

## SHORT REPORT

# Screening for thyroid dysfunction with free T4 instead of thyroid stimulating hormone (TSH) improves efficiency in older adults in primary care

MADELEEN BOSMA<sup>1,2,3</sup>, ROBERT S. DU PUY<sup>4</sup>, BART E.P.B. BALLIEUX<sup>1</sup>

<sup>1</sup>Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands

<sup>2</sup>Department of Clinical Chemistry, St. Antonius Hospital, Nieuwegein, The Netherlands

<sup>3</sup>Eurofins Salux, Eurofins Clinical Diagnostics NL, Leiden, The Netherlands

<sup>4</sup>Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands

Address correspondence to: Madeleen Bosma, Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands. Tel: +31-(0)71-5298268. Email: [m.bosma.l@lumc.nl](mailto:m.bosma.l@lumc.nl)

---

## Abstract

Subclinical hypothyroidism (SCHT) is defined as a consistently elevated thyroid stimulating hormone (TSH) with a free T4 (fT4) within the reference range. This diagnosis may lead to additional monitoring, levothyroxine therapy and increased patient concerns, despite lack of evidence of treatment benefit in older adults. In order to avoid this diagnosis, we evaluated the efficiency of fT4-based screening for thyroid dysfunction, in older adults in primary care and compared it with TSH-based screening. Individuals aged >65years in primary care were selected for this retrospective study when both TSH and fT4 were individually requested irrespective of the TSH value. Exclusion criteria were C-reactive protein > 10 mg/l or a history of thyroid hormone monitoring in the previous year. Screening based on fT4 instead of TSH decreased reflex testing from 23.8% to 11.2%. The positive predictive value (PPV) for clinical hypothyroidism increased from 17.3% to 52.2%. The negative predictive value was 96.1% with TSH-based screening versus 97.8% with fT4-based screening. Elevation of the TSH cutoff value from 4.2 to 6.5 mU/l resulted in a reflex test percentage of 12.5% and a PPV of 31.0%. Our results suggest that screening for thyroid dysfunction in older individuals in primary care can be improved by screening based on fT4 instead of TSH or by adjusting the TSH cutoff value. Adjustment of the screening strategy may be of interest to health policy makers because of potential cost reduction. From a patient perspective, medical concerns and unnecessary biochemical follow-up might be reduced by circumventing the diagnosis SCHT.

**Keywords:** thyroid disease, screening, older adults, thyroid stimulating hormone (TSH), subclinical hypothyroidism, older people

## Key Points

- Screening for thyroid dysfunction with free T4 instead of thyroid stimulating hormone (TSH) improves efficiency in older adults in primary care.
  - Adjustment of the screening strategy will reduce costs, unnecessary follow-up and patient concerns.
  - An alternative strategy is to elevate the TSH cutoff value; this also led to improved positive predictive value (PPV) for clinical hypothyroidism
-

## Introduction

Screening for thyroid dysfunction is frequently performed in older individuals presenting with fatigue and general malaise to exclude thyroid disease. The diagnosis of thyroid disease is largely biochemical. Hence, thyroid stimulating hormone (TSH) and free thyroxine (fT4) cutoff values, which vary significantly between laboratories, are important determinants of diagnosis, monitoring and treatment. Current screening guidelines for thyroid dysfunction in primary care suggest a TSH assay followed by analysis of fT4, if TSH results are outside the reference interval (reflex testing). TSH is a sensitive marker. Linear changes in fT4 lead to exponential changes in TSH [1]. Hence, TSH has traditionally been the first choice screening test.

A consistently elevated TSH with an fT4 within the reference range results in the diagnosis subclinical hypothyroidism (SCHT). Although associations between SCHT and some areas of medicine such as cardio-renal metabolic disease are still under investigation [2–4], in older persons (65 years and over) SCHT is generally believed not to be associated with quality of life, symptom burden, physical or cognitive functioning, other diseases or mortality [5–7]. Additionally, multiple studies have shown that older adults with SCHT do not benefit from levothyroxine supplementation [6, 8, 9]. Still, in daily practice, the diagnosis SCHT leads to additional monitoring, unnecessary treatment and increased patient concerns.

Therefore, fT4-based screening to identify clinical thyroid disease would potentially be a more efficient screening strategy for older adults with possible symptoms of thyroid disease. This strategy could reduce costs and patient concerns. Several reports dating from the '80's and '90's have already suggested such a testing algorithm [10–12]. Alternatively, TSH decision limits could be adjusted to reduce unnecessary reflex testing and thus reduce unwarranted monitoring [13].

In this retrospective study we evaluated the efficiency of fT4-based screening for thyroid dysfunction in older adults in primary care compared with TSH-based screening. Furthermore, the alternative approach of elevating TSH decision limits was investigated in parallel.

## Methods

Patient data (fT4, TSH, C-reactive protein [CRP], date of birth) from the laboratory information system of the St. Antonius Hospital (Nieuwegein, the Netherlands) were analysed retrospectively. Individuals aged >65 years in primary care were selected when both TSH and fT4 were individually requested irrespective of the TSH value. Reflex testing for thyroid function (i.e. fT4 testing is performed only when the TSH is outside the reference range) was not implemented in the period that the data collection took place: 2018–19. In order to exclude patients with active inflammatory disease and patients treated and/or monitored for thyroid dysfunction, individuals with inflammation (CRP > 10 mg/l) or a history of thyroid hormone monitoring in the previous year

were excluded. Solely the first TSH and fT4 measurements for each patient were included in the analysis. The institutional medical ethics committee (MEC-U) approved the study.

TSH and fT4 were analysed in lithium–heparin plasma on a Roche Cobas platform (Roche Diagnostics, Mannheim, Germany) using the Elecsys FT4 III and the Elecsys TSH assay, respectively.

Three models were evaluated in the dataset: the current screening protocol (fT4 reflex test when TSH > 4.2 mU/l) (model 1), elevation of the TSH decision limit (model 2) and fT4-based screening (TSH reflex test when fT4 < 12 pmol/l) (model 3).

SCHT was defined as a TSH value above the reference interval (>4.2 mU/l) combined with a fT4 value inside the reference interval (12–22 pmol/l). Overt hypothyroidism was defined as either an increased TSH value combined with a decreased fT4 value or a TSH value > 10 mU/l.

Reference intervals were based on the instructions for use of the Roche assays. The positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity and cut-off value (receiving operating curve [ROC] curve) analyses were performed in Microsoft Excel version 2011 (Microsoft, Redmond, WA, USA) and SPSS version 26 (IBM, Armonk, NY, USA).

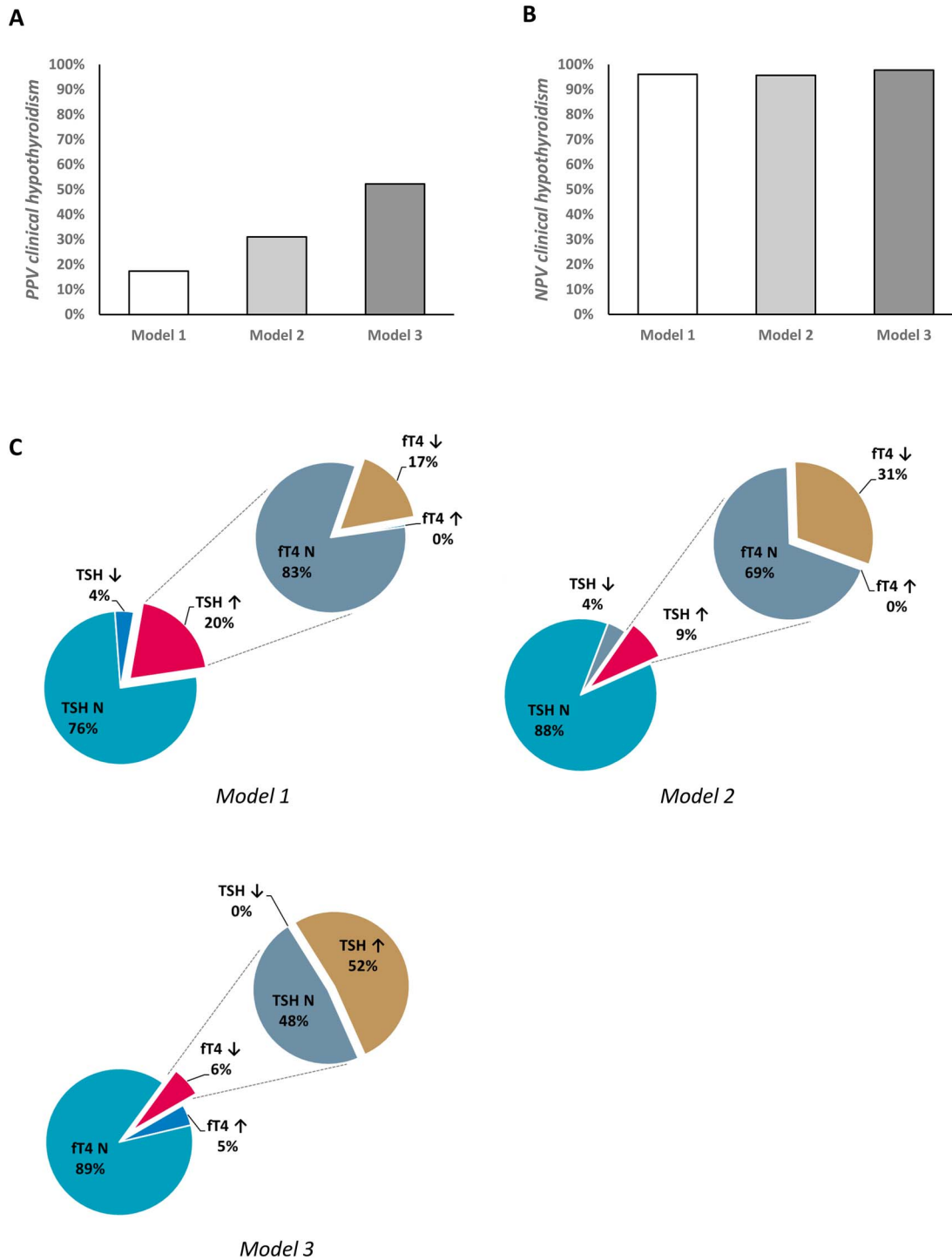
## Results

The dataset existed of 1,016 older adults, of which  $n = 35$  (3.4%) were diagnosed with clinical hypothyroidism (TSH > 4.2 mU/l and fT4 < 12 pmol/l). When data were analysed based on the current guidelines (model 1, fT4 reflex test following TSH < 0.27 or > 4.2 mU/l), fT4 reflex testing followed TSH analysis in 242 individuals (23.8%), of which 202 (19.9%) had TSH values above the reference interval and 40 (3.9%) had TSH values below the reference interval. Of the individuals with elevated TSH values, 35 (17.3%) also showed decreased fT4 values. Hence, the PPV for TSH as a screening test for clinical hypothyroidism was 17.3%. The NPV was 96.1% (Figure 1 and Table 1).

Using an optimized TSH cutoff value of 6.5 mU/l (model 2) (ROC curves in Supplementary Figure 1), the number of reflex tests decreased to 127 (12.5%). This resulted in an improved PPV of 31.0% and an NPV of 95.7%. Raising the TSH cutoff also reduced the number of individuals diagnosed with clinical hypothyroidism (27 individuals versus 35 in model 1) (Figure 1 and Table 1).

Using fT4 as the screening test (model 3) the number of TSH reflex tests was 112 (11.2%), the PPV 52.2% and the NPV for overt clinical hypothyroidism was 97.8% (Figure 1 and Table 1). In this scenario, the number of individuals diagnosed with clinical hypothyroidism was identical to model 1.

With regard to hyperthyroidism, despite low numbers in the study population, the proposed strategy with screening based on fT4 (model 3) did not lead to missed cases of clinical hyperthyroidism (Supplementary Table 1), since



**Figure 1.** Graphic display of the PPV (A), NPV (B) and (C) the proportions of subjects with deviating values of TSH and fT4, and hence the proportion of subjects diagnosed with (subclinical) hypothyroidism, in the different models. *N* = normal (value within the reference range).

cutoff values for elevated fT4 and reduced TSH were not changed.

### Discussion

Given the findings, that treating SCHAT with L-Thyroxine in older people provides no benefit [6, 8, 9], efficacy of current

screening strategy based on TSH needs reassessment. In this retrospective study in older adults in primary care, the fT4-based screening approach reduced reflex testing by >50% compared with current TSH-based screening strategy. The PPV for clinical hypothyroidism of the initial screening result increased from 17.3% to 52.2%. The NPV was not hampered with an NPV of 96.1% for TSH-based screening

**Table 1.** Comparison of the performances of the four proposed strategies

	Current screening protocol: TSH > 4.2 mU/l model 1	Elevation of the TSH cutoff value to > 6.5 mU/l model 2	Screening based on fT4: fT4 < 12.0 pmol/l model 3
<b>Total number of reflex tests</b>	242 (23.8%)	127 (12.5%)	114 (11.2%)
<b>Reflex tests for hypothyroidism</b>	202 (19.9%)	87 (8.6%)	67 (6.6%)
<b>Clinical hypothyroidism</b>	35 (3.4%)	27 (2.7%)	35 (3.4%)
<b>SCHT</b>	167 (16.4%)	60 (5.9%)	–
<b>PPV clin. hypothyroidism</b>	17.3%	31.0%	52.2%
<b>NPV clin. hypothyroidism</b>	96.1%	95.7%	97.8% <sup>a</sup>
<b>Overall diagnostic efficiency</b>	80.4%	90.2%	94.8%
<b>Sensitivity clin. hypothyroidism</b>	52.5%	40.3%	62.5% <sup>a</sup>
<b>Specificity clin. hypothyroidism</b>	82.4%	93.7%	96.7%

<sup>a</sup>based on overt hypothyroidism (TSH > 10 mU/l).

and 97.8% for fT4-based screening. Earlier studies that suggested a testing algorithm based on fT4 date from the 80's and 90's, when screening for thyroid disease was not as frequent and a variety of biochemical- or immuno-assays was used [10–12]. Preliminary retrospective results of an ongoing population cohort study seem to confirm our results (Manuscript in preparation, M.E. Jongeneelen and Prof. Dr. J. Gussekloo).

A reduction in reflex testing could result in significant reductions in costs and patient concerns. Since treatment of SCHT is not indicated [6, 8, 9], this approach may reduce initiation of levothyroxine treatment and it would result in less frequent follow-up.

Additional benefit of an fT4-based approach is that hypothalamic/pituitary hypothyroidism, probably missed by TSH-based screening, is detected. However, using an fT4-based approach, individuals with a TSH value >10 mU/l (with a presumed higher risk of progression to clinical hypothyroidism) and individuals with triiodothyronine toxicosis could potentially be missed. Yet, the latter group is usually clinically identifiable and will probably not be included in initial screening in primary care. With regard to individuals with a TSH value >10 mU/l, individuals should be advised to see their GP again when complaints persist and should then be offered a combined fT4–TSH test. Likewise, when there is a high suspicion of thyroid disease, a direct combination of fT4 plus TSH should be offered.

It should also be noted that chronic disease, frequent in older adults, could be associated with non-thyroidal illness (NTI) and in severe cases may result in decreased fT4, thus generating false positives. Yet, TSH elevations are often found in the recovery phase after NTI mimicking SCHT [14].

Alternatively, the PPV of current screening for thyroid dysfunction could be improved by elevation of the TSH cut-off value. Several other authors have proposed this approach [13, 15, 16]. However, with this strategy (model 2), individuals with hypothyroidism and a limited elevation of TSH would be missed. Furthermore, in our population the increase in PPV was only limited compared with the fT4-based strategy.

Some limitations have to be acknowledged. Firstly, due to the retrospective nature of the study, we cannot exclude that

individuals included in the study may have had a history of thyroid disease, which we could not filter from the available laboratory data, for example in cases where patients have visited other laboratories for checkups. Second, the study group was relatively small. Further prospective studies with sufficient power are needed to validate our findings. Third, the determination of optimal cutoff values for TSH and fT4 (which currently vary significantly within geographical areas and within method groups) should also be investigated in follow-up studies. Yet, this is hampered by the fact that a gold standard test for hypothyroidism is lacking, the combination of test results from two tests, fT4 and TSH, determines the diagnosis.

In conclusion, our results suggest that screening for thyroid dysfunction in older individuals in primary care can be improved by screening based on fT4 instead of TSH or by adjusting the TSH cutoff value. Adjustment of the screening strategy may be of interest to health policy makers because of potential cost reduction. From a patient's perspective, medical concerns and unnecessary biochemical follow up might be reduced by circumventing the diagnosis SCHT.

**Supplementary Data:** Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

**Acknowledgements:** Dr W. den Elzen (Dept Clinical Chemistry, Amsterdam University Medical Center) and Prof. Dr J. Gussekloo (Dept Public Health and Primary Care, Leiden University Medical Center) are kindly acknowledged for their conceptual contribution to this study.

**Declaration of Conflicts of Interest:** None

**Declaration of Sources of Funding:** None

## References

1. Sheehan M. Biochemical testing of the thyroid: TSH is the best and, oftentimes, only test needed – a review for primary care. *Clin Med Res* 2016; 14: 83–92.
2. 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: 650–9.

3. Tsuda S, Nakayama M, Matsukuma Y *et al.* SCHT is independently associated with poor renal outcomes in patients with chronic kidney disease. *Endocrine* 2021; 73: 141–50.
4. Iglesias P, Bajo MA, Selgas R, Díez JJ. Thyroid dysfunction and kidney disease: an update. *Rev Endocr Metab Disord* 2017; 18: 131–44.
5. Zijlstra LE, Jukema JW, Westendorp RGJ *et al.* Levothyroxine treatment and cardiovascular outcomes in older people with SCHT: pooled individual results of two randomised controlled trials. *Front Endocrinol* 2021; 12: 674841. <https://doi.org/10.3389/fendo.2021.674841>.
6. Stott DJ, Rodondi N, Kearney PM *et al.* Thyroid hormone therapy for older adults with SCHT. *N Engl J Med* 2017; 376: 2534–44.
7. Biondi B, Cappola AR, Cooper DS. SCHT: a review. *JAMA* 2019; 322: 153–60.
8. Bekkering GE, Agoritsas T, Lytvyn L *et al.* Thyroid hormones treatment for SCHT: a clinical practice guideline. *BMJ* 2019; 365: 1–9.
9. 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 SCHT. *JAMA* 2019; 322: 1977–86.
10. De los Santos ET, Starich GH, Mazzaferri EL. Sensitivity, specificity, and cost-effectiveness of the sensitive thyrotropin assay in the diagnosis of thyroid disease in ambulatory patients. *Arch Intern Med* 1989; 149: 526–32.
11. Goldstein BJ, Mushlin AI. Use of a single thyroxine test to evaluate ambulatory medical patients for suspected hypothyroidism. *J Gen Intern Med* 1987; 2: 20–4.
12. Helfand MSJ. Screening for thyroid dysfunction: which test is best? *JAMA* 1993; 270: 2297–8.
13. Henze M, Brown SJ, Hadlow NC, Walsh JP. Rationalizing thyroid function testing: Which TSH cutoffs are optimal for testing free T4? *J Clin Endocrinol Metabol* 2017; 102: 4235–41.
14. Gurnell M, Halsall DJ, Chatterjee VK. What should be done when thyroid function tests do not make sense? *Clin Endocrinol (Oxf)* 2011; 74: 673–8.
15. Vadiveloo T, Donnan PT, Murphy MJ, Leese GP. Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in tayside, Scotland: the thyroid epidemiology, audit, and research study (TEARS). *J Clin Endocrinol Metabol* 2013; 98: 1147–53.
16. Fatourechi V. Editorial: upper limit of normal serum thyroid-stimulating hormone: a moving and now an aging target? *J Clin Endocrinol Metabol* 2007; 92: 4560–2.

**Received 5 January 2022; editorial decision 22 July 2022**