infrastructure allowed for a pre-planned combined analysis. Consequently, the combined cohorts are presented and analysed as a single study group throughout the manuscript.

The full study protocols and statistical analysis plans for the TRUST Thyroid trial [20,21] and the IEMO 80-plus Thyroid trial [22] have been published previously. In short, persons aged

65 years and older (TRUST Thyroid Trial) or 80 years and older (IEMO 80-plus Thyroid Trial) with a diagnosis of subclinical hypothyroidism, defined as elevated TSH levels (≥ 4.6 and ≤ 19.9 mIU/L), measured on at least two occasions with an interval between > three months and three years apart, and free thyroxine (fT4) within normal local laboratory ranges, were enrolled. Exclusion criteria included, but were not limited to, the use of levothyroxine, antithyroid medication, clinical diagnosis of dementia, recent hospitalization for major illness and terminal illness.[20] Participants were randomised in a 1:1 ratio to levothyroxine or placebo.

The study medication consisted of levothyroxine sodium tablets and matching placebo tablets taken orally daily. The intervention group started with 50 micrograms daily (or 25 micrograms for participants with body weight < 50 kg or a history of coronary heart disease) and the control group with a matching placebo for six to eight weeks. Two optional up- or down-titrations after six-to-eight-week intervals, and one at 12 months of follow-up ensured adequate treatment while avoiding potential over-treatment, mirroring routine clinical practice. An adaptive mock titration schedule was applied for the placebo group. Participants, General Practitioners (GPs) and study personnel remained blinded to treatment allocation and thyroid function test results throughout the study. Ethics approval was obtained from the medical ethics committees and competent authorities in the Netherlands, United Kingdom, Switzerland, and Ireland. Written informed consent was obtained from all participants.

Study outcome

The predetermined outcome of the present analysis was the change from baseline haemoglobin levels after 12 months of treatment. Haemoglobin levels were measured as part of a complete blood count in automated analysis systems in EDTA anticoagulated blood samples and measured within four hours of sample collection (the Netherlands: Sysmex XN10/20; Scotland: Sysmex XN10; Switzerland: Siemens Advia; Ireland: Advia 2120i until October 2015, and afterwards Sysmex XN-9000), at baseline and after 12 months. Anaemia was defined according to the World Health Organization criteria (Hb < 12 g/dL for women and Hb < 13 g/dL for men) and anaemia severity as mild (women: 11.0 ≤ Hb ≤ 11.9 g/dL, men: 11.0 ≤ Hb ≤ 12.9 g/dL), moderate (all: 8.0 ≤ Hb ≤ 10.9 g/dL) or severe (all: Hb < 8.0 g/dL).[23]

Additional variables

Thyroid function tests (TSH and fT4 levels) were performed by certified medical laboratories at each site (the Netherlands: Elecsys 2010 (Roche), Architect (Abbott), UniCel DxL (Beckman), Immulite 2000XPI (Siemens) or Vitros ECoQ (Ortho Clinical Diagnostics); United Kingdom: Architect (Abbott); Switzerland: Elecsys 2010 (Roche) and Ireland: Cobas 8001 or E601 (Roche) or Architect (Abbott)).[20,22] fT4 levels at 12 months were available for a subsample of participants (total n=115, United Kingdom n=75, the Netherlands n=37, Ireland n=3). Socio-demographic data (age, sex and race), information on lifestyle factors (smoking

and alcohol intake), the presence of prior conditions and medication use (prescribed and over-the-counter) were recorded by research nurses during baseline and follow-up visits. Comorbidity was defined as having a history of one or more of the following diagnoses: myocardial infarction, angina, stroke, transient ischemic attacks, heart failure, peripheral vascular disease, revascularization, atrial fibrillation, hypertension, diabetes mellitus, epilepsy, dementia, osteoporosis or other disease.[20,22]

Statistical methods

To investigate cross-sectional associations between thyroid hormones and haemoglobin or anaemia status, linear mixed effect regression models and logistic regression models were used, adjusting for possible confounders age and sex at baseline. Treatment effects were calculated using linear mixed effect regression models (differences in the change of haemoglobin levels between baseline and at 12 months) and logistic regression models (odds of developing anaemia or resolving anaemia during follow-up), adjusting for study site, study, treatment dose at randomization, age, sex, smoking status and alcohol consumption, with treatment (levothyroxine or placebo) as the independent variable, haemoglobin level or anaemia (stratified for sex) as the dependent variable, age and alcohol as covariates and sex and smoking status as fixed factors. All confounding variables were assessed for statistical interaction with all p-for-interactions above the 0.05 significance threshold.

Subgroup analyses were performed, stratified by sex, age (< 80 vs. ≥ 80 years), TSH level by quartiles (4.6 to 5.10 mIU/L, 5.11 to 5.76 mIU/L, 5.77 to 6.99 mIU/L and >7 mIU/L), haemoglobin level by quartiles (men: 9.35 to 13.70, 13.71 to 14.50, 14.51 to 15.20 and >15.21 g/dL; women: 9.67 to 12.73, 12.74 to 13.40, 13.41 to 14.02 and >14.02 g/dL) and the presence of anaemia at baseline. Differences in median haemoglobin levels at baseline were analysed across quartiles of TSH and fT4 using Kruskall-Wallis tests. Analyses of the treatment effects were performed in the modified intention to treat (ITT) population, based on participants with data available for haemoglobin levels. A per protocol analysis in participants with uninterrupted use of study medication throughout the studies was included as a sensitivity analysis. Analyses were repeated while excluding participants using concomitant anti-anaemic medications (iron supplements, parenteral iron preparations, vitamin B12, folic acid, erythropoietin, or other anti-anaemic drug, n=47). All data analyses were performed with SPSS Statistics Software version 22.0 for Windows (IBM, Armond, NY, USA). A p-value less than 0.05 was considered statistically significant.

RESULTS

For both studies, 2,989 individuals were assessed for eligibility and a total of 842 participants (737 in TRUST, 105 in IEMO) were randomised in the main studies. For the present analysis, 669 participants (592 in TRUST, 77 in IEMO) whose haemoglobin levels were measured at baseline were included (placebo n=337, levothyroxine n=332, Figure 1). Table 1 presents the baseline characteristics of the study participants. The median age of the population was 75 years (range 65 to 97) in both treatment groups. Men constituted 47.5% and 49.1% of all participants in the placebo and levothyroxine groups, respectively. More than 83% of the study population had one or more comorbid condition. The mean haemoglobin level in the placebo group was 13.8 g/dL (standard deviation [SD] 1.3; mean in women 13.4 g/dL [SD 1.1], mean in men 14.3 g/dL [SD 1.3]) and in the levothyroxine group 13.8 g/dL (SD 1.3; mean in women 13.4 g/dL [SD 1.1]; mean in men 14.3 g/dL [SD 1.4]). At baseline, 13.1% of the participants had anaemia in the placebo group; 11.4% had anaemia in the levothyroxine group.

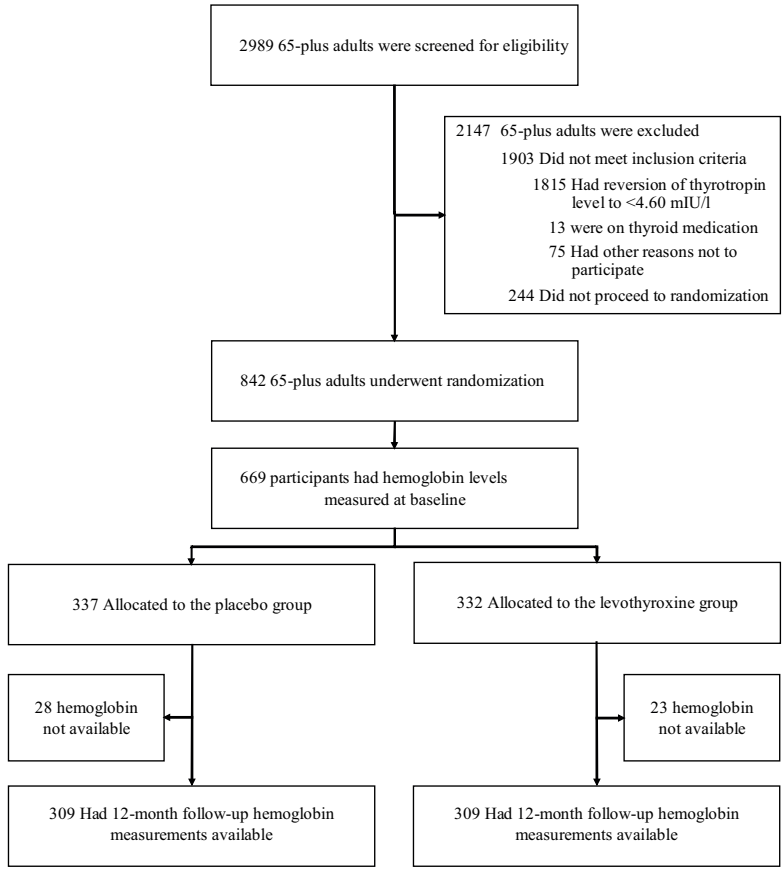


Figure 1. Recruitment, Randomization, and Patient Flow of the Participants

Table 1. Baseline characteristics of the study participants

Characteristic	Placebo group	Levothyroxine group
	n=337	n=332
Socio-demographic characteristics:		
Age, median (range)	75 (65 - 97)	75 (65 - 93)
Male, no. (%)	160 (47.5)	163 (49.1)
MMSE, median (IQR)	29 (28 - 30)	29 (27 - 29)
Barthel Index, mean (SD)	19.6 (0.9)	19.7 (0.9)
Instrumental activities of daily living score, mean (SD)	13.6 (1.1)	13.7 (0.8)
Handgrip strength, kg	27.6 (11.1)	28.3 (10.6)
Body mass index, kg/m^2	27.6 (4.4)	28.1 (5.2)
Systolic blood pressure, mmHg	143.3 (19.3)	141.6 (19.0)
Diastolic blood pressure, mmHg	75.4 (11.9)	74.1 (11.6)
Current smoker, no. (%)	24 (7.1)	25 (7.5)
Current alcohol user, no. (%)[a]	213 (63.2)	187 (56.3)
Comorbidities, no. (%)[b]	280 (83.1)	282 (84.9)
0-1	161 (47.8)	144 (43.4)
2-4	160 (47.5)	167 (50.3)
4+	16 (4.7)	21 (6.3)
Caucasian, no. (%)	331 (98.2)	328 (98.8)
Concurrent anti-anaemic medication use, no. (%)[c]	21 (6.2)	26 (7.8)
Thyroid function tests		
TSH, median (IQR), mIU/L	5.8 (5.1 - 6.8)	5.8 (5.1 - 7.1)
fT4, pmol/L	13.4 (1.9)	13.4 (2.0)
Haematological parameters		
Haemoglobin, g/dL[d]	13.8 (1.3)	13.8 (1.3)
Men	14.3 (1.3)	14.3 (1.4)
Women	13.4 (1.1)	13.4 (1.1)
Anaemia status, no. (%)[e]		
No anaemia	293 (86.9)	294 (88.6)
Mild anaemia	36 (10.7)	31 (9.3)
Moderate anaemia	8 (2.4)	7 (2.1)
Severe anaemia	0 (0)	0 (0)

Continuous data are presented as mean (standard deviation) and, if stated otherwise, as median (interquartile range); categorical data are presented as number (percentage)

[a] Defined as an average of one or more self-reported consumed units of alcohol per week.

[b] Comorbidity was defined as having a history of one or more of the following: myocardial infarction, angina, stroke, transient ischemic attacks, heart failure, peripheral vascular disease, revascularisation, atrial fibrillation, hypertension, diabetes mellitus, epilepsy, dementia, osteoporosis or other.

[c] Concurrent medication use was defined as using one or more of the following medications: iron supplements, parenteral iron preparations, vitamin B12, folic acid, erythropoietin, or other anti-anaemic drugs.

[d] Conversion factor of haemoglobin from g/dL to mmol/L = value*0.6206.

[e] Anaemia defined by WHO categories: women haemoglobin [Hb] < 12.0 g/dL, men Hb < 13.0 g/dL. Anaemia severity was graded as: mild (women: 11.0 ≤ Hb ≤ 11.9 g/dL, men: 11.0 ≤ Hb ≤ 12.9 g/dL), moderate (all: 8.0 ≤ Hb ≤ 10.9 g/dL), severe (all: Hb < 8.0 g/dL).

Table 2 shows the relation between thyroid function and haematological parameters at baseline. No association was observed between TSH or fT4 levels and haemoglobin levels in univariable and multivariable regression models. Although a small association between higher fT4 and the odds of having anaemia was observed at baseline (odds ratio [OR] 1.13, 95%CI 1.01 to 1.26, p=0.041), in the multivariable models higher levels of TSH or fT4 were not associated with increased odds of having anaemia. Median haemoglobin levels at baseline were comparable across quartiles of TSH and fT4 (Figure 2).

Table 2. Baseline associations between thyroid function and hematologic parameters in 669 older participants with subclinical hypothyroidism.

	Estimate	95% CI	p-value
Thyroid function and haemoglobin	**Beta**		
TSH			
Univariable	-0.01	-0.06 to 0.04	0.649
Multivariable	-0.01	-0.06 to 0.04	0.736
fT4			
Univariable	-0.05	-0.10 to 0.00	0.065
Multivariable	-0.03	-0.07 to 0.02	0.267
Thyroid function and anaemia (n=82)	**Odds Ratio**		
TSH			
Univariable	1.04	0.93 to 1.16	0.523
Multivariable	1.03	0.91 to 1.17	0.595
fT4			
Univariable	1.13	1.01 to 1.26	0.041
Multivariable	1.08	0.96 to 1.22	0.200

Betas and corresponding confidence intervals were calculated using linear mixed effect regression models and represent the difference in haemoglobin level (g/dL) per unit increase in TSH (mIU/L) or fT4 (pmol/L), Odds ratios with corresponding confidence intervals were calculated using logistic regression models, per unit increase in TSH or fT4. Multivariable models were adjusted for possible confounders age and sex.
Units: TSH mIU/L; fT4 pmol/L.

Although significant reductions in TSH (-1.98 mIU/L, 95% CI -2.30 to -1.66, $p < 0.001$) and increases in fT4 (2.33, 95% CI 1.87 to 2.80, $p < 0.001$) over 12 months of follow-up were observed in the levothyroxine group compared to the placebo group, no differences between the groups were observed in the change in haemoglobin levels after 12 months of follow-up (between group difference -0.03 g/dL, 95% CI -0.16 to 0.11, p=0.703, Figure 3).

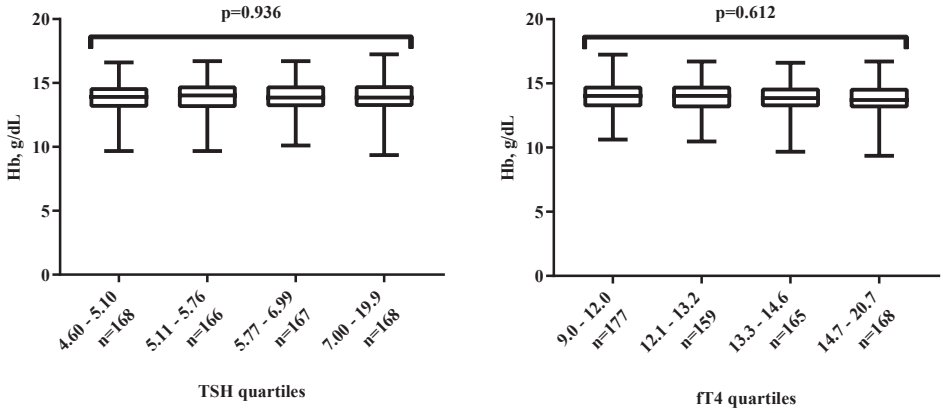


Figure 2. Baseline Hb levels per quartiles of TSH and fT4
Legend: Boxplots represent the median, interquartile range, minimum and maximum in Hb levels per quartile of TSH (mIU/L) and fT4 (pmol/L). Difference in medians were evaluated using Kruskall-Wallis test.

Results were not substantially different in the per protocol analysis for participants that remained on the assigned treatment after 12 months and had no major protocol violations (haemoglobin treatment effect -0.05 g/dL, 95% CI -0.19 to 0.10, p=0.530). Similar results were also found when subgroups were stratified by sex, age, quartiles of TSH levels, quartiles of haemoglobin levels in men and women, or the presence of anaemia at baseline (Table 3).

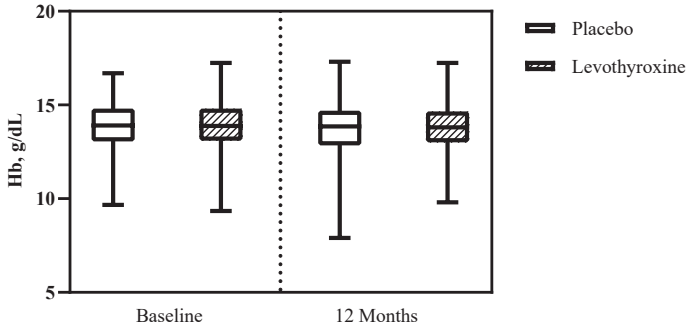


Figure 3. Levothyroxine treatment effect on Hb levels after 12 months of treatment
Legend: Boxplots represent the median, interquartile range, minimum and maximum in Hb levels. Treatment effect -0.03 g/dL (95%CI -0.16 to 0.11). Treatment effect (95% CI) was calculated using linear mixed effect regression models adjusting for study site, study, treatment dose at randomization, age and sex and represents the additional change in haemoglobin level after 12 months of treatment with levothyroxine, compared to placebo.

Table 3. Changes in TSH, fT4 and haemoglobin levels in relevant subgroups in older participants with subclinical hypothyroidism on levothyroxine treatment, compared to placebo.

		Baseline		12-month follow-up			
	N	Placebo	Levothyroxine	Placebo	Levothyroxine	Treatment effect (95%CI)[a]	p-value
Effect on thyroid hormone levels[b]							
TSH, mIU/L[b]	660	6.3 (1.7)	6.5 (2.1)	5.5 (2.5)	3.6 (2.1)	-1.98 (-2.30 to -1.66)	< 0.001
fT4, pmol/L [b]	177	13.4 (1.9)	13.4 (2.0)	12.4 (1.7)	14.7 (2.8)	2.33 (1.87 to 2.80)	< 0.001
Effect on haemoglobin levels							
Sex							
Men	298	14.3 (1.3)	14.3 (1.4)	14.2 (1.5)	14.2 (1.5)	-0.05 (-0.26 to 0.15)	0.599
Women	320	13.4 (1.1)	13.4 (1.1)	13.4 (1.3)	13.4 (1.1)	0.01 (-0.16 to 0.17)	0.950
Age							
< 80 years	466	14.0 (1.2)	13.9 (1.3)	14.0 (1.3)	13.8 (1.3)	-0.06 (-0.20 to 0.08)	0.412
≥ 80 years	152	13.4 (1.4)	13.6 (1.4)	13.3 (1.5)	13.5 (1.5)	0.10 (-0.23 to 0.44)	0.540
TSH							
4.60 – 5.10	157	13.8 (1.2)	13.8 (1.3)	13.6 (1.3)	13.6 (1.4)	-0.09 (-0.32 to 0.13)	0.418
5.11 – 5.76	152	13.8 (1.4)	13.9 (1.2)	13.8 (1.7)	13.9 (1.3)	-0.03 (-0.35 to 0.29)	0.863
5.77 – 6.99	158	13.9 (1.2)	13.8 (1.3)	13.8 (1.5)	13.9 (1.4)	0.16 (-0.08 to 0.41)	0.188
> 7.0	151	13.8 (1.3)	13.8 (1.5)	13.8 (1.2)	13.8 (1.3)	-0.13 (-0.40 to 0.14)	0.336
Haemoglobin quartiles split by sex							
Men							
9.35 – 13.70	67	12.4 (0.8)	12.3 (1.1)	12.5 (1.2)	12.5 (1.4)	-0.26 (-0.95 to 0.44)	0.455
13.71 – 14.50	93	14.2 (0.3)	14.1 (0.3)	14.3 (0.7)	14.0 (0.7)	-0.20 (-0.70 to 0.30)	0.419
14.51 – 15.20	60	15.0 (0.2)	14.9 (0.2)	14.4 (1.5)	14.8 (0.7)	-0.09 (-0.52 to 0.34)	0.680

Table 3. Changes in TSH, fT4 and haemoglobin levels in relevant subgroups in older participants with subclinical hypothyroidism on levothyroxine treatment, compared to placebo. (*continued*)

		Baseline		12-month follow-up			
15.21 – 17.24	73	15.8 (0.4)	15.8 (0.5)	15.5 (0.8)	15.4 (1.1)	0.03 (-0.30 to 0.35)	0.877
Women							
9.67 – 12.73	79	11.9 (0.8)	12.1 (0.5)	12.0 (1.3)	12.1 (0.9)	0.04 (-0.35 to 0.43)	0.842
12.74 – 13.40	84	13.1 (0.2)	13.2 (0.2)	13.0 (0.8)	13.3 (0.6)	0.02 (-0.21 to 0.25)	0.858
13.41 – 14.02	76	13.8 (0.2)	13.7 (0.2)	13.8 (0.5)	13.7 (0.7)	-0.19 (-0.55 to 0.18)	0.317
14.03 – 16.30	81	14.7 (0.4)	14.7 (0.6)	14.6 (0.9)	14.4 (0.8)	-0.13 (-0.55 to 0.29)	0.540
Anaemia at baseline							
No anaemia	549	14.1 (1.0)	14.1 (1.1)	14.0 (1.2)	14.0 (1.2)	0.01 (-0.12 to 0.15)	0.886
Anaemia	69	11.6 (0.8)	11.6 (0.8)	11.9 (1.3)	11.8 (1.2)	-0.33 (-0.87 to 0.21)	0.233

[a] Treatment effects were calculated using linear mixed effect regression models adjusting for study site, study, treatment dose at randomization, age and sex and represent the additional change in haemoglobin level (g/dL) after 12 months of treatment with levothyroxine, compared to placebo. In stratified analysis the stratifying variable was not an adjusting variable for that analysis.

[b] Treatment effects represent the additional change in TSH (mIU/L) or fT4 (pmol/L) after 12 months of treatment with levothyroxine, compared to placebo. Anaemia defined by WHO categories: Women haemoglobin (Hb) <12.0 g/dL, men Hb <13.0 g/dL.

In the 618 (92.4%) participants for whom haemoglobin measurements were available at 12 months, levothyroxine treatment was not associated with decreased odds of developing anaemia (placebo n=17/272, levothyroxine n=19/277, OR 1.16 [95% CI 0.59 to 2.29], p=0.675) or increased odds of resolution of anaemia (placebo n=11/37, levothyroxine n=7/32, OR 1.80 [95% CI 0.54 to 5.93], p=0.337). Similar results were found in sensitivity analyses restricted to participants with anaemia without anti-anaemic medication (data not shown).

No confounding variables in any of the models demonstrated significant statistical interaction (data not shown).

DISCUSSION

In this combined analysis of two randomised trials of older adults with a diagnosis of subclinical hypothyroidism, levothyroxine treatment for 12 months was not associated with an increase in haemoglobin levels. Additionally, no changes in haemoglobin levels were observed in relevant subgroup analyses including sex, age and baseline haemoglobin levels. No baseline associations were identified between TSH and haemoglobin levels or the presence of anaemia. A clinically insignificant increased odds of having anaemia at baseline was identified for fT4 in the univariable analysis, that was no longer present when correcting for the influence of sex and age, and is regarded as a chance finding.

These findings are in contrast with differences in haemoglobin levels between persons with different levels of thyroid function identified in earlier observational studies,[5,9-11,14,15,17-19] but in line with recent systematic review and cohort studies.[24,25] The addition of the experimental results from this manuscript underpin the proposition that subclinical hypothyroidism and anaemia may not be causally related. In line with earlier results from the TRUST and IEMO trials, in which no beneficial effects of levothyroxine treatment were demonstrated for a range of clinically relevant outcomes including thyroid-specific and generic quality of life, grip strength, blood pressure and body mass index,[21,26] the lack of a beneficial effect on haemoglobin levels is an added finding suggesting a limited clinical value of treating subclinical hypothyroidism with levothyroxine mono-therapy in older persons.

Interestingly, Christ-Crain and colleagues observed increases in erythropoietin levels upon levothyroxine treatment in a small placebo-controlled RCT of women (mean age 59 years [SD 1]) with subclinical hypothyroidism, while haemoglobin levels and haematocrit remained unchanged after 48 weeks of treatment.[27] One may hypothesize that effects of thyroid function on haemoglobin levels may only become apparent in those with overt hypothyroidism or severe anaemia.[6-8,15,28-30] Indeed, a number of studies have shown a beneficial

effect of thyroid hormone treatment in patients with overt hypothyroidism on erythropoietin levels.[9,14,17] Nevertheless, the absence of change in haemoglobin levels in our study – even in those with the lowest haemoglobin levels - may suggest that in older persons with subclinical hypothyroidism hematopoietic processes are quite robust to changes such as levothyroxine treatment.

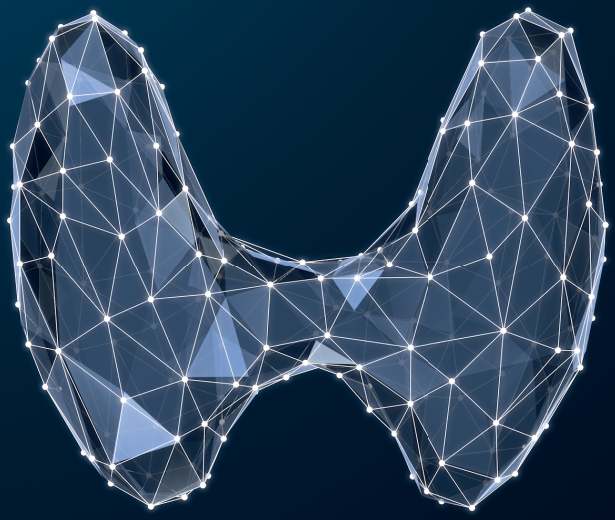
The present study used data from the largest randomised controlled trials (RCTs) to date on levothyroxine treatment for subclinical hypothyroidism in community-dwelling older adults, with 12 months of follow-up, but some limitations must be acknowledged. First, despite the sample size of these RCTs, few participants (n=21) had a baseline TSH level of more than 10 mIU/L, i.e. the upper end of the subclinical hypothyroid spectrum, in which a few earlier studies identified additional risks of unwanted health effects.[4] The majority of participants in our RCTs had mild subclinical hypothyroidism (TSH between 4.6 and 10.0 mIU/L). Second, the number of participants with anaemia at baseline was rather low, leading to insufficient power to study the effects of levothyroxine on haemoglobin in those with anaemia. However, an absence of treatment effect was consistent across stratifications based on quartiles of haemoglobin levels at baseline. Third, fT4 levels were not routinely measured at 12 months and were only available in a subset of participants. Still, these results illustrate that, apart from persistently lowering TSH, levothyroxine treatment did result in a significant increase in fT4. The lack of effect of levothyroxine on haemoglobin levels in this study can therefore not be explained by a lack of increase in thyroid hormone function. Fourth, additional markers of erythropoiesis or other potential causes of anaemia such as ferritin, iron, folate, vitamin B12 or kidney function, were not available, restricting further exploration of underlying pathophysiological mechanisms. Fifth, although sensitivity analyses were performed excluding those using anti-anaemic medication such as iron, vitamin B12 or erythropoietin, no information was available on blood transfusions or venesection.

In conclusion, treatment with levothyroxine does not improve haemoglobin levels in individuals with a diagnosis of subclinical hypothyroidism aged 65 years and older. Whether anaemia in patients with more marked hypothyroidism is responsive to treatment with levothyroxine needs further experimental studies.

REFERENCES

1. Beghe C, Wilson A, Ershler WB. Prevalence and outcomes of anemia in geriatrics: a systematic review of the literature. *Am J Med.* 2004;116 Suppl 7A:3S-10S.
2. Cooper DS, Biondi B. Subclinical thyroid disease. *Lancet (London, England).* 2012;379(9821):1142-1154.
3. Gaskell H, Derry S, Andrew Moore R, McQuay HJ. Prevalence of anaemia in older persons: systematic review. *BMC Geriatr.* 2008;8:1.
4. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA.* 2010;304(12):1365-1374.
5. Das KC, Mukherjee M, Sarkar TK, Dash RJ, Rastogi GK. Erythropoiesis and erythropoietin in hypo- and hyperthyroidism. *J Clin Endocrinol Metab.* 1975;40(2):211-220.
6. Fandrey J, Pagel H, Frede S, Wolff M, Jelkmann W. Thyroid hormones enhance hypoxia-induced erythropoietin production in vitro. *Experimental hematology.* 1994;22(3):272-277.
7. Fein HG, Rivlin RS. Anemia in thyroid diseases. *The Medical clinics of North America.* 1975;59(5):1133-1145.
8. Golde DW, Bersch N, Chopra IJ, Cline MJ. Thyroid hormones stimulate erythropoiesis in vitro. *Br J Haematol.* 1977;37(2):173-177.
9. Bremner AP, Feddema P, Joske DJ, et al. Significant association between thyroid hormones and erythrocyte indices in euthyroid subjects. *Clin Endocrinol (Oxf).* 2012;76(2):304-311.
10. den Elzen WP, de Craen AJ, Mooijaart SP, Gussekloo J. Low thyroid function and anemia in old age: the Leiden 85-plus study. *J Am Geriatr Soc.* 2015;63(2):407-409.
11. Evans ES, Rosenberg LL, Simpson ME. Erythropoietic response to calorigenic hormones. *Endocrinology.* 1961;68:517-532.
12. Gianoukakis AG, Leigh MJ, Richards P, et al. Characterization of the anaemia associated with Graves' disease. *Clin Endocrinol (Oxf).* 2009;70(5):781-787.
13. Horsley V. The Brown Lectures on Pathology. *Br Med J.* 1885;1(1261):419-423.
14. Horton L, Coburn RJ, England JM, Himsworth RL. The haematology of hypothyroidism. *The Quarterly journal of medicine.* 1976;45(177):101-123.
15. Kawa MP, Grymula K, Paczkowska E, et al. Clinical relevance of thyroid dysfunction in human haematopoiesis: biochemical and molecular studies. *European journal of endocrinology / European Federation of Endocrine Societies.* 2010;162(2):295-305.
16. Kendrick TS, Payne CJ, Epis MR, et al. Erythroid defects in TRalpha-/- mice. *Blood.* 2008;111(6):3245-3248.
17. Tudhope GR, Wilson GM. Anaemia in hypothyroidism. Incidence, pathogenesis, and response to treatment. *The Quarterly journal of medicine.* 1960;29:513-537.
18. Vitale G, Fatti LM, Prolo S, et al. Screening for hypothyroidism in older hospitalized patients with anemia: a new insight into an old disease. *J Am Geriatr Soc.* 2010;58(9):1825-1827.
19. Wopereis DM, Du Puy RS, van Heemst D, et al. The Relation Between Thyroid Function and Anemia: A Pooled Analysis of Individual Participant Data. *J Clin Endocrinol Metab.* 2018;103(10):3658-3667.
20. Stott DJ, Gussekloo J, Kearney PM, et al. Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST). *BMC Endocr Disord.* 2017;17(1):6.
21. Stott DJ, Rodondi N, Kearney PM, et al. Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism. *The New England journal of medicine.* 2017;376(26):2534-2544.

22. Du Puy RS, Postmus I, Stott DJ, et al. Study protocol: a randomised controlled trial on the clinical effects of levothyroxine treatment for subclinical hypothyroidism in people aged 80 years and over. *BMC Endocr Disord.* 2018;18(1):67.
23. Nutritional anaemias. Report of a WHO scientific group. *World Health Organization technical report series.* 1968;405:5-37.
24. Floriani C, Feller M, Aubert CE, et al. Thyroid Dysfunction and Anemia: A Prospective Cohort Study and a Systematic Review. *Thyroid.* 2018;28(5):575-582.
25. M'Rabet-Bensalah K, Aubert CE, Coslovsky M, et al. Thyroid dysfunction and anaemia in a large population-based study. *Clin Endocrinol (Oxf).* 2016;84(4):627-631.
26. Mooijaart SP, Du Puy RS, Stott DJ, et al. Association Between Levothyroxine Treatment and Thyroid-Related Symptoms Among Adults Aged 80 Years and Older With Subclinical Hypothyroidism. *JAMA.* 2019;322(20):1-11.
27. Christ-Crain M, Meier C, Huber P, Zulewski H, Staub JJ, Muller B. Effect of restoration of euthyroidism on peripheral blood cells and erythropoietin in women with subclinical hypothyroidism. *Hormones (Athens, Greece).* 2003;2(4):237-242.
28. Maggio M, De Vita F, Fisichella A, et al. The Role of the Multiple Hormonal Dysregulation in the Onset of "Anemia of Aging": Focus on Testosterone, IGF-1, and Thyroid Hormones. *Int J Endocrinol.* 2015;2015:292574.
29. Perrin MC, Blanchet JP, Mouchiroud G. Modulation of human and mouse erythropoiesis by thyroid hormone and retinoic acid: evidence for specific effects at different steps of the erythroid pathway. *Hematol Cell Ther.* 1997;39(1):19-26.
30. Sullivan PS, McDonald TP. Thyroxine suppresses thrombocytopoiesis and stimulates erythropoiesis in mice. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine (New York, NY).* 1992;201(3):271-277.



Chapter 9

General discussion

Subclinical hypothyroidism in community-dwelling older people (arbitrarily 65 years and older) is a long-debated subject.[1] In the absence of robust scientific evidence, and with decades of arguably trial-and-error experimentation with levothyroxine treatment, opinions about whether it actually constitutes a diseased state, whether monitoring is required and whether it requires levothyroxine treatment in community-dwelling older people are spread out wide and bolstered, amongst patients, physicians and researchers alike.

This thesis set out to 1) establish whether subclinical hypothyroidism in older people is a neutral, beneficial or detrimental condition by establishing if subclinical hypothyroidism in older people is associated with a) clinically relevant outcomes and b) biologically relevant outcomes, and, 2) investigate if levothyroxine treatment for subclinical hypothyroidism in older people provides long-term benefits in clinically or biologically relevant outcomes. In the next paragraphs, the main findings and their implications from a physician-, patient- and societal-perspective are discussed as well as suggested scientific areas for future exploration.

MAIN FINDINGS

Part 1: subclinical hypothyroidism in older people is a neutral condition not associated with clinically or biologically relevant outcomes

Subclinical hypothyroidism is, by definition, a biochemically defined disease state. However, one may argue that to warrant further physician action, be it monitoring, diagnostic or therapeutic, the definition should not only rely on the traditional reference ranges alone, but should extend to also include clinically or biologically relevant outcomes.

In **chapter 2** we establish that in 4 prospective, international cohorts of community-dwelling oldest old (80 years and older, N=2116), participants with thyroid dysfunction, including subclinical hypothyroidism, had no significant differences in clinically relevant outcomes at baseline and after 5 years of follow-up; i.e. disability in activities of daily living, cognitive functioning, depressive symptoms, grip strength or mortality risk compared to euthyroid age-matched counterparts.

Traditionally, antithyroperoxidase antibody (TPOAb) positivity has always been included in diagnostic and therapeutic algorithms as an additional predictor of poor outcome and a criterion to start levothyroxine treatment. The primary driving force being a potential 2% annual increase in progression from subclinical to overt hypothyroidism.[2] In **chapter 3** we investigate TPOAb positivity in participants in the Leiden 85-plus Study, a population-based cohort study of residents of Leiden aged 85 years and older, and fail to reach the same conclusion. We found that TPOAb levels were indeed associated with higher TSH levels, as

commonly found in subclinical hypothyroidism. Nevertheless, they did not predict a 5-year change in thyroid function (in particular progression to overt hypothyroidism), physical function, disability in ADL, cognitive function or depressive symptoms. Although elevated TPOAb levels were associated with 10-year cumulative survival benefit, suggesting a possible beneficial effect of subclinical hypothyroidism, the effect was independent of circulating fT4 levels making it less likely that any favourable effects are mediated through thyroid function.

One of the most commonly found biological co-occurrences is between subclinical hypothyroidism and erythropoiesis – the production of red blood cells. In **chapter 4,** the results from 16 international longitudinal cohorts in the Thyroid Studies Collaboration (N=42,162) were pooled in an individual participant data meta-analysis to establish whether subclinical thyroid dysfunction was associated with anaemia. Although cross-sectionally the presence of subclinical hypothyroidism was accompanied by a higher odds ratio of having anaemia in subgroup analyses of older adults, the analysis failed to demonstrate an increased hazard of developing anaemia over a median follow-up of 5.7 years.

In conclusion, in older community-dwelling people, subclinical hypothyroidism is not associated with clinically relevant outcomes or biologically relevant outcomes.

Part 2: levothyroxine treatment for subclinical hypothyroidism in older people does not provide long-term benefits in clinically or biologically relevant outcomes

Until recently the combined randomised controlled trial evidence for levothyroxine for subclinical hypothyroidism was limited with only 12 trials (with 350 people in total included) in the most recent Cochrane Review[3]; few in numbers, small cohorts, with varying age and reference ranges used, outcomes measured and medications used. The author's recommended that "until better data are available, clinical judgment and patients' preferences remain the best manner to decide" and that large, standardised, international trials were necessary to increase the evidence base.

The TRUST study (**Chapter 5**, N=737, mean age 74 years) shows that after 12 months of levothyroxine treatment the TSH levels in serum had changed significantly in the verum arm, but that the mean change in hypothyroid symptoms, tiredness, quality of life, hand-grip strength, blood pressure, body-mass index or adverse events was no different when compared to the placebo group. Treatment with levothyroxine for persistent subclinical hypothyroidism in participants aged 65 years and older provided no apparent benefits.

There is ample data to suggest that thyroid function, and any possible consequences, are mediated by age. To this end we designed the IEMO 80-plus randomised controlled trial as

an ancillary study to the TRUST study, sharing trial infrastructure and protocols to allow for a joint analysis of all participants aged 80 years and older with persistent subclinical hypothyroidism (**Chapter 6**).

The IEMO 80-plus study (N=251, mean age 85 years) shows, in line with the results from the TRUST trial, that after 12 months of levothyroxine treatment indeed the serum TSH levels had changed significantly, but that no change in hypothyroid symptoms, tiredness, quality of life or any of the other endpoints could be identified compared to the placebo group (**Chapter 7**).

In an additional pre-planned combined analysis using data from both the TRUST and IEMO 80-plus thyroid trials we discovered that treatment with levothyroxine resulted in no increases in haemoglobin levels (**Chapter 8**). The results were not different when stratifying by sex, age, TSH level, or presence of anaemia at baseline, suggesting that in subclinical hypothyroidism, thyroid function and haemoglobin levels is not causally related and should not be used to influence physician practice and policies.

In conclusion, treatment with levothyroxine for subclinical hypothyroidism in older people does not provide benefits in hypothyroid symptoms, tiredness, haemoglobin levels and a range of secondary outcomes.

Is subclinical hypothyroidism in community-dwelling older people a disease?

Considering the aforementioned two conclusions, one may wonder if it's necessary to re-evaluate how we look at subclinical hypothyroidism in community-dwelling older people. Although the International Classification of Diseases version 11 has no code listed for subclinical hypothyroidism,[4] and the International Classification of Primary Care version 1 lists subclinical hypothyroidism as an 'aberrant laboratory result',[5] experts have been disagreeing for years whether subclinical hypothyroidism in older people actually constitutes a disease.

The Oxford medical dictionary defines disease as 'Any illness or abnormal condition of the body with a specific cause (which may or may not be known), excluding physical trauma, that has recognizable signs and symptoms.' The Merriam-webster dictionary extends the definition with 'an impairment of the normal state of the living animal or plant body or one of its parts that interrupts or modifies the performance of the vital functions, is typically manifested by distinguishing signs and symptoms, and is a response to environmental factors (as malnutrition, industrial hazards, or climate), to specific infective agents (as worms, bacteria, or viruses), to inherent defects of the organism (as genetic anomalies), or to combinations of these factors'.

As mentioned earlier subclinical hypothyroidism, particularly in older people, is mostly asymptomatic or is accompanied by varying and unspecific signs and symptoms. The diagnosis is often established after a chance laboratory finding. Keeping the Oxford criteria of disease in mind we may find it difficult to defend that subclinical hypothyroidism fulfils all the criteria required to meet the definition. A more lenient viewpoint could be to include the domains 'defects of the organism' and 'interrupts or modifies the performance of vital functions' according to the Merriam-webster definition. However, as presented in the previous chapters, our research demonstrates that in community-dwelling populations of older people, to the best of our current knowledge, subclinical hypothyroidism is not related to defects of the organism (i.e. continued production of fT4) and does not interrupt or modify vital functions compared to euthyroidism (i.e. no influence on clinically or biologically relevant outcomes).

All evidence considered, it may be concluded that in community-dwelling populations of older people, the state we currently describe as subclinical hypothyroidism is not a disease but a strictly biochemical diagnosis that is not associated with detrimental nor beneficial health outcomes, but with neutral health effects at best.

Understanding age-adjusted reference ranges

In parallel with the growing evidence-base that shows that subclinical hypothyroidism in older people is a neutral condition, some experts have solicited that the thyroid function reference ranges should be adjusted. Although the TSH distribution depends among other things on the population studied and assays used, commonly a range between 0.4 and 4.5 mIU/L is defined as normal.[6] Some,[7-9] but not all,[10-15] studies have demonstrated an age-related increase in median TSH levels among presumed healthy older individuals prompting several experts to advocate a change in guidelines towards age- and sex-specific TSH reference ranges that widen in interval and increase in median with increasing age and differ per sex. If this is implemented, a proportion of older people currently diagnosed with subclinical hypothyroidism would be reclassified as euthyroid. There are, however, several caveats to this approach that are sometimes overlooked. Perhaps there is a better alternative.

Currently the TSH reference range serves dual purposes: 1) describing a distribution in the general population, as well as 2) the basis for treatment decisions for physicians. If we consider that these purposes may not necessarily require the same numbers, an alternative, elegant solution presents itself. Given the results in this thesis, it could be perfectly valid to continue describing a population distribution using the 2.5^{th} and 97.5^{th} percentiles like we do today, and nothing would change in the classification. However, subclinically hypothyroid older people do not seem to benefit from levothyroxine treatment, particularly with TSH<10 mIU/L. A secondary set of treatment thresholds, not based on the statistical distribution but based on scientifically substantiated expected benefit or prevention, would then serve as

levothyroxine treatment cut-offs. Much like how the current American and European thyroid guidelines already advocate more leniency towards treatment initiation if TSH levels are >10 mIU/L based on preventing a higher cardiovascular disease risk identified in a landmark IPD-meta analysis.[6] Using this approach doctors can finally start to "treat patients, not numbers".[16] This approach would be very challenging, requiring full international collaboration, discussion and consensus, on top of sufficiently powered and set-up randomised controlled trials and IPD-meta analyses, but is perhaps a more appropriate solution to the reference range dilemma.

IMPLICATIONS OF FINDINGS

Patients, physicians and researchers have been debating the impact and management of subclinical hypothyroidism in older people for decades, and it would be presumptuous to state that the evidence in this thesis ends that debate. Still, the choices and considerations (i.e. inclusion of older people, internationally sampled populations, long follow-up times, harmonised reference ranges and outcomes) that went into the research in this thesis make it of particular value for physicians serving older, community-dwelling patients with subclinical hypothyroidism. For this group as a whole, the main findings of the two research objectives demonstrate that in older community-dwelling people, the presence or resolution of subclinical hypothyroidism is not associated with clinically or biologically relevant outcomes and that routine treatment with levothyroxine does not provide long term benefits.

These findings may be of considerable importance in how care is provided to community-dwelling older people. The provision of high-value care in modern healthcare systems is inextricably linked with coordinating and balancing simultaneous commitments to physicians, patients, and society. In the next few paragraphs several implications for each are discussed.

Subclinical hypothyroidism and implications for physicians

Many physicians struggle with managing subclinical hypothyroidism in older people. A qualitative study by Gibbons *et al.* in 2009 was one of the first to lay bare that although physicians tried a patient-centred approach, they reported having little knowledge of the disease, possibly due to uncertainty regarding prognosis and variations in advice regarding treatment.[17] In 2012 Allport *et al.* concluded that British GPs are potentially uncertain how to interpret symptoms and thyroid function tests in older adults.[18] In 2015 a survey from den Elzen *et al.* confirmed this by establishing large GP treatment variations by country (the Netherlands, Germany, England, Ireland, Switzerland and New Zealand) and by patient characteristics, with some treating patients outside of guideline recommendations.[19] At the time of writing this thesis (2021), the latest international, professional, practice guidelines date from 2012

and 2013 (American Thyroid Association 2012[20], European Thyroid Association 2013[21], Dutch College of General Practitioners 2013[22]). It is a worrying thought that more than 10 years after the first paper of Gibbons demonstrated the difficulties in daily medical practice, the scientific community has not been able to provide physicians with more robust guiding principles and decision support other than 8-year-old low- to moderate-quality recommendations.

Potential for changes in clinical practice

The results presented in this thesis reinforce the recommendations that physicians may serve their older community-dwelling patients with subclinical hypothyroidism best through:

1) resolving the uncertainty of potential future health consequences; i.e. not attributing subclinical hypothyroidism to clinically or biologically relevant outcomes and preventing medicalization,
2) unburdening patients and practices from lifelong, periodic and invasive diagnostic schemes; i.e. exercising restraint when it comes to follow-up thyroid function or TPOAb testing, and by
3) protecting patients from unnecessary lifelong pharmaceutical management and from potential overtreatment; i.e. not initiating routine thyroid hormone substitution therapy.

In 2019 an international group of researchers published an independent and unendorsed clinical practice guideline, based on a systematic review of 21 trials (including some reported in this thesis) performed in (older) adults with subclinical hypothyroidism.[23,24] They concluded that 'almost all adults with subclinical hypothyroidism would not benefit from treatment with thyroid hormones', in line with the findings of this manuscript. To this date, however, the aforementioned official international guidelines remain unaltered. Updating these should be a first step in giving physicians a major advantage in the expected increase in subclinical hypothyroidism-related daily practice cases. If not for the entire subclinically hypothyroid group as a whole yet, then at least, based on the outcomes of this thesis, for older people.

Subclinical hypothyroidism and implications for patients

Medical professionals have had debates about subclinical hypothyroidism for decades, but perhaps it's patients that struggle with their affliction and opposing opinions the most. Unfortunately, patient perspectives are not always incorporated in research or guidelines. To avoid paternalistic medical care, researchers will have to reallocate some of their focus into understanding how our older patients actually feel, think and act[25], to tailor medical care to the needs and wants of older patients with subclinical hypothyroidism. Regardless of whether subclinical hypothyroidism is or is not detrimental to health, and regardless of whether levothyroxine treatment is or isn't indicated. Although guideline experts have ad-

vocated incorporating patient values, preferences and perspectives into practice guidelines for years,[26,27] no formal qualitative studies for subclinical hypothyroidism are available to date.

To bridge the gap, researchers will need to properly investigate how older patients feel about subclinical hypothyroidism, including, but not limited to: Would they experience lifelong pill taking and attending periodic testing as burdensome and a dependency? How would their lives be affected if a prominent supplier of levothyroxine suddenly stops shipments? Would they consider a trial of deprescribing levothyroxine in time?

This is potentially an area where most is to be gained. In an online survey of 11.166 American hypothyroid patients, participants scored their satisfaction with their treating physician (both GP and specialist) a disappointing 5 out of 10 (10 = 'completely satisfied'), the doctor´s knowledge 5 out of 10 (10 = 'very knowledgeable') and 71% had switched doctors at least once due to thyroid-related dissatisfaction.[28] It is not unlikely that older patients with subclinical hypothyroidism experience a similar divide between them and their treating physician. This may drive patients away, for instance to explore alternatives to levothyroxine treatment. Any such trend is inherently perpendicular to the values and standards of good medical care.

Additionally, the gap between patients and physicians may be expected to widen over time. For example, several countries including America, Canada, the UK, Australia and the Netherlands, have made their primary care electronic healthcare records (EHR) available for online access by patients and caregivers. Evaluation studies of early open EHR experiments have demonstrated that patients expressed a particular interest in accessing laboratory test results including thyroid function tests.[29] Some studies found this phenomenon to have no effect on anxiety levels,[30,31] while other studies report more negative emotions, more uncertainty and increased anxiety.[32] In half of all test results outside of the reference ranges, participants carried out online searches for additional information, and only half called the physician's office for advice. It is unclear where and how patients with abnormal thyroid function tests will find the sources on which to base any future decision and how this will affect them and their caregivers but it is likely that tensions between patients and physicians will resultingly increase. A striking example can be found in the use of T4+T3-combination therapy. In recent years the number of patients with thyroid dysfunction using a combined therapy of T4 and T3 has increased,[33] although this is not advocated in the guidelines. The narrative has changed from a constructive dialogue, to simply asking for this treatment to sometimes outright demanding this treatment modality, often caused by 'indiscriminate statements on the internet'.[34] By reopening the dialogue, and listening to patient perspectives and preferences, physicians may be able to address these issues appropriately, with or without medication.

Subclinical hypothyroidism and implications for society

Subclinical thyroid (dys-)function has a large impact on most healthcare systems worldwide and its global impact is expected to rise, both diagnostically and therapeutically. If screening, follow-up and treatment are not warranted perhaps the current medical practices leave room for improvement.

Diagnostics

TSH testing (interestingly, not fT4 testing) is listed in the World Health Organization list of Essential In Vitro Diagnostic laboratory tests and is one of the most ordered lab tests worldwide. For example in the United States of America, TSH testing was responsible for $484 million of Medicare spending in 2017, higher than any other laboratory test.[35] In the Netherlands over 2.7 million TSH tests were performed in over 2.2 million individual Dutch people in primary care in 2017,[36] and this number is expected to have risen since. In only 725.000 people however, additional fT4 tests were performed when TSH was found to be abnormal. The large number of TSH tests, relative to the small number of subsequent fT4 tests and users of levothyroxine medication, suggests that the majority of the TSH tests was performed from a screening perspective.[36] This is in contrast with the traditional Bayesian approach where, based on a priori suspicion of disease, laboratory testing is used to make a diagnosis more or less likely. The screening approach, rather than a diagnostic approach, suggests a fundamental uncertainty in physicians and potential patients that impacts healthcare systems and that is hard to correct without additional and robust scientific evidence. There is no telling how much more tests will be performed in the future if the over-the-counter TSH home testing kits for self-diagnosis, that have recently become available on the market, gain in popularity.

As the global demographic continues to age, and TSH levels have been shown to rise with age, screening practices for thyroid disease, both physician-requested and commercially, increases the odds of incidental medical findings, likely causing an overestimation of (mild) subclinical hypothyroidism in primary care. This may cause increased levels of anxiety and uncertainty for patients, increased costs, time consumption and work for physicians, and increased burden on healthcare systems through additional diagnostics, follow-up or treatment.

Optimizing thyroid function screening practices

Historically TSH measurements have always come first in screening practices due to better laboratory assays. A log-linear relationship between TSH and fT4 makes that little variations in TSH-levels correspond with major changes in fT4 and are therefore more sensitive. Only when TSH levels are identified outside of the normal reference ranges do we measure the fT4 levels to properly diagnose the amount of thyroid dysfunction; a process called reflex-testing. However, if for the majority of the general older population, subclinical thyroid disease is of

little clinical importance (only TSH results abnormal), yet identifying overt thyroid disease is (both fT4 and TSH results abnormal), it may be interesting to explore new diagnostic screening strategies. Including, for instance, a diagnostic algorithm that measures fT4 first and, in the event of it being high or low, reflex test TSH levels. This new and bold approach may reduce the number of identified people with subclinical hypothyroidism, reduce the amount of unnecessarily alarmed patients, reduce unnecessary lab analyses, reduce unnecessary treatment for subclinical thyroid disorders and improve limited resource spending. A Leiden-based initiative dubbed the "RESTORE"-study has been proposed to investigate a similar new diagnostic strategy.

Treatment

Levothyroxine is one of the most prescribed drugs worldwide. In the United States of America 122 million prescriptions for levothyroxine were dispensed in 2017, more than any other prescription drug.[37] In the Netherlands the amount of users of levothyroxine drugs in the Netherlands has risen from 447.880 in 2014 to 501.170 in 2018, with more than half of all patients (227.629) aged 65 years and older.[38] The total costs amount to 9.3 million Euro's per year and these have risen by over 800.000 euros over the past few years.

It is currently unknown how many of these prescriptions are given for subclinical hypothyroidism but it stands to reason that this constitutes a significant proportion. In older populations, up to 60% of patients with subclinical hypothyroidism revert back to a euthyroid state within 5 years without any intervention at all, and our results in chapters 5, 7 and 8 demonstrate that even in the case of persistent subclinical hypothyroidism treatment with levothyroxine is not associated with beneficial health effects. This indicates that a sizeable proportion of people with subclinical hypothyroidism is treated with levothyroxine unnecessarily, and could even suffer negative health consequences from it.

Possibility for deprescribing

Since international thyroid guidelines do not recommend regular re-evaluation of treatment indications, levothyroxine supplementation is usually continued for life,[39,40] often even without treatment evaluation.[41] The risk of levothyroxine overtreatment increases with age and years of treatment duration and it is estimated that up to 41% of older levothyroxine users show signs of over-suppletion.[42] Without compelling evidence of effect and with a high risk for overtreatment, guidelines may need to be evaluated to reduce the amount of routine thyroid function tests ordered and to decrease the ever-growing amount of levothyroxine prescriptions for older people with subclinical hypothyroidism.

Ultimately policy and guideline committees may also want to reassess whether the older, community-dwelling patients with subclinical hypothyroidism that are already on levothy-

roxine treatment for many years can safely be withdrawn from treatment while maintaining their health and well-being. Currently one self-controlled observational study ("RELEASE"-study), undertaken by the LUMC (The Netherlands), is investigating to what extent 360 levothyroxine users aged 60 years and older can successfully and safely be deprescribed. The results from this analysis should prove instrumental in optimizing thyroid care for older people.

LIMITATIONS TO GENERALIZABILITY

Although the results from this thesis may help shape the discourse of subclinical hypothyroidism in older people in medical care in general, it is hardly the be-all and end-all of subclinical thyroid dysfunction. It should be noted that the conclusions reached apply primarily to community-dwelling populations of older people; i.e. the vast majority of people with subclinical hypothyroidism. Because thyroid function is a continuum it cannot be excluded that in subgroups of patient populations on either end of the subclinical hypothyroidism spectrum, such as those under the direct care of an endocrinologist, other clinical and biological outcomes or levothyroxine treatment effects may be identified. This limitation extends to more subgroups including, but not limited to, iodine-, selenium- or iron-deficient populations, patients with significant pre-existent cardiovascular disease,[6,43] patients using thyroid influencing therapies (amiodarone, lithium or radiotherapy) or patients with TSH levels >10 mIU/L.[44] Although resource-intensive and time-consuming, additional studies aiming particularly at more specialised groups are needed before the results of this thesis can be fully generalised to these groups that are by and large more the exception than the rule.

FUTURE RESEARCH

Understanding subclinical hypothyroidism in older people is complex and we have only just uncovered the tip of the proverbial iceberg, with the majority of our understanding still left hidden below the water level. The findings from research objectives 1 and 2 demonstrate that for older community-dwelling people with subclinical hypothyroidism currently employed diagnostic and levothyroxine treatment strategies are not routinely indicated but this, however, does not mean that researchers can now comfortably sit on their hands. In order to optimise the management of subclinical hypothyroidism researchers may want to focus on generalizing findings to younger age groups, on particular populations (such as more marked thyroid dysfunction with TSH >10 mIU/L) that may or may not warrant particular management strategies and on the potential for deprescribing for current levothyroxine users. Without presuming to be exhaustive, a list of especially interesting future research ideas

may include understanding age-adjusted reference ranges, optimizing thyroid function screening practices and how to move forward with the results from this thesis.

Understanding the biology of the ageing thyroid

A future research area of particular interest is how and why subclinical hypothyroidism in old age happens in the first place. Many experts agree that in general median TSH levels increase with age (regardless of clinical significance). How these thyroid ageing processes actually work 'in vivo', however, has proven hard to investigate. Regardless of ethical limitations (for example requiring invasive biopsies), older people are generally exposed to a myriad of non-physiological mediators, both presently and over the course of their life. These include, but are not limited to, thyroid-influencing medication (such as tyrosine kinase inhibitors in cancer treatment), thyroid-influencing therapies (such as head/neck radiation), reduced intake of nutrients (such as iodine and selenium, both crucial for thyroid hormone function), thyroid-influencing comorbid conditions (such as autoimmune diseases like rheumatoid arthritis or vascular disease like atherosclerosis) and chronic low-grade inflammation (dubbed 'inflammaging').[45,46] Currently the complex series of events and interactions that ultimately lead to subclinical hypothyroidism in old age are yet to be uncovered but may prove vital to understanding the condition and its consequences in full. The European research project THYRAGE is currently attempting to elucidate the effects of thyroid hormone on a wide range of age-related diseases.[47] One particularly appealing theory is that the increased TSH levels and reduced thyroid hormone signalling in old age are actually the hallmarks of an evolutionary advantageous natural selection in which after the reproductive phase, the focus shifts from sexual maturity and fitness to functional optimization and somatic repair, leading to increased health and longevity in old age.[46]

CONCLUDING REMARKS

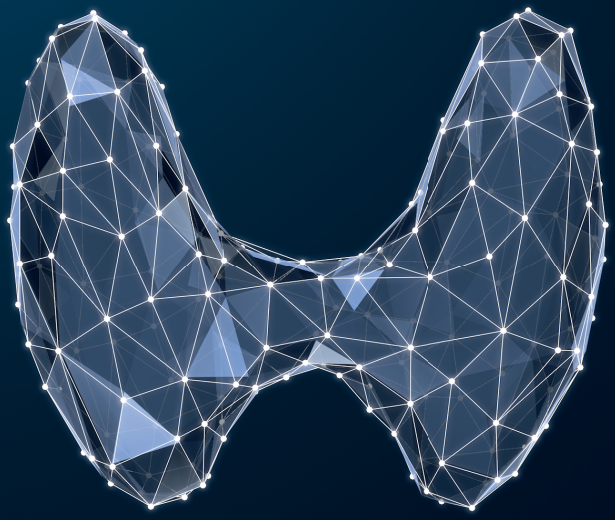
The findings from this thesis demonstrate that in older community-dwelling people, subclinical hypothyroidism is a neutral condition not associated with clinically or biologically relevant outcomes. Levothyroxine treatment does not provide long-term benefits. Physicians provide the best possible medical care by employing a more conservative management style reducing thyroid function testing and levothyroxine prescriptions. Reflecting on the statements in the general introduction, physicians: 'should not handle this elevated TSH level finding as a disease, may want to reassure their patients that this is not abnormal and does not explain any potential symptoms, may want to conservatively monitor thyroid function over time and should not start levothyroxine treatment instead.'

REFERENCES:

1. Peeters RP, Brito JP. Subclinical hypothyroidism: to treat or not to treat? *European journal of endocrinology.* 2020;183(6):D15-D24.
2. Vanderpump MP, Tunbridge WM, French JM, et al. The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. *Clinical endocrinology.* 1995;43(1):55-68.
3. Villar HC, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. *The Cochrane database of systematic reviews.* 2007;2007(3):CD003419.
4. World Health Organization. International Classification of Diseases, 11th Revision (ICD-11). In: World Health Organization; 2018.
5. WONCA Working Party: International Classification. International Classification of Primary Care. In: World Health Organization; 1987.
6. Rodondi N, den Elzen WP, Bauer DC, et al. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *Jama.* 2010;304(12):1365-1374.
7. Bjoro T, Holmen J, Kruger O, et al. Prevalence of thyroid disease, thyroid dysfunction and thyroid peroxidase antibodies in a large, unselected population. The Health Study of Nord-Trondelag (HUNT). *European journal of endocrinology.* 2000;143(5):639-647.
8. Boucai L, Surks MI. Reference limits of serum TSH and free T4 are significantly influenced by race and age in an urban outpatient medical practice. *Clinical endocrinology.* 2009;70(5):788-793.
9. Bremner AP, Feddema P, Leedman PJ, et al. Age-related changes in thyroid function: a longitudinal study of a community-based cohort. *The Journal of clinical endocrinology and metabolism.* 2012;97(5):1554-1562.
10. Guan H, Shan Z, Teng X, et al. Influence of iodine on the reference interval of TSH and the optimal interval of TSH: results of a follow-up study in areas with different iodine intakes. *Clinical endocrinology.* 2008;69(1):136-141.
11. Mariotti S, Barbesino G, Caturegli P, et al. Complex alteration of thyroid function in healthy centenarians. *The Journal of clinical endocrinology and metabolism.* 1993;77(5):1130-1134.
12. Mariotti S, Franceschi C, Cossarizza A, Pinchera A. The aging thyroid. *Endocrine reviews.* 1995;16(6):686-715.
13. Olsen T, Laurberg P, Weeke J. Low serum triiodothyronine and high serum reverse triiodothyronine in old age: an effect of disease not age. *The Journal of clinical endocrinology and metabolism.* 1978;47(5):1111-1115.
14. Tunbridge WM, Evered DC, Hall R, et al. The spectrum of thyroid disease in a community: the Whickham survey. *Clinical endocrinology.* 1977;7(6):481-493.
15. van den Beld AW, Visser TJ, Feelders RA, Grobbee DE, Lamberts SW. Thyroid hormone concentrations, disease, physical function, and mortality in elderly men. *The Journal of clinical endocrinology and metabolism.* 2005;90(12):6403-6409.
16. Jonklaas J, Razvi S. Reference intervals in the diagnosis of thyroid dysfunction: treating patients not numbers. *The lancet Diabetes & endocrinology.* 2019;7(6):473-483.
17. Gibbons V, Lillis S, Conaglen J, Lawrenson R. The reality of subclinical hypothyroidism in general practice. *Journal of primary health care.* 2009;1(3):215-221.
18. Allport J, McCahon D, Hobbs FD, Roberts LM. Why are GPs treating subclinical hypothyroidism? Case note review and GP survey. *Primary health care research & development.* 2013;14(2):175-184.

19. den Elzen WP, Lefebre-van de Fliert AA, Virgini V, et al. International variation in GP treatment strategies for subclinical hypothyroidism in older adults: a case-based survey. *The British journal of general practice : the journal of the Royal College of General Practitioners.* 2015;65(631):e121-132.
20. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists.* 2012;18(6):988-1028.
21. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. *European thyroid journal.* 2013;2(4):215-228.
22. NHG-werkgroep: Van Lieshout J F-SB, Bolsius EJM, Boer AM, Burgers JS, Bouma M, Sijbom M,. NHG-STANDAARD Schildklieraandoeningen. In: Nederlands Huisartsen Genootschap; 2013.
23. Bekkering GE, Agoritsas T, Lytvyn L, et al. Thyroid hormones treatment for subclinical hypothyroidism: a clinical practice guideline. *BMJ (Clinical research ed).* 2019;365:l2006.
24. Feller M, Snel M, Moutzouri E, et al. Association of Thyroid Hormone Therapy With Quality of Life and Thyroid-Related Symptoms in Patients With Subclinical Hypothyroidism: A Systematic Review and Meta-analysis. *Jama.* 2018;320(13):1349-1359.
25. Nexo MA, Watt T, Cleal B, et al. Exploring the experiences of people with hypo- and hyperthyroidism. *Qualitative health research.* 2015;25(7):945-953.
26. Montori VM, Brito JP, Murad MH. The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines. *Jama.* 2013;310(23):2503-2504.
27. Zhang Y, Coello PA, Brozek J, et al. Using patient values and preferences to inform the importance of health outcomes in practice guideline development following the GRADE approach. *Health and quality of life outcomes.* 2017;15(1):52.
28. Peterson SJ, Cappola AR, Castro MR, et al. An Online Survey of Hypothyroid Patients Demonstrates Prominent Dissatisfaction. *Thyroid : official journal of the American Thyroid Association.* 2018;28(6):707-721.
29. Redelmeier DA, Kraus NC. Patterns in Patient Access and Utilization of Online Medical Records: Analysis of MyChart. *Journal of medical Internet research.* 2018;20(2):e43.
30. Davis Giardina T, Menon S, Parrish DE, Sittig DF, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. *Journal of the American Medical Informatics Association : JAMIA.* 2014;21(4):737-741.
31. Mak G, Smith Fowler H, Leaver C, Hagens S, Zelmer J. The Effects of Web-Based Patient Access to Laboratory Results in British Columbia: A Patient Survey on Comprehension and Anxiety. *Journal of medical Internet research.* 2015;17(8):e191.
32. Giardina TD, Baldwin J, Nystrom DT, Sittig DF, Singh H. Patient perceptions of receiving test results via online portals: a mixed-methods study. *Journal of the American Medical Informatics Association : JAMIA.* 2018;25(4):440-446.
33. Jonklaas J, Tefera E, Shara N. Short-Term Time Trends in Prescribing Therapy for Hypothyroidism: Results of a Survey of American Thyroid Association Members. *Frontiers in endocrinology.* 2019;10:31.
34. Wiersinga WM. THERAPY OF ENDOCRINE DISEASE: T4 + T3 combination therapy: is there a true effect? *European journal of endocrinology.* 2017;177(6):R287-R296.
35. US Department of Health and Human Services. *Medicare payments for clinical diagnostic laboratory tests in 2017: year 4 of baseline data.* . 2018.
36. Zorginstituut Nederland. *Systematische analyse Endocriene Ziekten, voedings- en stofwisselingsstoornissen.* Diemen2018.

37. IQVIA Institute for Human Data Science. *Medicine use and spending in the US: a review of 2017 and outlook to 2022*. 2018.
38. Zorginstituut Nederland. *GIPdatabank: Aantal gebruikers 2012-2016 voor ATC-subgroep H03: Schildkliermiddelen*. Diemen2018.
39. Hall R, Scanlon MF. Hypothyroidism: clinical features and complications. *Clinics in endocrinology and metabolism*. 1979;8(1):29-38.
40. Taylor PN, Iqbal A, Minassian C, et al. Falling threshold for treatment of borderline elevated thyrotropin levels-balancing benefits and risks: evidence from a large community-based study. *JAMA internal medicine*. 2014;174(1):32-39.
41. Mandel SJ, Brent GA, Larsen PR. Levothyroxine therapy in patients with thyroid disease. *Annals of internal medicine*. 1993;119(6):492-502.
42. Somwaru LL, Arnold AM, Joshi N, Fried LP, Cappola AR. High frequency of and factors associated with thyroid hormone over-replacement and under-replacement in men and women aged 65 and over. *The Journal of clinical endocrinology and metabolism*. 2009;94(4):1342-1345.
43. Gencer B, Collet TH, Virgini V, et al. Subclinical thyroid dysfunction and the risk of heart failure events: an individual participant data analysis from 6 prospective cohorts. *Circulation*. 2012;126(9):1040-1049.
44. Korevaar TIM, Chaker L, Peeters RP. Improving the clinical impact of randomised trials in thyroidology. *The lancet Diabetes & endocrinology*. 2018;6(7):523-525.
45. Boelaert K. Thyroid dysfunction in the elderly. *Nature reviews Endocrinology*. 2013;9(4):194-204.
46. Rozing MP, Houwing-Duistermaat JJ, Slagboom PE, et al. Familial longevity is associated with decreased thyroid function. *The Journal of clinical endocrinology and metabolism*. 2010;95(11):4979-4984.
47. van Heemst D, Remaud S, Williams G, Dentice M, Gereben B, Timmerman P. Resetting the THYRoid axis for prevention of AGE-related diseases and co-morbidities. In: European Commission; 2016.



Chapter 10

English Summary

The thyroid gland plays a crucial role in regulating nearly all bodily processes. A particular state of thyroid function, dubbed subclinical hypothyroidism (with high thyroid-stimulating hormone [TSH] and normal free thyroid hormone [fT4] in blood tests), is often asymptomatic or accompanied with unspecific symptoms. Subclinical hypothyroidism becomes more prevalent with advancing age and it is unclear how this condition should be interpreted in community-dwelling older people (arbitrarily 65 years and older). The consequences and treatment considerations regarding subclinical hypothyroidism in community-dwelling older people are long-debated subjects. In the absence of robust and conclusive scientific evidence, and with decades of arguably trial-and-error experimentation with levothyroxine treatment (artificial thyroid hormone), opinions about whether monitoring is required and whether subclinical hypothyroidism requires levothyroxine treatment or not are spread out wide and bolstered.

Some observational studies have associated subclinical hypothyroidism in older people with negative consequences such as diminished physical and cognitive function, increased risk of depression, increased risk of progression to more serious thyroid disease and even increased mortality. In part because of these potential negative consequences, some patients are treated with levothyroxine whilst others are not. At the same time, other observational studies fail to arrive at the same conclusions or find evidence to the contrary. Only a few trials have investigated whether thyroid hormone supplementation provides any benefits. Most studies however suffer from key issues, such as low numbers of participants, various age ranges, different laboratory reference ranges and short follow-up times that limit generalizability to the vast majority of community-dwelling older people with subclinical hypothyroidism.

To improve the care of subclinical hypothyroidism in older people, the aims of this thesis were two-fold:

1. to establish whether subclinical hypothyroidism in older people is associated with clinically relevant outcomes and biologically relevant outcomes that may render it a neutral, beneficial or detrimental condition.
2. to investigate if levothyroxine treatment for subclinical hypothyroidism in older people provides long-term benefits.

Part 1: Consequences of subclinical hypothyroidism in older people

In the first part of this thesis, we first established whether subclinical hypothyroidism is actually associated with negative outcomes in community-dwelling older people, that warrant any physician action.

In **Chapter 2** we set out to measure clinical health outcomes in 4 prospective cohorts including 2,116 participants aged 80 years and older in the Netherlands, New Zealand, United

Kingdom and Japan. We found that participants with subclinical hypothyroidism scored the same in activities of daily living, cognition, mood and physical function as their euthyroid counterparts, and did not suffer from increased mortality. Moreover, no differences in functioning were found after 5 years of follow-up. The influence of subclinical hypothyroidism on clinically relevant outcomes therefore seems limited in community-dwelling older people.

It has been demonstrated that the presence of antithyroperoxidase antibodies (TPOAb) is associated with autoimmune thyroid diseases such as Hashimoto's thyroiditis or Graves' disease. Antibodies are usually produced by the human body to fight of harmful infections, but TPOAb target the thyroid gland instead. It is thought that TPOAb positivity is a predictor of progression to overt hypothyroidism and that these cases of subclinical hypothyroidism may actually suffer more negative consequences. Therefore, in **Chapter 3** we investigated the associations between TPOAb levels and clinical outcomes in community-dwelling older people in the Leiden 85-plus Study; a cohort including 488 residents of Leiden aged 85 at the start of the study. Although TPOAb positivity was indeed associated with higher thyroid-stimulating hormone levels at baseline and after 3 years of follow-up, it was not associated with an increased risk of change in thyroid function (i.e. progression to overt hypothyroidism), nor with activities of daily living, mood and physical function. Contradictory to expectations, positive TPOAb predicted a slight survival benefit after 10-years, but it is unlikely that this is caused by thyroid hormones. Accordingly, the consequences of positive TPOAb for clinically relevant outcomes in subclinical hypothyroidism in older people is minimal.

In observational studies subclinical hypothyroidism is often found in co-occurrence with biological outcomes such as anaemia. Therefore, in **Chapter 4** we analysed the data from 16 international longitudinal cohort studies in the Thyroid Studies Collaboration consisting of over 23,000 participants aged 65 years and older, who had measurements of thyroid function and anaemia. Although the chance of having anaemia at the start of the study was slightly higher in the presence of subclinical hypothyroidism, there was no increased risk of developing anaemia over almost six years of follow-up time.

The results from the first part of the thesis demonstrate that subclinical hypothyroidism in community-dwelling older people is not associated with clinically relevant or biologically relevant outcomes.

Part 2: Treatment outcomes for subclinical hypothyroidism in older people

In the second part of the thesis, the focus was shifted from observational to experimental studies, in order to discover if levothyroxine treatment for subclinical hypothyroidism in older people provided long-term benefits in clinically or biologically relevant outcomes.

In the TRUST study, an international randomised controlled trial involving 737 community-dwelling adults with subclinical hypothyroidism aged 65 years and older, we sought to understand whether levothyroxine treatment for at least one year would provide clinical benefits. In this study, written down in **Chapter 5,** we found that there was no change in hypothyroid symptoms, tiredness, adverse events or a battery of secondary clinical outcomes after at least one year of treatment.

Because thyroid function changes with age, we hypothesised that the effects may be profoundly different in the oldest old. Therefore in **Chapters 6** and **7** we investigated any potential clinical benefits in even older adults; aged 80 years and older. To allow for a joint analysis of all participants aged 80 years and older, the IEMO 80-plus study was designed as a complementary trial to the TRUST study, described in **Chapter 6**. In **Chapter 7** the results demonstrate that treatment with levothyroxine in participants aged 80 years and older was not associated with changes in symptoms, tiredness, adverse events or an extensive set of secondary clinical outcomes after at least one year of treatment. This is in line with the results from **Chapter 5**.

Finally, in **Chapter 8**, we studied whether levothyroxine treatment was associated with changes in biological outcomes – i.e. the occurrence or resolution of anaemia, or any change in haemoglobin at all. By again combining the data from the TRUST and IEMO 80-plus thyroid trials we discovered that there was no change in anaemia status after a minimum of one-year levothyroxine treatment for subclinical hypothyroidism in older people.

The results from the second part of the thesis show that levothyroxine treatment does not lead to benefits in clinically or biologically relevant outcome measures. These results do not support routine treatment for all older adults with subclinical hypothyroidism.

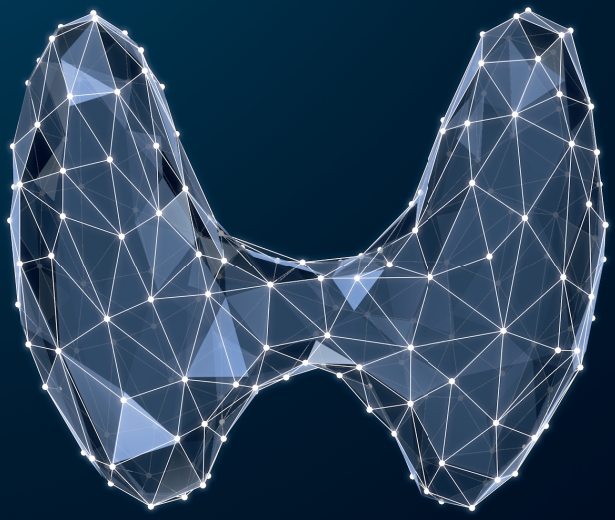
Discussion

All evidence considered, it may be concluded that in community-dwelling populations of older people, the state we currently describe as subclinical hypothyroidism is not a disease but a strictly biochemical diagnosis that is not associated with detrimental nor beneficial health outcomes. Treatment with levothyroxine does not provide benefits. It is important to keep in mind that the conclusions of this thesis do apply to the vast majority of community-dwelling older people, but may not be generalisable to certain specific subgroups of older people with subclinical hypothyroidism, such as those under the care of a medical specialist.

Physicians could play a crucial role by preventing medicalisation, unburdening patients from lifelong invasive diagnostic schemes, exercising restraint in pharmaceutical management and reducing potential overtreatment. Laboratory screening protocols may need to be updated

(and TSH-reflex testing potentially even reversed after additional research) and deprescribing studies should investigate if current levothyroxine users can taper and stop treatment without harm. Additionally, future studies should investigate whether the findings of this thesis may be generalised to more uncommon populations and to younger age groups. Finally, more research will be needed to explore patient values, preferences and perspectives about subclinical hypothyroidism, age-adjusted laboratory reference ranges and the biology of the ageing thyroid. We encourage guideline committees to update the guideline recommendations to realign with the findings, conclusions and recommendations of this thesis.

For now, the findings from this thesis suggest that the best medical care is provided by employing a more conservative management style, reducing thyroid function testing and levothyroxine prescriptions, but that there is still much left to be discovered.



Chapter 11

Nederlandse samenvatting

De schildklier speelt een cruciale rol in het reguleren van nagenoeg alle lichamelijke processen. Een specifieke afwijking in schildklierfunctie, genaamd subklinische hypothyreoïdie (met een hoog 'Thyroïd Stimulerend Hormoon' [TSH] en normaal 'vrij schildklier hormoon' [fT4] in bloed onderzoek), komt vaak voor zonder, of alleen met aspecifieke, symptomen. Subklinische hypothyreoïdie komt vaker voor bij een toenemende leeftijd en het is onduidelijk hoe deze afwijking moet worden geïnterpreteerd bij thuiswonende ouderen (arbitrair 65 jaar en ouder). De consequenties van, en behandelopties voor, subklinische hypothyreoïdie bij thuiswonende ouderen zijn onderwerpen waar al decennia over gedebatteerd wordt. In de afwezigheid van robuust en eenduidig wetenschappelijk bewijs, en gesteund door tientallen jaren van kleinschalige experimenten met levothyroxine behandeling (kunstmatig schildklier hormoon), zijn de meningen verdeeld over de zin en onzin van diagnostiek en of levothyroxine behandeling nodig is voor subklinische hypothyreoïdie.

Enkele observationele studies hebben associaties gevonden tussen subklinische hypothyreoïdie bij ouderen en afgenomen fysiek en cognitief functioneren, toegenomen risico op depressie, toegenomen risico op ontwikkeling naar ernstigere schildklierziekten en zelfs toegenomen risico op sterfte. Mede op basis van deze mogelijke negatieve uitkomsten worden vele oudere patiënten behandeld met levothyroxine. Tegelijkertijd worden in andere observationele onderzoeken deze associaties niet gevonden, of juist precies tegenovergestelde effecten gevonden. Slechts een handvol experimentele studies hebben onderzocht of het toedienen van schildklierhormoon daadwerkelijk leidt tot enig voordeel. De meeste studies hebben last van enkele methodologische gebreken die verhinderen dat de resultaten kunnen worden toegepast op de meerderheid van de thuiswonende ouderen met subklinische hypothyreoïdie, zoals kleine aantallen deelnemers, verschillende leeftijdscategorieën, verschillende laboratorium referentiewaarden en korte follow-up tijd.

Om de zorg voor subklinische hypothyreoïdie bij ouderen te verbeteren, heeft dit proefschrift twee doelstellingen:

1. vaststellen of subklinische hypothyreoïdie bij ouderen is geassocieerd met klinisch relevante uitkomstmaten en biologisch relevante uitkomstmaten.
2. onderzoeken of behandeling met levothyroxine voor subklinische hypothyreoïdie bij ouderen lange-termijn voordelen oplevert.

Deel 1: De consequenties van subklinische hypothyreoïdie bij ouderen

In het eerste deel van dit proefschrift onderzoeken we of subklinische hypothyreoïdie inderdaad is geassocieerd met negatieve uitkomsten die medisch handelen bij thuiswonende ouderen rechtvaardigen.

In **hoofdstuk 2** onderzoeken we klinisch relevante gezondheidsuitkomsten in 4 prospectieve cohort studies, met in totaal 2,116 deelnemers van 80 jaar en ouder, in Nederland, Nieuw Zeeland, het Verenigd Koninkrijk en Japan. Deelnemers met subklinische hypothyreoïdie behalen dezelfde resultaten in activiteiten in het dagelijks leven, cognitie, stemming en fysieke functie als deelnemers met normale schildklierfunctie. Er is geen sprake van toegenomen sterfte. Tevens zijn na 5 jaar follow-up geen verschillen in functioneren waarneembaar tussen de deelnemers met subklinische hypothyreoïdie en de deelnemers met normale schildklierfunctie. De invloed van subklinische hypothyreoïdie op relevante gezondheidsuitkomsten bij thuiswonende ouderen lijkt daarom beperkt.

Bij auto-immuun schildklierziekten zoals de ziekte van Hashimoto of de ziekte van Graves worden niet zelden antithyroperoxidase antilichamen (TPOAb) gevonden. Antilichamen worden normaal gesproken door het lichaam gevormd om een schadelijk virus of bacterie onschadelijk te maken, maar bij TPOAb richten deze zich (per ongeluk) op de schildklier. Er zijn aanwijzingen dat de aanwezigheid van TPOAb voorspellend is op de ontwikkeling van subklinische hypothyreoïdie naar klinische hypothyreoïdie en dat bij deze patiënten er mogelijk meer negatieve consequenties zijn dan bij patiënten die de antistoffen niet hebben. Om die reden zijn in **hoofdstuk 3** de associaties tussen TPOAb spiegels en klinische uitkomsten onderzocht bij ouderen in de Leiden 85-plus Studie; een cohortstudie met 488 inwoners uit Leiden die bij aanvang van de studie 85 jaar waren. Hoewel verhoogd TPOAb inderdaad is geassocieerd met hogere waarden van TSH bij aanvang en na 3 jaar follow-up, blijkt er geen verhoogd risico in verandering van schildklierfunctie (bijvoorbeeld progressie van subklinische naar klinische hypothyreoïdie) waarneembaar. Er is geen relatie tussen TPOAB en (veranderingen in) activiteiten in het dagelijkse leven, stemming of fysieke functie. Tegen de verwachtingen in lijkt er juist sprake van een minimaal betere 10-jaars overleving, al is het onwaarschijnlijk dat dit door schildklierhormonen wordt veroorzaakt. De consequenties van verhoogde TPOAb waarden op klinische relevante uitkomsten bij subklinische hypothyreoïdie bij ouderen lijken minimaal.

In observationele studies wordt subklinische hypothyreoïdie vaak samen gevonden met andere biologische afwijkingen zoals bloedarmoede (anemie). In **hoofdstuk 4** analyseren wij de samengevoegde data uit 16 internationale longitudinale cohort studies in de Thyroid Studies Collaboration, bestaande uit 23,000 deelnemers van 65 jaar of ouder met metingen van schildklierfunctie en van bloedarmoede. De kans op het hebben van bloedarmoede is bij aanvang licht verhoogd als ook subklinische hypothyreodie aanwezig is, maar er blijkt geen verhoogd risico op het ontwikkelen van bloedarmoede na bijna zes jaar tijd.

De resultaten uit het eerste deel van dit proefschrift wijzen erop dat subklinische hypothyreoïdie bij thuiswonende ouderen niet is geassocieerd met klinisch relevante of biologisch relevante uitkomsten.

Deel 2: De uitkomsten van behandeling voor subklinische hypothyreoïdie bij thuiswonende ouderen.

In het tweede deel van dit proefschrift is de aandacht verlegd van observationele naar experimentele studies, om te achterhalen of behandeling met levothyroxine voor subklinische hypothyreoïdie bij thuiswonende ouderen lange termijn voordelen op klinische of biologisch relevante uitkomstmaten oplevert.

In de TRUST studie, een internationaal opgezette gerandomiseerde gecontroleerde studie met 737 thuiswonende ouderen met subklinische hypothyreoïdie van 65 jaar en ouder, onderzoeken wij of minimaal één jaar behandeling met levothyroxine tabletten klinische voordelen oplevert. In deze studie, beschreven in **hoofdstuk 5**, vinden wij echter geen verbetering in symptomen van hypothyreoïdie, vermoeidheid, bijwerkingen of een reeks aan secundaire klinische uitkomsten na minimaal een jaar behandeling.

Omdat de schildklier functie verandert met de leeftijd kunnen de effecten in de oudste ouderen anders zijn dan in jongere groepen. Om die reden onderzoeken wij in **hoofdstuk 6** en **7** mogelijke klinische voordelen bij ouderen die nog wat ouder zijn; 80 jaar en ouder. In hoofdstuk 6 beschrijven wij de studie opzet van de IEMO 80-plus schildklier studie, die opgezet is als complementaire studie met de TRUST studie, zodat de data gezamenlijk geanalyseerd kunnen worden. In **hoofdstuk 7** laten de resultaten zien dat behandeling met levothyroxine in deelnemers van 80 jaar en ouder niet geassocieerd is met veranderingen in symptomen van hypothyreoïdie, vermoeidheid, bijwerkingen of een reeks aan secundaire klinische uitkomsten na minimaal een jaar behandeling. Dit is in lijn met de bevindingen in **hoofdstuk 5**.

Ten slotte onderzoeken wij in **hoofdstuk 8** of behandeling met levothyroxine geassocieerd is met veranderingen in biologische uitkomsten zoals het optreden of juist verdwijnen van bloedarmoede. Door opnieuw de data van de TRUST en IEMO studies te combineren is duidelijk geworden dat de behandeling met levothyroxine voor subklinische hypothyreoïdie bij ouderen geen invloed heeft op het ontstaan of verdwijnen van bloedarmoede status na minimaal een jaar behandeling.

De resultaten uit het tweede deel van dit proefschrift wijzen erop dat behandeling met levothyroxine niet leidt tot voordelen op klinische of biologisch relevante uitkomstmaten.

Deze resultaten pleitten tegen het routinematig behandelen van alle ouderen personen met subklinische hypothyreoïdie.

Discussie

De bevindingen in dit proefschrift ondersteunen de stelling dat subklinische hypothyreoïdie geen ziekte is, maar een pure laboratorium diagnose die geen negatieve of positieve gezondheidseffecten heeft. Behandeling met levothyroxine levert geen voordelen op. Een kanttekening is dat hoewel deze conclusies van toepassing zijn op de overgrote meerderheid van de thuiswonende ouderen, zij wellicht niet gelden voor bepaalde gespecialiseerde subgroepen van ouderen met subklinische hypothyroïdie, zoals patiënten die onder behandeling zijn van een medisch specialist.

Artsen kunnen een belangrijke rol vervullen bij het voorkomen van medicalisering, door patiënten met subklinische hypothyreoïdie te ontlasten van regelmatig bloedprikken, door terughoudend te zijn met medicijnen voorschrijven en door potentiële overbehandeling terug te brengen. Screenend laboratorium onderzoek zou wellicht herzien kunnen worden (door bijvoorbeeld TSH-reflex testen om te draaien) en studies moeten worden opgezet om de effecten van het minderen en eventueel stoppen met levothyroxine te onderzoeken. Tevens is verder onderzoek nodig om uit te wijzen of de bevindingen uit dit proefschrift ook mogen worden gegeneraliseerd naar meer 'ongewone' populaties en naar jongere leeftijdsgroepen. Ten slotte zou in toekomstig onderzoek meer aandacht moeten worden besteed aan de waarden, voorkeuren en perspectieven van patiënten met subklinische hypothyreoïdie, naar leeftijdsafhankelijke referentiewaarden en naar de biologie van de verouderende schildklier. Wij moedigen richtlijn commissies aan om de bevindingen, conclusies en aanbevelingen in dit proefschrift in richtlijnen op te nemen.

Tot slot, de bevindingen van dit proefschrift laten zien dat goede medische zorg bestaat uit een meer terughoudend behandelbeleid met minder schildklier diagnostiek en levothyroxine voorschriften, maar dat er nog een heleboel te ontdekken valt.

BIBLIOGRAPHY

Publications in this thesis

1. **Du Puy RS**, Poortvliet RKE, Mooijaart SP, den Elzen WPJ, Jagger C, Pearce SHS, Arai Y, Hirose N, Teh R, Menzies O, Rolleston A, Kerse N, Gussekloo J. *Outcomes of Thyroid Dysfunction in People Aged Eighty Years and Older: An Individual Patient Data Meta-Analysis of Four Prospective Studies (Towards Understanding Longitudinal International Older People Studies Consortium)*. Thyroid : official journal of the American Thyroid Association. 2020.
2. **Du Puy RS**, Poortvliet RKE, Snel M, den Elzen WPJ, Ballieux B, Dekkers OM, Mooijaart SP, Gussekloo J. *Associations of Elevated Antithyroperoxidase Antibodies with Thyroid Function, Survival, Functioning, and Depressive Symptoms in the Oldest Old: The Leiden 85-plus Study.* Thyroid : official journal of the American Thyroid Association. 2019;29(9):1201-1208.
3. Wopereis DM, **Du Puy RS**, van Heemst D, Walsh JP, Bremner A, Bakker SJL, Bauer DC, Cappola AR, Ceresini G, Degryse J, Dullaart RPF, Feller M, Ferrucci L, Floriani C, Franco OH, Iacoviello M, Iervasi G, Imaizumi M, Jukema JW, Khaw KT, Luben RN, Molinaro S, Nauck M, Patel KV, Peeters RP, Psaty BM, Razvi S, Schindhelm RK, van Schoor NM, Stott DJ, Vaes B, Vanderpump MPJ, Volzke H, Westendorp RGJ, Rodondi N, Cobbaert CM, Gussekloo J, den Elzen WPJ, Thyroid Studies C. *The Relation Between Thyroid Function and Anemia: A Pooled Analysis of Individual Participant Data.* The Journal of clinical endocrinology and metabolism. 2018;103(10):3658-3667.
4. Stott DJ, Rodondi N, Kearney PM, Ford I, Westendorp RGJ, Mooijaart SP, Sattar N, Aubert CE, Aujesky D, Bauer DC, Baumgartner C, Blum MR, Browne JP, Byrne S, Collet TH, Dekkers OM, den Elzen WPJ, **Du Puy RS**, Ellis G, Feller M, Floriani C, Hendry K, Hurley C, Jukema JW, Kean S, Kelly M, Krebs D, Langhorne P, McCarthy G, McCarthy V, McConnachie A, McDade M, Messow M, O'Flynn A, O'Riordan D, Poortvliet RKE, Quinn TJ, Russell A, Sinnott C, Smit JWA, Van Dorland HA, Walsh KA, Walsh EK, Watt T, Wilson R, Gussekloo J, Group TS. *Thyroid Hormone Therapy for Older Adults with Subclinical Hypothyroidism.* The New England journal of medicine. 2017;376(26):2534-2544.
5. **Du Puy RS**, Postmus I, Stott DJ, Blum MR, Poortvliet RKE, Den Elzen WPJ, Peeters RP, van Munster BC, Wolffenbuttel BHR, Westendorp RGJ, Kearney PM, Ford I, Kean S, Messow CM, Watt T, Jukema JW, Dekkers OM, Smit JWA, Rodondi N, Gussekloo J, Mooijaart SP. *Study protocol: a randomised controlled trial on the clinical effects of levothyroxine treatment for subclinical hypothyroidism in people aged 80 years and over.* BMC endocrine disorders. 2018;18(1):67.
6. Mooijaart SP, **Du Puy RS**, Stott DJ, Kearney PM, Rodondi N, Westendorp RGJ, den Elzen WPJ, Postmus I, Poortvliet RKE, van Heemst D, van Munster BC, Peeters RP, Ford I, Kean S, Messow CM, Blum MR, Collet TH, Watt T, Dekkers OM, Jukema JW, Smit JWA, Langhorne P, Gussekloo J. *Association Between Levothyroxine Treatment and Thyroid-Related*

Symptoms Among Adults Aged 80 Years and Older With Subclinical Hypothyroidism. Jama. 2019;322(20):1-11.

7. Mooijaart SP, **Du Puy RS**. *Treatment of Older Adults With Subclinical Hypothyroidism-Reply*. Jama. 2020;323(11):1097-1098.

Other publications

8. Chavannes NH, **Du Puy RS**, Bai C. *Suggestions for health information technology trials for respiratory disorders in low- and middle-income country settings: what can we learn from trials in high-income country settings?* NPJ primary care respiratory medicine. 2015;25:15045.
9. Stott DJ, Gussekloo J, Kearney PM, Rodondi N, Westendorp RG, Mooijaart S, Kean S, Quinn TJ, Sattar N, Hendry K, **Du Puy RS**, Den Elzen WP, Poortvliet RK, Smit JW, Jukema JW, Dekkers OM, Blum M, Collet TH, McCarthy V, Hurley C, Byrne S, Browne J, Watt T, Bauer D, Ford I. *Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST)*. BMC endocrine disorders. 2017;17(1):6.
10. O Riordan D, Aubert CE, Walsh KA, Van Dorland A, Rodondi N, **Du Puy RS**, Poortvliet RKE, Gussekloo J, Sinnott C, Byrne S, Galvin R, Jukema JW, Mooijaart SP, Baumgartner C, McCarthy V, Walsh EK, Collet TH, Dekkers OM, Blum MR, Kearney PM. *Prevalence of potentially inappropriate prescribing in a subpopulation of older European clinical trial participants: a cross-sectional study.* BMJ open. 2018;8(3):e019003.

Bibliography available online at ORCID

https://orcid.org/0000-0001-8909-7686

ACKNOWLEDGEMENTS

'*There is no "I" in research*'. De spreekwoordelijke promotie-reis die heeft geleid tot dit proefschrift had nooit een finishvlag gezien zonder de hulp, inspiratie en aanwijzingen van velen die hebben geholpen om de juiste koers te blijven varen of juist even van de gebaande paden af te wijken.

Jacobijn, Simon en Rosalinde, jullie hebben mij als promotor en copromotoren altijd ondersteund waar nodig en uitgedaagd waar mogelijk, waardoor ik tijdens het promotietraject nooit gestopt ben met exploreren en leren. Met jullie hulp heb ik niet alleen grote stappen gezet als academicus, maar ook als persoon. Ook de leden van mijn begeleidingscommissie zijn hierin instrumentaal gebleken: Wendy, zonder jouw enthousiasme en aanmoediging was dit traject waarschijnlijk nooit gestart, en Petra, zonder jouw hulp en flexibiliteit was het misschien nooit afgerond. Bedankt.

Een groot deel van het ervaren werkplezier is te danken aan alle collega's bij de afdeling PHEG. Alle AIOTHO's, POIOTHO's, promovendi en in het bijzonder natuurlijk de promovendi-'kabouters' op de P0. Dank voor de vele momenten om te sparren, de spontane borrels, mislukte escape-rooms en de hilarische avonturen op het werk en daarbuiten. Ik kijk uit naar jullie verdediging!

De TRUST en IEMO studies hebben een centrale plaats in dit proefschrift en waren nooit zo succesvol geworden zonder de hulp van een groot aantal mensen. In de eerste plaats veel dank aan de andere Principal Investigators, professoren Rudi Westendorp, David Stott, Nicolas Rodondi en Patricia Kearney, dat ik op jullie spreekwoordelijke trein mocht opstappen. Zonder de medewerking van alle secretaresses, onderzoeksmedewerkers, data-managers, artsen, verpleegkundigen en student-assistenten hadden deze studies slechts een concept gebleven. Angelique, Boes, Brenda, Corrie, Daisy, Esther, Henk, Hilda, Hilde, Inge, Iris, Marjan, Martin, Merel, Necmiye, Niki, Olga, Rob en Sanne, dank dat jullie zo'n geweldig team waren en dat wij zo vaak op jullie ervaring, wijsheid en humor mochten leunen. Het was een genot om met jullie samen te werken. Hoewel het teveel mensen zijn om op persoonlijke titel te bedanken, wil ik graag ook de bijdragen van alle (huisartsen-)laboratoria, de klinisch chemici, de huisartsen die de deelnemers benaderden en de leden van het ouderenberaad benoemen. Ten slotte uiteraard ontzettend veel dank aan alle ouderen en hun families die deel hebben genomen aan de studies, jullie bijdrage heeft de toekomst van schildklierziekten blijvend verbeterd.

Graag bedank ik alle secretaresses, opleiders (van externe stages), huisartsbegeleiders, tutoren, gedragswetenschappers, tutorgroepjes en mede-AIOS van de huisartsopleiding

die mij de ruimte boden en flexibiliteit toonden om dit traject naast de huisartsopleiding af te ronden. Ivalda en Bart, in jullie praktijk heb ik weer mogen ervaren hoeveel de kliniek en wetenschap elkaar nodig hebben. Irvine, bij jou in de praktijk heb ik echt geleerd wat voor arts ik nou ben en vooral wil blijven. Tevens wil ik de SBOH bedanken voor het mogelijk maken van een gecombineerd promotietraject tijdens de opleiding tot huisarts.

De paranimfen. Jonathan, ik denk regelmatig aan de avonden met net te veel Belgische bieren en slap gekeuvel, onze eerlijke maar niet geheel geslaagde poging tot een bandje en onze academische avonturen aan de andere kant van de wereldbol. Bedankt voor alles. Charlotte, ik ken geen andere collega die mij zo creatief weet uit te dagen; zowel academisch als op een bierviltje, die sterrenwaardige gerechten koken kan, die vaker in de media is geweest dan dat de meesten van ons een publicatie hebben, en die mij zo vaak en terecht, letterlijk en figuurlijk, wist te corrigeren als jij. Ik voel me vereerd om jullie aan mijn zijde te hebben staan. Ontzettend bedankt.

Lieve familie en vrienden, dank voor al jullie momenten van enthousiasme, raad, discussie en (sportieve) gezelligheid. Bob, wat kunnen weekendjes mountainbiken en sterrenkijken toch helpen om de blik te verruimen. Bedankt. Lieve pap, mam, broer en zus, dank voor jullie oneindige geduld, eeuwige steun en ontelbare barbecues en cache-wandelingen.

Lieve Laura, jij was er bij vanaf het begin en hebt met mij de afgelopen 14 jaar alle pieken gevierd en alle dalen gekoesterd. Dank voor alle goede zorgen, die mijn leven en dit proefschrift kleur hebben gegeven. Hoewel dít avontuur nu af is, is onze reis nog lang niet klaar, en ik kan niet wachten om te zien waar onze rit ons nog brengen gaat.

CURRICULUM VITAE

Robert Du Puy was born on February 9 1989, in Voorburg, the Netherlands. He grew up in Zoetermeer, graduated from the Alfrink College (bilingual atheneum, International Baccalaureate Higher Level English) in 2007 and in that same year started medical school at the Leiden University Medical Center. During his years in medical school, he proudly served as a member of the organizing committee of the Leiden International (Bio)Medical Student Conference 2009-2011. He completed medical school in 2014 and started working as junior researcher at the department of Public Health and Primary Care for a short while, before working as a PhD-student at the same department during the period 2015-2019. The research project focused on subclinical hypothyroidism in community-dwelling older adults, and to that end he joined the Dutch division of the international TRUST and IEMO 80+ consortia, under the supervision of Prof. J. Gussekloo, Dr. Simon P. Mooijaart and Dr. Rosalinde K.E. Poortvliet. Nearing the end of the PhD period he was appointed as junior teacher for the Master Vitality & Ageing at the Leiden University Medical Center in 2019 and completed his Basiskwalificatie Onderwijs training at the Leiden University in 2021. He was involved with successfully applying for two successive studies that build upon the work in this thesis, the RELEASE-study (2018) and the RESTORE-study (2020). Robert is currently in training to become a General Practitioner at the Stichting Beroepsopleidingen Huisartsen, Utrecht, the Netherlands, and is scheduled to graduate as a licensed GP in 2022.