1 Subclinical Thyroid Dysfunction and the Risk of Cognitive Decline: a Meta-Analysis of Prospective Cohort Studies

2 Authors: Carole Rieben, MD; Daniel Segna, MD; Bruno R. da Costa, PhD; Tinh-Hai Collet, MD; Layal Chaker, MD; Carole E. Aubert, MD; Christine Baumgartner, MD; Osvaldo P. Almeida, MD, PhD; Eef Hogervorst, PhD; Stella Trompet, PhD; Kamal Masaki, MD; Simon P. Mooijaart, MD, PhD; Jacobijn Gussekloo, MD, PhD; Robin P. Peeters, MD, PhD; Douglas C. Bauer, MD; Drahomir Aujesky, MD, MSc; Nicolas Rodondi, MD, MAS.

3 Affiliations: a Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Switzerland; b Institute of Primary Health Care (BIHAM), Faculty of Medicine, University of Bern, Bern, Switzerland; c Service of Endocrinology, Diabetes and Metabolism, University Hospital of Lausanne, Lausanne, Switzerland; d University of Cambridge Metabolic Research Laboratories, Wellcome Trust-MRC Institute of Metabolic Science, Addenbrooke’s Hospital, Cambridge, UK; e Rotterdam Thyroid Center, Department of Epidemiology, Rotterdam, The Netherlands; f School of Psychiatry and Clinical Neurosciences, University of Western Australia, Perth, Australia; g School of Sport, Exercise and Health Sciences, Loughborough University, Leicestershire, UK; h Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands; i Kuakini Medical Center and the Department of Geriatric Medicine, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii, USA; j Department of Gerontology and Geriatrics, LUMC, Leiden and Institute for Evidence-based Medicine in Old Age (IEMO), Leiden, The Netherlands; k Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, The Netherlands; l Department of Medicine, University of California, San Francisco, USA.

4 Corresponding author and address for reprint requests: Nicolas Rodondi, MD, MAS, Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, 3010 Bern, Switzerland; e-mail: nicolas.rodondi@insel.ch; Phone +41 31 632 41 63; Fax +41 31 632 88 85

5 Abbreviated title: Subclinical Thyroid Dysfunction and Cognitive Decline

6 Key terms: Subclinical thyroid dysfunction, subclinical hyperthyroidism, subclinical hypothyroidism, cognitive decline, dementia

7 Word Count: Text: 3907

8 Abstract: 248

9 Table Count: 2; Figure Count: 2

10 Supplemental Table Count: 2; Supplemental Figures Count: 2

11 Reference Count: 62

12 Financial Support: Swiss National Science Foundation (SNSF 320030-150025).

13 Disclosure summary: The authors have nothing to disclose.

14 Registration: The protocol of this meta-analysis has been published in the PROSPERO (PROSPERO Record: CRD42015019819) register.
ABSTRACT

Context: While both overt hyper- and hypothyroidism are known to lead to cognitive impairment, data on the association between subclinical thyroid dysfunction and cognitive function are conflicting.

Objective: To determine the risk of dementia and cognitive decline associated with subclinical thyroid dysfunction among prospective cohorts.

Data Sources: Search in MEDLINE, EMBASE until November 2014.

Study Selection: Two physicians identified prospective cohorts that assessed thyroid function and cognitive outcomes (dementia; Mini-Mental State Examination, MMSE).

Data Extraction: Data were extracted by one reviewer following standardized protocols and verified by a second reviewer. The primary outcome was dementia and decline in cognitive function was the secondary outcome.

Data Synthesis: Eleven prospective cohorts followed 16,805 participants during a median follow-up of 44.4 months. Five studies analyzed the risk of dementia in subclinical hyperthyroidism (n=6410), six in subclinical hypothyroidism (n=7401). Five studies analyzed MMSE decline in subclinical hyperthyroidism (n=7895), seven in subclinical hypothyroidism (n=8960). In random-effects models, the pooled adjusted RR for dementia in subclinical hyperthyroidism was 1.67 (95% confidence interval [CI] 1.04-2.69) and 1.14 (95%CI 0.84-1.55) in subclinical hypothyroidism versus euthyroidism, both without evidence of significant heterogeneity (I²=0.0%). The pooled mean MMSE decline from baseline to follow-up (mean 32 months) did not significantly differ between subclinical hyper- or hypothyroidism versus euthyroidism.

Conclusions: Subclinical hyperthyroidism might be associated with an elevated risk for dementia, while subclinical hypothyroidism is not, and both conditions are not associated with faster decline in MMSE over time. Available data are limited, and additional large, high-quality studies are needed.
The prevalence of subclinical hypothyroidism (SHypo) reaches up to 10% in the elderly population, while subclinical hyperthyroidism (SHyper) has a prevalence of 2.4%, and 4.3% in the population aged ≥80 years.\textsuperscript{1, 2} SHypo is biochemically defined as elevated serum thyroid-stimulating hormone (TSH, thyrotropin) levels, but free thyroxin (fT4) levels within laboratory reference ranges\textsuperscript{3}, SHyper is defined as decreased serum TSH concentrations and normal fT4 and fT3 levels. The subclinical forms of thyroid dysfunction have previously been associated with increased risk of heart failure and coronary heart disease.\textsuperscript{4-6} Furthermore, SHyper may negatively influence bone and mineral metabolism.\textsuperscript{7}

While both overt hyper- and hypothyroidism are known to lead to cognitive impairment and clinical guidelines recommend screening for thyroid dysfunction among patients with cognitive disorders\textsuperscript{8}, data on the association between subclinical thyroid dysfunction (SCTD) and cognitive function remained conflicting. In the general population, the prevalence of mild cognitive impairment is 3-22%, with a higher prevalence among adults >70 years (14-18%).\textsuperscript{9-11} Mild cognitive impairment, a cognitive decline not normal for age but with essentially preserved functional activities, is believed to be the earliest clinical symptom of cognitive disorders and may be the stage to intervene with preventive therapies.\textsuperscript{11, 12} The progression rate from cognitive impairment to dementia in the general population aged > 65 years is around 6-10% per year.\textsuperscript{11} SHyper has also been associated with increased risk of dementia,\textsuperscript{13} with one retrospective cohort reporting a hazard ratio of 1.6 (95% confidence interval [CI] 1.2-2.3) for dementia.\textsuperscript{14} SHypo might also be associated with alterations in cognitive function,\textsuperscript{13, 15, 16} with one case-control study reporting a nearly 4-fold increase in the odds ratio of dementia (OR=3.8, 95%CI 1.6-9.1).\textsuperscript{17}

However, data on the association between SCTD and cognitive function are conflicting.\textsuperscript{18-20} Two recent meta-analyses yielded discrepant findings for SHypo, one showing a significant risk of cognitive alteration (composite endpoint of reduced Mini-Mental State Examination (MMSE), Wechsler Memory Scale-Revised, total memory quotient and Wechsler Adult Intelligence Scale scores) for SHypo individuals younger than 75 years with an OR of 1.56 (95%CI 1.07-2.27),\textsuperscript{21} whereas the other found no decline in
MMSE in SHypo patients aged ≥60 years (pooled estimate [ES] 0.03, 95%CI −0.001–0.07).\textsuperscript{22} As both meta-analyses were limited by pooling heterogeneous study designs (prospective and retrospective data), and did neither examine the risk of dementia nor cognitive function associated with SHyper, we conducted a meta-analysis to determine whether SHyper and SHypo were associated with an increased risk of dementia or decline in cognitive function in prospective cohorts, the gold standard for observational data.
METHODS

Data sources and Searches
To perform this systematic review, we followed a pre-defined protocol registered on PROSPERO (Record: CRD42015019819). We conducted a systematic literature search in MEDLINE and EMBASE databases from inception until November 2014 searching for articles related to SCTD and cognitive decline and dementia. The Medical Subject Headings (Mesh) in Ovid MEDLINE included “thyroid disease”, “hypothyroidism”, “hyperthyroidism”, “thyroid hormones”, “thyrotropin”, “subclinical hyperthyroidism”, “subclinical hypothyroidism”, “subclinical dysthyroidism”, “subclinical thyroid”, “cognition”, “dementia”, “memory”, “Alzheimer”, “cognitive”, “cohort studies”, “cohort”, “controlled clinical trial”, “epidemiologic methods”, “review”. We applied a cohort filter designed by the British Medical Journal knowledge information specialists but did not use any other filters or restrictions including year limitations or language restrictions. A similar strategy with similar terms was used for EMBASE. Additionally, we searched the bibliographies of included articles, as well as key articles in the field, and contacted several authors for unpublished subgroup data.

Study Selection (Figure 1)
Two reviewers (CR, DS) independently screened titles and abstracts of the search results and selected publications. In a second step, the same two reviewers independently evaluated the full-text publications of the retrieved studies according to the following pre-specified eligibility criteria: prospective cohort studies among participants ≥18 years, including a SCTD and a euthyroid control group with thyroid function tests at baseline and assessment of cognitive function during follow-up, with published risk estimates or sufficient information to calculate them. We excluded studies examining solely participants with overt thyroid disease. Disagreements were resolved by an independent third author (NR). SHyper was defined as decreased or undetectable TSH and normal fT4, and SHypo as elevated TSH and normal fT4. Cohort-specific TSH- and fT4-cutoff levels were used to determine thyroid status. For dementia definition, we accepted all
validated assessments of memory and cognitive function, and did not exclude studies that reported other scales than MMSE. For our analyses, we also collected information on clinical dementia \textit{(Supplement table 1)}. Additionally, we gathered data on MMSE results at baseline and follow-up visits. Studies were included if they provided information on either dementia or MMSE outcomes, or both.

**Data Extraction and Quality Assessment**

Standardized data extraction forms were used to collect information from the included cohorts concerning patient characteristics, thyroid hormone levels, and scales for tests or criteria used to define memory function, dementia or Alzheimer’s disease (AD). Data were extracted by one reviewer (CR) and verified by a second independent reviewer (DS). Discrepancies were solved by a third author (NR). Two reviewers (CR, DS) independently assessed study quality using key indicators of cohort study quality\textsuperscript{24,25}: origin of population (convenience versus population-based, the latter defined as a random sample of the general population), methods of outcome ascertainment and adjudication (considered as adequate if in each potential case performed by an expert panel blinded regarding the thyroid status and following defined outcome criteria), completeness of follow-up, assessment of the proportional hazard assumption and adjustment for confounders.

**Data Synthesis and Statistical Analysis**

We performed four main analyses on the association of 1) subclinical hyperthyroidism and dementia, 2) subclinical hypothyroidism and dementia, 3) subclinical hyperthyroidism and MMSE and 4) subclinical hypothyroidism and MMSE. We expressed the estimates of the association between SCTD (i.e. SHyper or SHypo) and outcomes as risk ratios (RR) for dementia or as between-group differences in mean changes from baseline for MMSE scores (MD). Only prospective data were analyzed. A RR>1 indicates a higher risk of an event in SCTD compared to euthyroids and a MD>0 indicates higher decline of MMSE in SCTD compared to euthyroids. To account for the different lengths of follow-up time across studies, we standardized MD per 1 year unit. We used most adjusted estimates provided by the studies as primary analysis. We
used an inverse variance random-effects meta-analysis to pool estimates across studies. Estimates of the association between SCTD and dementia were pooled on a log scale and latter exponentiated to be reported as RR. To evaluate heterogeneity across the studies, we used the Q statistic with a conservative p-value of 0.10. Furthermore, we calculated the I² statistic, indicating the proportion of variability in estimates of effects across studies that is not due to chance alone (<25% low, 25-75% increasing, >75% high heterogeneity between studies). We visually evaluated publication bias through funnel plots and, statistically, with the Egger’s test. To explore the sources of heterogeneity, we conducted several sensitivity analyses. Due to the small number of studies in each group, subgroup analysis with interaction tests could not be meaningfully performed. All P-values were two-sided. All analyses were conducted using STATA software, version 13.1 (College Station, Texas).
RESULTS

Study Selection

Of the 1505 reports initially identified, 1471 remained after removing duplicates. We excluded 1435 records on the basis of their abstracts and 25 after full text examination (Figure 1). Eleven studies met pre-specified eligibility criteria and were included in the analyses. Among those, 3 studies provided information on both dementia and MMSE outcomes (Supplement Table 1, Section A), 4 studies reported information on dementia outcomes only (Supplement Table 1, Section B), and 4 assessed MMSE outcomes only (Supplement Table 1, Section C). The agreement among the reviewers was 98.63% for the first screen of abstracts ($\kappa=0.75$) and 89.74% for the full-text search ($\kappa=0.71$).

Study Characteristics

Eleven cohorts reported data on 16,805 participants (Table 1). Two cohorts only included men. Mean age was 70 years or higher, except for one study. The follow-up time ranged from 12 to 152 months (median follow-up 44.4 months). Eight cohorts excluded participants treated with thyroid hormones or medications altering thyroid hormone levels, while three excluded the participants taking thyroid hormones or thyroid altering medication in sensitivity analysis.

Description and Quality of Studies

The quality of studies was heterogeneous. Nine cohorts were population-based and two were convenience samples (Supplement Table 1). All the cohorts used third generation TSH assays, except one using second generation tests and one that did not report test details. Four studies had a formal adjudication committee for dementia diagnosis. Seven studies provided information on attrition during follow-up. Six studies provided information on non-violation of the proportional hazard assumption. All studies reported adjusted data with various confounders, except one study that provided us unadjusted data.
**Subclinical Hyperthyroidism and Dementia**

Among five cohorts analyzing the association between SHyper and dementia (n=6410, 329 cases of dementia, mean follow-up 68.3 months),\(^{29-31,37,38}\) the pooled risk ratio [RR] of dementia was 1.67 (95%CI 1.04-2.69, \(I^2=0.0\%\), p for heterogeneity=0.82) among SHyper patients compared with euthyroidism (Figure 2).

Sensitivity analyses (Table 2) excluding one study with a convenience-based sample, one study that followed both patients with and without thyroid hormone replacement, or studies without or not reported formal adjudication for dementia, yielded similar results. As the Framingham study only analyzed the relationship with dementia using TSH tertiles (highest tertile: TSH 1.9-9.9 mU/L) and did not measure fT4,\(^{34}\) we added this study in a sensitivity analysis and found comparable results. A sensitivity analysis excluding 475 overlapping patients between two cohorts\(^{31,38}\) yielded similar results; we did not include these data in the main analysis, as they examined different follow-up duration and were not based on peer-reviewed published results (the investigators sent us these data separately). The relationship between SHyper and AD was assessed by three studies only (n=3186, 108 AD cases, mean follow-up 75.0 months).\(^{30,31,38}\) The pooled RR for AD was 1.67 (95%CI 0.79-3.51, \(I^2=16.8\%\), p for heterogeneity=0.30).

**Subclinical Hypothyroidism and Dementia**

Among six studies analyzing the relationship between SHypo and dementia (n=7401, 416 cases of dementia, mean follow-up 64.6 months),\(^{29-32,37,38}\) the pooled RR for dementia was 1.14 (95%CI 0.84-1.55, \(I^2=0.0\%\), p for heterogeneity=0.49) (Figure 2). No individual study showed a statistically significant association. Sensitivity analyses (Table 2) excluding a study with a convenience-based sample, studies with TSH cut-off <4.5mU/l and potentially including individuals in the euthyroid range, two studies that followed both patients with and without thyroid hormone replacement, studies without or not reported formal adjudication process for dementia, one study with additional unadjusted data, or 475 overlapping participants between two cohorts\(^{31,38}\) yielded similar results. The addition of the Framingham study\(^{34}\) to the main analysis yielded similar results. Four studies analyzed the relationship between SHypo and AD (n=3823,
151 AD cases, mean follow-up 69.36 months. The pooled RR for AD was 0.95 (95%CI 0.52-1.71, $I^2=0.0\%$, p for heterogeneity=0.89).

Subclinical Hyperthyroidism and MMSE

Among five studies reporting change in MMSE among participants with SHyper (n=7895, mean follow-up 33.6 months), the pooled mean MMSE decline in cognitive function per year was 0.01 points difference from baseline (95%CI -0.14-0.15; $I^2=23.5\%$, p for heterogeneity=0.27; Supplement Figure 1). Results remained similar after excluding one study using a convenience-based sample or one study that followed both patients with and without thyroid hormone replacement (Supplement Table 2). Because the results of the main analyses between SHyper and dementia did not seem concordant with the results of the meta-analysis looking at the decrease of MMSE in SHyper participants, we undertook a sensitivity analysis including the two studies examining the relationship of SHyper and both MMSE and dementia, which also showed no larger decline of MMSE among SHyper.

Subclinical Hypothyroidism and MMSE

Among seven studies reporting change in MMSE in SHypo (n=8960; mean follow-up 32.2 months), pooled mean MMSE per year declines did not significantly differ between SHypo and euthyroid groups (ES 0.01 points difference from baseline, 95%CI -0.10-0.12, $I^2=27.6\%$, p for heterogeneity=0.22; Supplement Figure 1). Sensitivity analyses (Supplement Table 2) excluding one study with a convenience-based sample, studies using TSH cut-offs <4.5mU/l, one study that followed both patients with and without thyroid hormone replacement, one study that might have subclinical hyperthyroid participants in the control group, or one study using unadjusted data yielded similar results.

Publication bias
Both graphical inspection and Egger’s tests indicated little evidence of publication bias for all associations, although the number of available studies was small (Supplement Figure 2).
DISCUSSION AND CONCLUSION

In this meta-analysis of 11 prospective cohorts, we found that SHyper, but not SHypo, might be associated with an elevated risk for dementia, while decline in MMSE over time was minimal for both conditions. SHyper showed also a similar pattern of higher risk for AD. Results for the association between SHyper and dementia remained similar when we pooled higher quality studies in sensitivity analysis, such as studies with formal adjudication process for the outcome assessment or population-based studies.

Our results for SHyper and risk for dementia are consistent with a non-systematic literature review summarizing results from 13 cross-sectional or case-control, and 10 cohort studies that found supportive evidence of an association between SHyper and cognitive impairment or dementia. Of these 10 cohort studies, four did not meet the eligibility criteria for our systematic review: one due to missing subgroups of thyroid dysfunction, two analyzed only euthyroid participants and one had a retrospective design. Several other individual studies reported an association between SHyper and an elevated risk for dementia as well: a retrospective nested case-control study including 2004 patients with SHyper reported a hazard ratio for dementia of 1.79 (95%CI 1.28-2.51), and a cross-sectional study found a positive association between SHyper and dementia in 1276 participants (33 with SHyper) aged ≥65 years (OR for dementia 4.1, 95%CI 1.3-13.1). Van Osch et al prospectively studied 178 patients with AD and 291 community-dwelling controls without objective cognitive impairment, and found an adjusted OR for AD of 2.36 (95%CI 1.19-4.67) in participants in the lowest (TSH<1.3mU/l) versus highest TSH tertile (TSH>2.1mU/l). Another population-based prospective cohort of 313 elderly adults with normal TSH that found that those with a decline of cognitive dysfunction had a mean TSH of 1.78mU/l, while those without decline had a mean TSH of 2.24mU/l (p=0.001).

Our findings might be consistent with the hypothesis that SHyper increases the risk of dementia, although decline in MMSE over time did not differ between SHyper and euthyroidism. In our meta-analysis, only two out of 11 studies published results on both dementia and MMSE in SHyper. Analyzing only these two studies showed no larger decline of MMSE among participants with SHyper. This discrepancy might be
explained by several factors: the length of follow-up of studies on SCTD and dementia was twice the duration of studies on SCTD and MMSE (mean follow-up time 66 vs. 33 months), different population investigated, the modest sensitivity of MMSE as a diagnostic tool (79%)\textsuperscript{47,48}, as well as for detecting mild cognitive impairment and subtle changes in specific cognitive domains, and the multimodal approach needed to diagnose dementia.\textsuperscript{49} Furthermore, different scores were used as gold standard depending on the type and version of diagnostic criteria (supplement table 1). Factors increasing the plausibility of the association between SHyper and dementia were that results remained similar when we pooled higher quality studies in sensitivity analysis, such as studies with formal adjudication process for the outcome assessment or population-based studies, and that SHyper also showed a pattern of higher risk for AD. However, higher quality studies are needed.

Several pathways could explain the association of thyroid dysfunction with cognition and dementia. Thyroid hormone has distinct effects on the cardiovascular system and thyroid dysfunction has been associated with several cardiovascular risk factors, including hypertension, dyslipidemia and atrial fibrillation.\textsuperscript{4,6} In turn, these cardiovascular risk factors are associated with a higher risk of dementia and Alzheimer’s Disease.\textsuperscript{50} Most studies included in our meta-analysis adjusted for cardiovascular risk factors. However, the number and type of variables that were adjusted for differed for each study. Other explanations for the association include direct effects of thyroid hormone, such as neurotoxicity and altered gene expression in pathways relevant for dementia. The exact pathophysiological link between thyroid dysfunction and dementia remains unclear and requires more research.

Recently, two meta analyses on SHypo and cognitive impairment were published, yielding discrepant results.\textsuperscript{21,22} The first review included 13 studies and found a significant higher risk for cognitive alteration (composite endpoint of incidence or prevalence of dementia or difference in MMSE, Wechsler Adult Intelligence scale and Wechsler Memory-Revised score) in SHypo individuals younger than 75 years (OR 1.56; 95%CI 1.07-2.27, p=0.02), and for dementia (OR 1.81; 95%CI 1.43-2.28, p<0.01).\textsuperscript{21} However, the authors
pooled different designs (cross sectional, case-control, cohort studies), used a composite endpoint of clinical
events and scales as primary outcome, and found a significant risk for the primary endpoint only in subclin-
ical hypothyroid individuals younger than 75 years. As results were calculated on the basis of mean age,
without availability of individual patient data, they may have been subject to potential aggregation bias
( ecological fallacy). Contrary to that meta-analysis, all studies in our meta-analysis but one (included
only in a sensitivity analysis) measured fT4 to define SCTD. The second meta-analysis analyzed 15 studies
(9 cross-sectional, 6 prospective cohort studies) and found no association between SHypo and decline in
cognitive function among people aged > 60 years (cross-sectional analysis: pooled ES for MMSE −0.01
points difference from baseline [95%CI −0.09-0.08]; prospective analysis, pooled MMSE change: 0.03
[95%CI −0.001-0.07] p=0.055, with heterogeneity [I²] of <0.001%),22 which is consistent with our findings.
In comparison to these two meta-analyses, we included only prospective cohorts (n=11) allowing us to re-
duce the bias that could arise due to differing study designs. To make literature search broad enough, we
excluded studies examining solely participants with overt thyroid disease but added no other exclusion cri-
teria.
Two small placebo controlled trials (n=89; n=94) found no evidence that treatment of SHypo with levothy-
roxine was associated with improved cognitive function.18,52 However, these trials had several limitations.
In the trial by Parle et al,52 recruitment was based on a single thyroid function test, so that euthyroid partic-
pants with transiently elevated TSH may have been included (50% in the placebo group were euthyroid at
12 months), which may have underpowered the trial to detect an effect of hormone replacement.52 Thyroxin
substitution lasted only for 12-months, which may have been too short to affect cognitive function. In the
trial by Jorde et al,18 one third of participants did not attend follow-up. Because of numerous exclusion cri-
teria, the study population was unusually healthy, with 57% of the participants having a TSH value between
3.50 and 4.99mU/l, so that it probably included many euthyroid participants. The ongoing TRUST (Thyroid
Hormone Replacement for Subclinical Hypothyroidism) trial (ClinicalTrials.gov: NCT01660126) may clar-
ify whether treatment with levothyroxine in SHypo is associated with better cognitive outcomes over time.53
There are several strengths of our meta-analysis. By combining the results of 11 prospective cohorts, we analyzed a total of 432 cases of dementia and 160 cases of AD in more than 15,000 participants. By contacting several authors of these studies, we obtained additional data that allowed us to derive more uniform subgroup and sensitivity analyses. In comparison to the two other meta-analyses, our results are less prone to bias due to pooling of heterogeneous study design and quality, because we only included prospective cohorts. We also conducted a detailed literature search in several electronic databases with as few limitations as possible in order to retrieve the maximum number of studies available on the topic, and were able to include a larger number of prospective cohorts than previous meta-analyses.

Our meta-analysis has also several limitations. Except for two studies, studies only examined Caucasians, limiting the generalizability to other populations. All studies were performed in participants with a mean age ≥65 years and the time of follow-up was relatively short, ranging between 12 and 70.8 months (152.4 months in the Framingham Study, added in a sensitivity analysis). All but two studies assessed thyroid function tests only at baseline, which is a limitation of most previously published large cohort studies on the risks of thyroid dysfunction. Some participants with SCTD at baseline may have normalized to euthyroidism or progressed to overt thyroid disease over time. Regarding the elderly participants in the included studies, we cannot exclude a certain degree of overdiagnosis of SHypo due to the physiological rise of TSH towards upper limits during normal ageing. All these non-differential misclassification of thyroid status might bias the results towards no difference. The limited sensitivity of MMSE for detecting subtle changes in specific cognitive domains may further limit our ability to detect a possible decline in cognitive function. A meta-analysis of observational studies requires cautious interpretation of the results and potential biases, and confounding and heterogeneity must be carefully investigated. The quality of the incorporated studies was variable. We performed several sensitivity analyses to address differences between the studies, as recommended, although they should be interpreted with caution given the small number of studies. In study level meta-analysis, interpretation of subgroup data should be performed with caution. Because of the small amount of studies, no meaningful subgroup analysis could be performed.
There are multiple confounders for cognitive decline and dementia, the most important is age, others are depression/mood or cardiometabolic risk factors. All cohorts adjusted for age and several other confounders, but there was heterogeneity in the choice of confounders, which may lead to residual confounding. Bias in the selection of included studies cannot be excluded. To limit selection bias, we conducted a detailed literature search in several electronic databases with broad inclusion. We performed graphical and statistical assessment to assess selective reporting, but these analyses were not very sensitive considering the small number of studies included. Although included cohorts enrolled community-dwelling adults in ambulatory visits, who are therefore less likely to have an acute disease, some participants with non-thyroidal illness may have been analyzed. Included studies addressed this problem differently: Two repeatedly measured thyroid values, one assessed and adjusted for rT3 (reverse triiodothyronine), and others adjusted for comorbidities. We cannot exclude that some participants had nonthyroidal illness.

In summary, our systematic review and meta-analysis indicates that Shyper, but not Shypo, might be associated with a modestly elevated risk of dementia. Neither Shyper nor Shypo were significantly associated with a faster decline in MMSE over time, as compared to euthyroidism. Available data were limited, and additional large, high-quality prospective cohort studies are needed.
Financial Support

Prof. N. Rodondi’s work is supported by a grant from the Swiss National Science Foundation (SNSF 320030-150025). Dr Tinh-Hai Collet’s research is supported by grants from the Swiss National Science Foundation (PBLAP3-145870, P3SMP3-155318).

Acknowledgements

The authors thank Francesc Formiga Pérez (Bellvitge University Hospital, Barcelona, Spain), Paola Forti (Department of Medical and Surgical Sciences, Bologna, Italy), Mohammad Arfan Ikram (Erasmus Medical Center, Rotterdam, The Netherlands), Bu Yeap (University of Western Australia and Fiona Stanley and Freemantle Hospitals, Perth, Australia) for supplying additional data from their studies.
References


509. Surks MI, Ortiz E, Daniels GH, et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. *Jama.* 2004;291(2):228-238.


## Table 1. Description of Included Studies for the Effect of Subclinical Thyroid Dysfunction on Dementia/Mini-Mental State Examination (MMSE)

<table>
<thead>
<tr>
<th>Study, Year of publication</th>
<th>Population</th>
<th>Women</th>
<th>Mean age; SD</th>
<th>Follow-up time</th>
<th>Age min-max. years</th>
<th>TSH cutoff level (mU/l)</th>
<th>fT4 measured</th>
<th>Thyroid hormone recipients excluded?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotterdam, 2000</td>
<td>1843</td>
<td>61.9</td>
<td>68.8; 7.5</td>
<td>25.2</td>
<td>55-93</td>
<td>&gt; 4.0</td>
<td>&lt; 0.4</td>
<td>yes</td>
</tr>
<tr>
<td>Leiden 85-Plus Study, 2004</td>
<td>558</td>
<td>66.0</td>
<td>85.0; 0.0</td>
<td>44.4</td>
<td>85</td>
<td>&gt; 4.8</td>
<td>&lt; 0.3</td>
<td>yes; in SA</td>
</tr>
<tr>
<td>Rotterdam Scan, 2006</td>
<td>1077</td>
<td>51.2</td>
<td>72.3; 7.4</td>
<td>66.0</td>
<td>60-90</td>
<td>&gt; 4.3</td>
<td>&lt; 0.4</td>
<td>yes</td>
</tr>
<tr>
<td>Health Ageing, 2008</td>
<td>1047</td>
<td>51.0</td>
<td>73.6; 6.2</td>
<td>24.0</td>
<td>64-94</td>
<td>&gt; 4.8</td>
<td>&lt; 0.3</td>
<td>yes</td>
</tr>
<tr>
<td>Framingham, 2008</td>
<td>1864</td>
<td>59.0</td>
<td>71.0; 7.0</td>
<td>152.4</td>
<td></td>
<td></td>
<td>1</td>
<td>no; in SA</td>
</tr>
<tr>
<td>HAAS, 2009</td>
<td>665</td>
<td>0.0</td>
<td>78.0; 4.7</td>
<td>56.4</td>
<td>71-93</td>
<td>&gt; 4.3</td>
<td>&lt; 0.4</td>
<td>yes</td>
</tr>
<tr>
<td>Japanese Study, 2010</td>
<td>229</td>
<td>65.0</td>
<td>80.9; 4.7</td>
<td>12.0</td>
<td></td>
<td>&gt; 4.0</td>
<td>NR</td>
<td>yes</td>
</tr>
<tr>
<td>Conselice, 2012</td>
<td>660</td>
<td>52.9</td>
<td>73.3; 6.0</td>
<td>45.6</td>
<td>65-91</td>
<td>&gt; 4.5</td>
<td>&lt; 0.45</td>
<td>yes; in SA</td>
</tr>
<tr>
<td>HIMS, 2012</td>
<td>3401</td>
<td>0.0</td>
<td>76.8; 3.5</td>
<td>70.8†</td>
<td>70-89</td>
<td>&gt; 4.0</td>
<td>&lt; 0.4</td>
<td>yes</td>
</tr>
<tr>
<td>PROSPER, 2013</td>
<td>5154</td>
<td>49.4</td>
<td>75.0</td>
<td>38.4</td>
<td>80-82</td>
<td>&gt; 4.5</td>
<td>&lt; 0.45</td>
<td>yes</td>
</tr>
<tr>
<td>OCTABAIX, 2014</td>
<td>307</td>
<td>54.6</td>
<td>85.0; 0.0</td>
<td>36.0</td>
<td>85</td>
<td>&gt; 5</td>
<td>&lt; 0.25</td>
<td>yes</td>
</tr>
</tbody>
</table>

**Abbreviations:** Conselice = Conselice Study of Brain Ageing; Framingham = The Framingham Study; fT4 = free thyroxine; HAAS = Honolulu-Asia Aging Study; Health Ageing = Health Ageing Study; HIMS = The Health in Men Study; Japanese Study = Cognitive function with subclinical hypothyroidism in elderly people without dementia: One year follow up; Leiden 85+ = Leiden 85-plus Study; NR = not reported; OCTABAIX = OCTABAIX Study; PROSPER = The PROSPER Study; Rotterdam Scan = Rotterdam Scan Study; Rotterdam = The Rotterdam Study; SA = sensitivity analysis; SHyper = subclinical hyperthyroidism; SHypo = subclinical hypothyroidism; TSH = thyrotropin.

† median

‡ The Framingham Study did not use TSH cut-offs for SCTD, but tertiles: tertile 1: 0.1-1.08mU/l for women, 0.10-0.90mU/l for men; tertile 2: 1.10-2.03mU/l for women, 0.99-1.80mU/l for men; tertile 3: 2.10-9.90mU/l for women, 1.09-9.90mU/l for men. Therefore, this study could not be included in the meta-analysis but was added to a sensitivity analysis.

§ Due to additional unpublished data provided by the authors, the studies could be incorporated in the meta-analysis on SCTD and MMSE; unadjusted data.

|| Due to additional unpublished data provided by the authors, the study could be incorporated in the meta-analysis on SCTD and dementia.
## Supplement Table 1: Quality Assessment of Included Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Population, setting</th>
<th>Assessment of dementia or Alzheimer’s Disease</th>
<th>Formal adjudication for dementia or Alzheimer’s Disease</th>
<th>Blinding of study investigators regarding TSH</th>
<th>Blinding of patients regarding TSH</th>
<th>Loss of FUP (reasons for incompleteness)</th>
<th>TSH assessment</th>
<th>TSH assay</th>
<th>Adjustment for potential confounders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Studies that provided information on both dementia and MMSE outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotterdam, 31 2000</td>
<td>population-based, ambulatory / subjects in institutions</td>
<td>D: DSM-III-R (3-step-screening: 1. MMSE (&lt;26), GMSS (&lt;0); 2. CAMDEX; 3. Neurologist, Neuropsychologist, Brain Image); AD: NINCDS-ADRDA</td>
<td>yes</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>BL</td>
<td>3rd</td>
<td>Age, sex, atrial fibrillation, education (years), cigarette smoking (never, former, current), depressive symptoms, APOE</td>
</tr>
<tr>
<td>Consetice, 32 2012</td>
<td>population-based, ambulatory</td>
<td>D: DSM-IV (Multi step screening: 1. Interview, IADL, GDS, MMSE (&lt;24; &gt;9), neurological examination, blood test; 2. Neuropsychological testing with Mental Deterioration Battery, Prose Memory Test); AD: NINCDS-ADRDA</td>
<td>yes</td>
<td>no</td>
<td>NR</td>
<td>NR</td>
<td>BL</td>
<td>3rd</td>
<td>*</td>
</tr>
<tr>
<td>OCTA-BAIX, 37 2014</td>
<td>population-based, ambulatory</td>
<td>NR (Assessments: BI, BADL, LI; MMSE &lt;23/35 pts; Mini Nutritional Assessment, Gait Rating Scale, QoL-VAS, CS)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>BL</td>
<td>3rd</td>
<td>**, Sex, education, marital status, cardiovascular risk factors such as treatment for high blood pressure above 140/90, diabetes mellitus, dyslipidemia, heart failure, medication, prevalence for stroke or dementia, number of drugs prescribed</td>
</tr>
<tr>
<td><strong>B. Studies that provided information on dementia outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rotterdam Scan, 38 2006</td>
<td>population-based, ambulatory / subjects in institutions</td>
<td>D: DSM-III-R (MR for all study subjects, then three step protocol for screening: 1. Screening MMSE (&lt;26) and GMSS (&lt;0); 2. CAMDEX; 3. Neuropsychological testing); AD: NINCDS-ADRDA</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>BL</td>
<td>3rd</td>
<td>Age, sex, educational level, depressive symptoms, cigarette smoking, cardiac medication, beta-blockers, systemic corticosteroid use, atrial fibrillation, diabetes mellitus, BMI; total cholesterol and HDL levels, creatinine, homocysteine level, T3/rT3, TPO-AB</td>
</tr>
<tr>
<td>Study</td>
<td>Population, setting</td>
<td>Assessment of dementia or Alzheimer’s Disease</td>
<td>Formal adjudication for dementia or Alzheimer’s Disease</td>
<td>Blinding of study investigators regarding TSH</td>
<td>Blinding of patients regarding TSH</td>
<td>Loss of FUP (reasons for incompleteness)</td>
<td>TSH assessment</td>
<td>TSH assay</td>
<td>Adjustment for potential confounders</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------</td>
<td>---------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
<td>------------------------------------------</td>
<td>----------------</td>
<td>-----------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td><strong>B. Studies that provided information on dementia outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAAS, 30, 2009</td>
<td>population-based, ambulatory</td>
<td>D: DSM-III-R (Multi step protocol: 1. CASI; 2. CERAD; 3. Neurologic Exam and Interview. In case of Dementia: neuroimaging and blood test); MMSE derived from CASI. AD: NINCDS-ADRDA; CERAD</td>
<td>yes</td>
<td>NR</td>
<td>NR</td>
<td>335 (dementia at BL [131], death or refusal [204])</td>
<td>BL</td>
<td>3rd</td>
<td>Age, age at death (autopsy-study), educational level, depressive symptoms, diabetes mellitus, smoking status, systolic and diastolic blood pressure, use of thyroid medication, thyroid-altering drugs (incl. beta-blocking agents and antiarrhythmics), BMI, biochemical markers (albumin, total and HDL cholesterol, APOE).</td>
</tr>
<tr>
<td>HIMS, 29, 2012</td>
<td>convenience-based, ambulatory</td>
<td>D: ICD-9/10</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>2848 (exclusion due to: history of thyroid disease [2139], thyroid drugs [71], missing data [82], hyperthyroidism [14], prevalent dementia [12], with SMMSE &lt; 24 [521], fT4 &lt; 9 pmol/l [1], fT4 &gt; 25 pmol/l [8])</td>
<td>BL</td>
<td>3rd</td>
<td>Age, BMI, smoking-status, education, SMMSE, social support, medical comorbidities, sensorial impairment</td>
</tr>
<tr>
<td>Framingham, 34, 2008</td>
<td>population-based, ambulatory</td>
<td>D: DSM-IV and Clinical Dementia Rating ≥1 and symptoms of dementia ≥ 6 months (Multi step protocol: 1. Neuropsychological Testing, MMSE; 2. Neurological- and neuropsychological examinations); AD: NINCDS-ADRDA</td>
<td>yes</td>
<td>NR</td>
<td>NR</td>
<td>295 (reason unclear)</td>
<td>2 years before BL</td>
<td>3rd</td>
<td>Age, APOE, education level (dichotomized: high school completion yes/no), plasma homocysteine, current smoking, body-mass index, prevalent stroke, atrial fibrillation</td>
</tr>
<tr>
<td><strong>C. Studies that provided information only on MMSE outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leiden 85-Plus Study, 33, 2004</td>
<td>population-based, ambulatory</td>
<td>no assessment of dementia or AD, only MMSE outcomes reported</td>
<td>dementia not assessed</td>
<td>NR</td>
<td>NR</td>
<td>279 (death [209], refusal [70])</td>
<td>BL + 3 year FUP</td>
<td>3rd</td>
<td><strong>Sex, self-reported educational level (&lt;6 vs &gt; 6 years of education), albumin, CRP, MMSE at baseline, subjective health, type 2 diabetes, myocardial infarction, stroke, COPD, arthritis, Parkinson disease</strong></td>
</tr>
<tr>
<td>Health Ageing, 36, 2008</td>
<td>population-based, ambulatory</td>
<td>no assessment of dementia or AD, only MMSE outcomes reported</td>
<td>dementia not assessed</td>
<td>yes</td>
<td>NR</td>
<td>148</td>
<td>BL</td>
<td>2nd</td>
<td>Age, sex, education, mood, baseline MMSE, high blood pressure, smoking, history of diabetes mellitus, heart attack, stroke, study site</td>
</tr>
<tr>
<td>Study</td>
<td>Population, setting</td>
<td>Assessment of dementia or Alzheimer's Disease</td>
<td>Formal adjudication for dementia or Alzheimer's Disease</td>
<td>Blinding of study investigators regarding TSH</td>
<td>Blinding of patients regarding TSH</td>
<td>Loss of FUP (reasons for incompleteness)</td>
<td>TSH assessment</td>
<td>TSH assay</td>
<td>Adjustment for potential confounders</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------------</td>
<td>----------------</td>
<td>----------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Japanese Study,²⁵ 2010</td>
<td>population-based, ambulatory</td>
<td>no assessment of dementia or AD, only MMSE outcomes reported</td>
<td>dementia not assessed</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>650 (exclusion at BL: antithyroid drugs [6], levothyroxine [159], amiodarone [20], high fT4 and high TSH [7], low fT4 and low TSH [1], overt thyroid disease [146], missing fT4 [303], missing TSH [8])</td>
<td>BL</td>
<td>NR</td>
</tr>
<tr>
<td>PROSPER,²⁰ 2013</td>
<td>convenience-based, ambulatory</td>
<td>no assessment of dementia or AD, only MMSE outcomes reported</td>
<td>dementia not assessed</td>
<td>NR</td>
<td>NR</td>
<td>650 (exclusion at BL: antithyroid drugs [6], levothyroxine [159], amiodarone [20], high fT4 and high TSH [7], low fT4 and low TSH [1], overt thyroid disease [146], missing fT4 [303], missing TSH [8])</td>
<td>BL at 6 Mont h</td>
<td>3rd</td>
<td>Sex, age, education, country, Apo E genotype</td>
</tr>
</tbody>
</table>

**Abbreviations:** AD = Alzheimer's disease; APOE = Apolipoprotein-E Ɛ4 allele status; BADL = Basic Activities of Daily Living; BI = Barthel Index; BL = Baseline; BMI = Body mass index; CAMDEX = Cambridge examination for mental disorders of the elderly; CASI = Cognitive Abilities Screening Instrument; CERAD = Consortium to Establish a Registry for Alzheimer Disease battery; Conselice = Conselice Study of Brain Ageing; COPD = Chronic obstructive pulmonary disease; CRP = C-reactive protein; CS = Charlson Score; D = dementia; DSM = Diagnostic and Statistical Manual of Mental Disorders; Framingham = The Framingham Study; fT4 = Free thyroxine; FUP = Follow-up; GDS = Geriatric Depression Scale; GMSS = Geriatric Mental State Schedule; HAAS = Honolulu-Asia Aging Study; HDL = High-density lipoproteins; Health Ageing = Health Ageing Study; HIMS = The Health in Men Study; IADL = Instrumental Activities of Daily Living; ICD = International Statistical Classification of Diseases and Related Health Problems; Japanese Study = Cognitive function with subclinical hypothyroidism in elderly people without dementia: One year follow up; Leiden 85+ = Leiden 85-Plus Study; LI = Lawton Index; MCI: Mild cognitive impairment; MMSE = Mini Mental State Exam; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; NR = Not reported; OCTABIX = OCTABAIX Study; PROSPER = The PROSPER Study; pts = points; QoL-VAS = Quality of Life Test Visual Analogue Scale; Rotterdam Scan = Rotterdam Scan Study; Rotterdam = The Rotterdam Study; SA = Sensitivity analysis; SHyper = Subclinical hyperthyroidism; SMMSE = Standardized mini mental state exam; TPO-AB = Thyroid peroxidase antibodies; TSH = thyrotropin.

* Unadjusted additional data has been used for this meta-analysis. The original study adjusted for age, sex, education, serum cholesterol, Geriatric Depression Scale Score, BMI, hypertension, diabetes mellitus, history of cardiovascular disease, plasma total homocysteine.

** No adjustment for age because of same aged cohort.
### Table 2: Sensitivity Analyses on the Association between Subclinical Thyroid Dysfunction (SCTD) and Dementia

<table>
<thead>
<tr>
<th>Subclinical Hyperthyroidism and Dementia</th>
<th>RR†</th>
<th>95% CI</th>
<th>p for heterogeneity</th>
<th>N of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis29-31,37,38</td>
<td>1.67</td>
<td>1.04, 2.69</td>
<td>0.82</td>
<td>5</td>
</tr>
<tr>
<td>Exclusion of one study using a convenience-based sample30,31,37,38</td>
<td>1.73</td>
<td>1.07, 2.80</td>
<td>0.82</td>
<td>4</td>
</tr>
<tr>
<td>Exclusion of one study enrolling patients with and without thyroid hormone replacement29,31,37,38</td>
<td>1.73</td>
<td>0.96, 3.11</td>
<td>0.68</td>
<td>4</td>
</tr>
<tr>
<td>Exclusion of studies without formal adjudication30,31</td>
<td>1.86</td>
<td>0.96, 3.62</td>
<td>0.46</td>
<td>2</td>
</tr>
<tr>
<td>Exclusion of overlapping 475 participants from 2 studies29-31,37,38‡</td>
<td>1.60</td>
<td>0.92, 2.78</td>
<td>0.82</td>
<td>5</td>
</tr>
<tr>
<td>Main analysis adding the Framingham Study29-31,34,37,38 §</td>
<td>1.46</td>
<td>1.08, 1.97</td>
<td>0.84</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclinical Hypothyroidism and Dementia</th>
<th>RR†</th>
<th>95% CI</th>
<th>p for heterogeneity</th>
<th>N of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis29-32,37,38</td>
<td>1.14</td>
<td>0.84, 1.55</td>
<td>0.49</td>
<td>6</td>
</tr>
<tr>
<td>Exclusion of one study using a convenience-based population30-32,37,38</td>
<td>1.31</td>
<td>0.91, 1.89</td>
<td>0.65</td>
<td>5</td>
</tr>
<tr>
<td>Exclusion of studies with TSH cut-off &lt;4.5mU/l32,37</td>
<td>1.36</td>
<td>0.91, 2.05</td>
<td>0.59</td>
<td>2</td>
</tr>
<tr>
<td>Exclusion of two studies enrolling patients with and without thyroid hormone replacement29,31,37,38</td>
<td>1.14</td>
<td>0.69, 1.90</td>
<td>0.29</td>
<td>4</td>
</tr>
<tr>
<td>Exclusion of studies without formal adjudication process30-32</td>
<td>1.14</td>
<td>0.73, 1.78</td>
<td>0.57</td>
<td>3</td>
</tr>
<tr>
<td>Exclusion of one study with unadjusted data29-31,37,38</td>
<td>1.06</td>
<td>0.71, 1.60</td>
<td>0.39</td>
<td>5</td>
</tr>
<tr>
<td>Exclusion of overlapping 475 participants from 2 studies29-32,37,38‡</td>
<td>1.10</td>
<td>0.80, 1.50</td>
<td>0.66</td>
<td>5</td>
</tr>
<tr>
<td>Main analysis adding the Framingham Study29-32,34,37,38 §</td>
<td>1.18</td>
<td>0.91, 1.52</td>
<td>0.60</td>
<td>7</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI = confidence interval; N = number; RR = risk ratio; SCTD = subclinical thyroid dysfunction; TSH = thyrotropin.

† RR>1 indicates higher risk of an event in SCTD than in euthyroidism. A positive mean difference indicates larger decrease in cognitive function in SCTD than in euthyroidism.

‡ Performed on data additionally provided by the author; we did not include these data in the main analysis, as they examined different follow-up duration and were not based on peer-reviewed published results (the investigators sent us these data separately).

§ As the Framingham Study did not use TSH cut-off for SCTD, we compared lowest versus highest tertiles (lowest tertile: 0.10-1.08 for women, 0.10-0.90 for men; highest tertile 2.10-9.90 for women, 1.90-9.90 for men). 

34
Supplement Table 2: Sensitivity Analyses on the Association between Subclinical Thyroid Dysfunction and Annualized Mean Change in Mini-Mental State Examination (MMSE)

<table>
<thead>
<tr>
<th>Subclinical Hyperthyroidism and MMSE</th>
<th>ES</th>
<th>95% CI</th>
<th>p for heterogeneity</th>
<th>N. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis20,31,33,36,37</td>
<td>0.01</td>
<td>-0.14, 0.15</td>
<td>0.27</td>
<td>5</td>
</tr>
<tr>
<td>Exclusion of one study using a convenience-based population31,33,36,37</td>
<td>0.06</td>
<td>-0.36, 0.49</td>
<td>0.20</td>
<td>4</td>
</tr>
<tr>
<td>Exclusion of one study enrolling patients with and without thyroid hormone replacement20,31,33,37</td>
<td>-0.01</td>
<td>-0.11, 0.09</td>
<td>0.68</td>
<td>4</td>
</tr>
<tr>
<td>Studies indicating results for both MMSE and dementia31,37</td>
<td>-0.09</td>
<td>-0.26, 0.08</td>
<td>0.80</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclinical Hypothyroidism and MMSE</th>
<th>ES</th>
<th>95% CI</th>
<th>p for heterogeneity</th>
<th>N. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main analysis20,31-33,35,37</td>
<td>0.01</td>
<td>-0.10, 0.12</td>
<td>0.22</td>
<td>7</td>
</tr>
<tr>
<td>Exclusion of one study using a convenience-based population31,33,35,37</td>
<td>0.06</td>
<td>-0.06, 0.18</td>
<td>0.48</td>
<td>6</td>
</tr>
<tr>
<td>Exclusion of studies with TSH cut-off &lt;4.5mU/l20,32,33,36,37</td>
<td>0.07</td>
<td>-0.13, 0.28</td>
<td>0.09</td>
<td>5</td>
</tr>
<tr>
<td>Exclusion of one study enrolling patients with and without thyroid hormone replacement20,31,33,35,37</td>
<td>-0.01</td>
<td>-0.12, 0.11</td>
<td>0.22</td>
<td>6</td>
</tr>
<tr>
<td>Exclusion of one study with unadjusted data20,31,33,35-37</td>
<td>-0.06</td>
<td>-0.13, 0.01</td>
<td>0.70</td>
<td>6</td>
</tr>
<tr>
<td>Studies indicating results for both MMSE and dementia31,32,37</td>
<td>0.06</td>
<td>-0.14, 0.26</td>
<td>0.18</td>
<td>3</td>
</tr>
<tr>
<td>Exclusion of the Japanese Study20,31,33,36,37†</td>
<td>0.01</td>
<td>-0.11, 0.13</td>
<td>0.16</td>
<td>6</td>
</tr>
</tbody>
</table>

**Abbreviations:** MD = mean difference in change from baseline for mini mental state exam score. MD >0 indicates higher decline of MMSE in SCTD than in euthyroidism.

Abbreviations: CI = confidence interval; ES = effect size, defined as annualized mean change in MMSE; MMSE = Mini-Mental State Examination; N = number; SCTD = subclinical thyroid dysfunction; TSH = thyrotropin.

† This study35 was excluded in the sensitivity analysis, because it might have included subclinical hyperthyroid participants in the control group.
Figure 1: Study Evaluation for Inclusion in the Meta-Analysis, adapted from PRISMA Statement Flow Diagram

Records identified through database searching* (n=1505)

Additional record identified through other sources† (n=0)

Records after duplicates removed (n=1471)

1471 records screened

1435 records excluded based on title and abstract (not a prospective study on the association between thyroid hormones and cognition)

36 full-text articles assessed for eligibility

25 of full-text articles excluded:
- No full text available because review, meeting abstract, poster or editorial (n=5)
- No thyroid function test (n=8)
- Reported the same study already selected without additional data to extract (n=1)
- No prospective study (n=11)

11 studies included in qualitative synthesis

10 studies included in quantitative synthesis (meta-analysis)

* Until November 11, 2014 for MEDLINE and November 14, 2014 for EMBASE.
† From key articles in the field and contact with authors
**Supplement Figure 1: Association between Subclinical Thyroid Dysfunction and Annualized Mean Change in Mini-Mental State Examination (MMSE)**

**Study**

<table>
<thead>
<tr>
<th>Euthyroid</th>
<th>SHyper</th>
<th>MD (95% CI)</th>
<th>Weight %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Ageing Study</td>
<td>823</td>
<td>7</td>
<td>0.67 (-0.01, 1.36)</td>
</tr>
<tr>
<td>Leiden 85-Plus Study</td>
<td>265</td>
<td>10</td>
<td>-0.27 (-1.42, 0.88)</td>
</tr>
<tr>
<td>OCTABAIX Study</td>
<td>186</td>
<td>5</td>
<td>-0.31 (-2.07, 1.44)</td>
</tr>
<tr>
<td>The PROSPER Study</td>
<td>4928</td>
<td>65</td>
<td>0.03 (-0.09, 0.15)</td>
</tr>
<tr>
<td>The Rotterdam Study</td>
<td>1520</td>
<td>86</td>
<td>-0.09 (-0.26, 0.09)</td>
</tr>
<tr>
<td>Overall (I-squared = 23.5%, p = 0.265)</td>
<td></td>
<td></td>
<td>0.01 (-0.14, 0.15)</td>
</tr>
</tbody>
</table>

**Legend, Supplement Figure 1:** MD = mean difference in change from baseline for MMSE score. MD>0 indicates higher decline of MMSE in SCTD than in euthyroidism. Abbreviations: 95% CI = 95% confidence interval; MD = annualized differences in mean change from baseline in MMSE; Euthyroid = euthyroidism; MMSE = Mini-Mental State Examination; n = number of patients with dementia per group; N = total number of patients per group; SHyper = subclinical hyperthyroidism; SHypo = subclinical hypothyroidism

**NOTE:** Weights are from random effects analysis
Figure 2: Association between Subclinical Thyroid Dysfunction and Dementia
SHyper and Dementia

<table>
<thead>
<tr>
<th>Study</th>
<th>Euthyroid n/N</th>
<th>SHyper n/N</th>
<th>Risk Ratio (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honolulu-Asia Aging Study</td>
<td>76/527</td>
<td>5/22</td>
<td>1.58 (0.71, 3.50)</td>
<td>35.14</td>
</tr>
<tr>
<td>OCTABAIX Study</td>
<td>39/164</td>
<td>1/4</td>
<td>1.05 (0.19, 5.87)</td>
<td>7.57</td>
</tr>
<tr>
<td>Rotterdam Scan Study</td>
<td>47/916</td>
<td>7/79</td>
<td>1.73 (0.81, 3.70)</td>
<td>38.74</td>
</tr>
<tr>
<td>The Health in Men Study</td>
<td>132/3018</td>
<td>0/19</td>
<td>0.57 (0.04, 8.84)</td>
<td>2.98</td>
</tr>
<tr>
<td>The Rotterdam Study</td>
<td>19/1568</td>
<td>3/91</td>
<td>2.72 (0.82, 9.03)</td>
<td>15.57</td>
</tr>
<tr>
<td>Overall (I-squared = 0.0%, p = 0.521)</td>
<td></td>
<td></td>
<td>1.67 (1.04, 2.69)</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Higher risk of dementia with Euthyroid Higher risk of dementia with SHyper

SHypo and Dementia

<table>
<thead>
<tr>
<th>Study</th>
<th>Euthyroid n/N</th>
<th>SHypo n/N</th>
<th>Risk Ratio (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conselloe Study of Brain Aging</td>
<td>62/524</td>
<td>18/120</td>
<td>1.27 (0.78, 2.06)</td>
<td>39.81</td>
</tr>
<tr>
<td>Honolulu-Asia Aging Study</td>
<td>76/527</td>
<td>2/19</td>
<td>0.73 (0.19, 2.75)</td>
<td>5.33</td>
</tr>
<tr>
<td>OCTABAIX Study</td>
<td>39/164</td>
<td>5/13</td>
<td>1.62 (0.77, 3.39)</td>
<td>17.15</td>
</tr>
<tr>
<td>Rotterdam Scan Study</td>
<td>47/918</td>
<td>2/17</td>
<td>2.30 (0.61, 8.79)</td>
<td>5.30</td>
</tr>
<tr>
<td>The Health in Men Study</td>
<td>132/3018</td>
<td>13/564</td>
<td>0.82 (0.47, 1.43)</td>
<td>30.06</td>
</tr>
<tr>
<td>The Rotterdam Study</td>
<td>19/1568</td>
<td>1/149</td>
<td>0.55 (0.07, 4.11)</td>
<td>2.34</td>
</tr>
<tr>
<td>Overall (I-squared = 0.0%, p = 0.489)</td>
<td></td>
<td></td>
<td>1.14 (0.84, 1.55)</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Higher risk of dementia with Euthyroid Higher risk of dementia with SHypo

Legend, Figure 2: RR > 1 indicates higher risk of an event in SCTD than in euthyroidism. Abbreviations: 95% CI = 95% confidence interval; Euthyroid = euthyroidism; n = number of patients with dementia per group; N = total number of patients per group; SHyper = subclinical hyperthyroidism; SHypo = subclinical hypothyroidism

NOTE: Weights are from random effects analysis
Supplement Figure 2: Funnel Plots and Egger’s Tests

**SHyper and Dementia**

- Funnel plot with p-value 95% confidence limits
- Egger’s test P = 0.564

**SHypo and Dementia**

- Funnel plot with p-value 95% confidence limits
- Egger’s test P = 0.75

**SHyper and MMSE**

- Funnel plot with p-value 95% confidence limits
- Egger’s test P = 0.14

**SHypo and MMSE**

- Funnel plot with p-value 95% confidence limits
- Egger’s test P = 0.14