Effect of Thyroid Hormone Therapy on Fatigability in Older Adults with Subclinical Hypothyroidism

A Nested Study within a Randomized Placebo-Controlled Trial

Stuber, Mirah J.; Moutzouri, Elisavet; Feller, Martin; Del Giovane, Cinzia; Bauer, Douglas C.; Blum, Manuel R.; Collet, Tinh Hai; Gussekloo, Jacobijn; Mooijaart, Simon P.; McCarthy, Vera J.C.; Aujesky, Drahomir; Westendorp, Rudi; Stott, David J.; Glynn, Nancy W.; Kearney, Patricia M.; Rodondi, Nicolas; Newman, Anne

Published in:
Journals of Gerontology - Series A Biological Sciences and Medical Sciences

DOI:
10.1093/gerona/glaa123

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Research Article

Effect of Thyroid Hormone Therapy on Fatigability in Older Adults With Subclinical Hypothyroidism: A Nested Study Within a Randomized Placebo-Controlled Trial

Mirah J. Stuber, MD,1 Elisavet Moutzouri, MD, PhD,1,2,* Martin Feller, MD, MSc,1,2 Cinzia Del Giovane, PhD,1 Douglas C. Bauer, MD,3 Manuel R. Blum, MD,1,2,4 Tinh-Hai Collet, MD,5 Jacobijn Gussekloo, MD, PhD,6,7 Simon P. Mooijaart, MD, PhD,6 Vera J. C. McCarthy, PhD,8 Drahomir Aujesky, MD, MSc,2 Rudi Westendorp, MD, PhD, MPH,9 David J. Stott, MD,10 Nancy W. Glynn, PhD,11 Patricia M. Kearney, MD, PhD,12 and Nicolas Rodondi, MD, MAS1,2

1Institute of Primary Health Care (BIHAM), University of Bern, Switzerland. 2Department of General Internal Medicine, Inselspital, Bern University Hospital, Switzerland. 3Departments of Medicine and Epidemiology and Biostatistics, University of California, San Francisco. 4Department of Health Research and Policy, Division of Epidemiology, Stanford University School of Medicine, California. 5Service of Endocrinology, Diabetes, Nutrition and Therapeutic Education, Geneva University Hospitals, Switzerland. 6Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Center, the Netherlands. 7Department of Public Health and Primary Care, Leiden University Center, the Netherlands. 8School of Nursing and Midwifery, University College Cork, Ireland. 9Department of Public Health and Center for Healthy Aging, University of Copenhagen, Denmark. 10Institute of Cardiovascular Medicine, University of Glasgow, Scotland. 11Department of Epidemiology, Center for Aging and Population Health, Graduate School of Public Health, University of Pittsburgh, Pennsylvania. 12School of Public Health, University College Cork, Ireland.

*Address correspondence to: Elisavet Moutzouri, MD, PhD, Department of General Internal Medicine, Inselspital, Bern University Hospital, 3010 Bern, Switzerland. E-mail: elisavet.moutzouri@extern.insel.ch

Trial Registration: This nested study is registered on ClinicalTrials.gov (number NCT02500342) with the title “Effect of Thyroid Hormone Replacement on Fatigability in Untreated Older Adults With Subclinical Hypothyroidism (TRUST FATIGUE).” (https://clinicaltrials.gov/ct2/show/NCT02500342?term=fatigue&cond=Subclinical+hypothyroidism&rank=1)

Received: January 9, 2020; Editorial Decision Date: May 6, 2020

Decision Editor: Anne Newman, MD, MPH

Abstract

Background: Fatigue often triggers screening for and treatment of subclinical hypothyroidism. However, data on the impact of levothyroxine on fatigue is limited and previous studies might not have captured all aspects of fatigue.

Method: This study is nested within the randomized, placebo-controlled, multicenter TRUST trial, including community-dwelling participants aged ≥65 and older, with persistent subclinical hypothyroidism (TSH 4.60–19.99 mIU/L, normal free thyroxine levels) from Switzerland and Ireland. Interventions consisted of daily levothyroxine starting with 50 μg (25 μg if weight <50 kg or known coronary heart diseases) together with dose adjustments to achieve a normal TSH and mock titration in the placebo group. Main outcome was the change in physical and mental fatigability using the Pittsburgh Fatigability Scale over 1 year, assessed through multivariable linear regression with adjustment for country, sex, and levothyroxine starting dose.

Results: Among 230 participants, the mean ± standard deviation (SD) TSH was 6.2 ± 1.9 mIU/L at baseline and decreased to 3.1 ± 1.3 with LT4 (n = 119) versus 5.3 ± 2.3 with placebo (n = 111, p < .001) after 1 year. After adjustment we found no between-group difference at 1 year on perceived physical (0.2; 95% CI −1.8 to 2.1; p = .88), or mental fatigability (−1.0; 95% CI −2.8 to 0.8; p = .26). In participants with higher...
In brief, participants were assigned to treatment or placebo group in a randomized, placebo-controlled parallel-group trial that included As previously published (14), the TRUST trial was a multicenter trial with the power to detect clinically meaningful benefits from LT4 replacement in older adults with Subclinical Hypothyroidism (SHypo) (6). Its findings indicated no effect of LT4 replacement on global fatigue, as assessed by the Thyroid Related Quality-of-Life Patient-Reported Outcome measure (ThyPRO) (6,9). However, assessing global fatigue without taking activity into account might result in misleading conclusions since older adults, in particular, tend to adjust their activity levels in a way that perceived fatigue remains tolerable (10). Assessing fatigability, a concept which anchors tiredness to a set of activities of defined intensity and duration, has proven to be a more sensitive measure for detecting fatigue rather than using global fatigue scores (10). In view of the remaining controversy on the benefit of LT4 replacement in SHypo regarding symptoms (11,12), this pre-registered study nested within the TRUST trial aimed at extending the TRUST trial findings with this novel concept of fatigability using the Pittsburgh Fatigability Scale (PFS) (10,13).

Method

This study is registered on ClinicalTrials.gov (number NCT02500342) as a nested study from the TRUST trial performed in Scotland, Ireland, the Netherlands, and Switzerland (6). This nested study was designed after the recruitment of the main trial started since the scale used for outcome measure was published afterward (January 2015) (13).

This nested study included the TRUST participants included at the two Swiss study centers (Inselspital, Bern University Hospital, and Centre Hospitalier Universitaire Vaudois, Lausanne University Hospital) and the Irish center (University College Cork, National University of Ireland, Ireland) (14). The trial was approved by the relevant ethics committees, and participants provided written informed consent.

The TRUST Trial

As previously published (14), the TRUST trial was a multicenter randomized, placebo-controlled parallel-group trial that included community-dwelling adults aged ≥ 65 years with untreated SHypo. In brief, participants were assigned to treatment or placebo group through permuted block randomization in a 1:1 ratio with stratification by sex, country, and starting dose. SHypo was defined as the presence of persistently elevated TSH levels (4.6–19.9 mIU/L) with fT4 within the assay reference range. TSH levels counted as persistently elevated if increased at a minimum of two occasions at least 3 months apart, over a maximum period of 3 years. The intervention consisted of daily LT4 doses, starting with 50 μg (or 25 μg in participants with bodyweight < 50 kg or with known coronary heart diseases, that is, symptoms of angina pectoris, or previous myocardial infarction), followed by dose adjustments to achieve a TSH level within the reference range (6). Blinding of participants was ensured by matching LT4 and placebo tablets as well as mock titrations in the placebo group, and blinding of clinicians and study centers through remote laboratory analysis of blood samples for TSH and dose titration as per computer algorithm.

The study was funded by the European Union FP7 and by the Swiss National Science Foundation (SNSF) for this nested study. LT4 as well as matching placebo were supplied free of charge by Merck KGaA (Darmstadt, Germany). The funders, 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 had no influence on the main and nested studies’ designs, analyses or reporting.

The primary outcomes of this study were defined prior to data collection as the mean follow-up scores for mental and physical fatigability, measured by the PFS (13), after 1 year of follow-up with adjustment for baseline scores. Exploratory secondary outcome measures (that were not pre-specified) were the number of participants reporting higher physical (≥15 points) or mental fatigability (≥13 points) after 1 year of treatment with adjustment for baseline numbers, based on previously established cut points (15).

Pittsburgh Fatigability Scale

The PFS is a reliable, sensitive scale, and the first-validated self-report tool to measure perceived fatigability in older adults (13). It is a 10-item questionnaire asking the participants to estimate the physical and mental fatigability they expect or image they would have after performance of different activities of specific intensity and duration. The two independent subscores (physical and mental) range from 0 to 50, with higher scores indicating greater fatigability. The cut points for higher fatigability are set at ≥15 points for the physical PFS subscore and at ≥13 points for the mental PFS subscore (15,16). The minimal clinically important difference is estimated at 2 to 3 points (13,17). The PFS has been translated into German and French, including two forward translations, synthesis of translations, and two back-translations.

Statistical Analysis

We performed the main analysis in a modified intention-to-treat population, which included all participants with outcome of interest, and not more than three missing answers in the PFS. For up to three missing items, we imputed data using an imputation method.
published by the scale's developers (18). Imputed data were based on the mean value of an individual's complete responses with adjustments for varying intensity levels of the different activities, sex differences as well as differences in reported fatigability levels between participants who have versus those who have not done each specific activity (18). Multivariable linear regression was conducted for the physical and mental PFS scores at 1 year with adjustment for the stratification variables used for the randomization (country, sex, starting dose of LT4), as well as for corresponding PFS baseline values. Goodness of model fit was assessed through distribution plots of residuals and scatter plots of standardized residuals against predicted values.

The between-group difference in number of participants with higher fatigability at follow-up was assessed using multivariable logistic regression adjusting for the same variables as described above.

We conducted several sensitivity analyses: (a) in a population limited to participants with complete outcome data (participants who have answered to all the questions from baseline and follow-up questionnaire), (b) using the inverse probability weighting method in order to adjust for a possible attrition bias due to loss to follow-up (12 months missing primary outcomes) (19); the covariates included in the logistic regression model to estimate inverse probability weightings were chosen based on clinical judgment: age, body-mass index, country, sex and the number of comorbidities, (c) in participants who, after 1 year, adhered to the trial regimen in accordance with the protocol (per-protocol population), (d) only in participants who reported higher fatigability at baseline, (e) two exploratory sensitivity analyses, one excluding participants with diabetes and one adjusting for the presence of diabetes, to adjust for the baseline imbalance of diabetes prevalence between the groups, (f) an analysis of square-root transformed data to account for skewness in baseline scores, (g) in participants with TSH levels in the upper quartile.

Power calculation (Analysis of covariance [ANCOVA] method) for 110 participants per allocation group assuming a standard deviation (SD) of 9 in baseline fatigability and a baseline to follow-up correlation of 0.7 resulted in a 93.3% power to detect a difference of three points (minimal clinically important difference) in the mean follow-up fatigability scores at a two-sided alpha level of 0.05 (20).

Results

Trial Population

The study flowchart is shown in Figure 1. Participants were enrolled from January 2014 to December 2015. The last participant completed the study follow-up in November 2016. From 1,273 participants screened, most reverted to normal TSH before randomization and 276 participants were randomized (n = 142 allocated to LT4) (6). Fifty-six participants were randomized before the present nested trial began. Overall, 38 participants had missing 12-month PFS. Thirteen withdrew from the study, two participants died, one participant withdrew due to administrative reasons, (n = 276). Characteristics were well balanced between the two groups except for diabetes mellitus (p = .12), which was more prevalent in the LT4 group. In the LT4 group, 17 (12.0%) participants started with a lower dose of 25 µg. The mean LT4 dose in the LT4 group was 47 µg (ranging from 25 µg to 50 µg) at baseline and 48 µg (range 0 µg to 100 µg) at 1-year follow-up. The mean ± SD TSH was 6.2 ± 1.9 mIU/L at baseline and decreased to 3.1 ± 1.3 in the LT4 group versus 5.3 ± 2.3 in the placebo group (p < .001) after 1 year (Table 2).

Fatigability

Baseline perceived physical and mental fatigability between the LT4 and the placebo group differed by 3.6 points (p = .003) and 2.3 points (p = .023), respectively (Table 2). Baseline fatigability imbalance was not associated with baseline diabetes imbalance. The mean follow-up PFS physical score was 14.8 ± 9.6 in the LT4 group and 12.4 ± 9.3 in the placebo group with an adjusted between-group difference of 0.2 (95% CI −1.8 to 2.1, p = .88). The mean follow-up PFS mental score was 6.0 ± 7.8 in the LT4 and 6.0 ± 8.0 in the placebo group, respectively, with an adjusted between-group

Figure 1. Randomization and follow-up. The Pittsburgh Fatigability Scale (PFS) has two separate subscores for mental and physical fatigability, each ranging from 0 to 50 with higher scores indicating greater fatigability.

Table 1 summarizes the baseline characteristics from all randomized participants (n = 276). Characteristics were well balanced between the two groups except for diabetes mellitus (p = .12), which was more prevalent in the LT4 group. In the LT4 group, 17 (12.0%) participants started with a lower dose of 25 µg. The mean LT4 dose in the LT4 group was 47 µg (ranging from 25 µg to 50 µg) at baseline and 48 µg (range 0 µg to 100 µg) at 1-year follow-up. The mean ± SD TSH was 6.2 ± 1.9 mIU/L at baseline and decreased to 3.1 ± 1.3 in the LT4 group versus 5.3 ± 2.3 in the placebo group (p < .001) after 1 year (Table 2).

Fatigability

Baseline perceived physical and mental fatigability between the LT4 and the placebo group differed by 3.6 points (p = .003) and 2.3 points (p = .023), respectively (Table 2). Baseline fatigability imbalance was not associated with baseline diabetes imbalance.

The mean follow-up PFS physical score was 14.8 ± 9.6 in the LT4 group and 12.4 ± 9.3 in the placebo group with an adjusted between-group difference of 0.2 (95% CI −1.8 to 2.1, p = .88). The mean follow-up PFS mental score was 6.0 ± 7.8 in the LT4 and 6.0 ± 8.0 in the placebo group, respectively, with an adjusted between-group
Table 1. Baseline Characteristics of all Participants Randomized (n = 276)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Placebo (N = 134)</th>
<th>Levothyroxine (N = 142)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>73.5 ± 6.3</td>
<td>73.9 ± 5.1</td>
</tr>
<tr>
<td>Range</td>
<td>65.1–92.9</td>
<td>63.3–88.6</td>
</tr>
<tr>
<td>Female sex—no. (%)</td>
<td>60 (44.8)</td>
<td>62 (43.4)</td>
</tr>
<tr>
<td>White race—no. (%)</td>
<td>132 (99)</td>
<td>141 (99)</td>
</tr>
<tr>
<td>Standard housing—no. (%)</td>
<td>130 (97)</td>
<td>140/142 (99)</td>
</tr>
<tr>
<td>Weight &lt;50 kg—no. (%)</td>
<td>3 (2)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Body-mass index, kg/m², mean ± SD</td>
<td>27.5 ± 4.6</td>
<td>28.0 ± 4.6</td>
</tr>
<tr>
<td>Previous medical conditions and clinical descriptors—no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>13 (9.7)</td>
<td>14 (9.9)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>16 (11.9)</td>
<td>15 (10.7)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>8 (6.0)</td>
<td>4 (2.8)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>14 (10.4)</td>
<td>24 (16.9)</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>17 (12.7)</td>
<td>14 (10.1)</td>
</tr>
<tr>
<td>Current smoking</td>
<td>12 (9.0)</td>
<td>8 (5.6)</td>
</tr>
<tr>
<td>No. of concomitant medicines—median (IQRg)</td>
<td>4 (2–6)</td>
<td>5 (3–5.5)</td>
</tr>
<tr>
<td>Laboratory results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyrotropin—mIU/L, mean ± SD</td>
<td>6.31 ± 2.00</td>
<td>6.12 ± 1.90</td>
</tr>
<tr>
<td>Range</td>
<td>5.70 (5.02–6.79)</td>
<td>5.60 (5.04–6.80)</td>
</tr>
<tr>
<td>Free thyroxine—pmol/L, mean ± SD</td>
<td>13.5 ± 1.9</td>
<td>13.7 ± 2.0</td>
</tr>
<tr>
<td>Levothyroxine dose of 25 μg at randomization (%)</td>
<td>14 (10.5)</td>
<td>17 (11.9)</td>
</tr>
</tbody>
</table>

Notes: *Race was reported by the participants.
1Standard housing was defined as non-sheltered community accommodation, whereas sheltered housing means a purposely-built grouped housing for older persons.
2Calculated as weight (kilograms) divided by squared height (meters).
3Defined as history of angina pectoris or previous myocardial infarction.
4Defined as current smoking at the time of baseline exam.
5IPQR denotes interquartile range.

Discussion
This nested study within a randomized, placebo-controlled, parallel-group, double-blind multicenter trial in 230 participants aged ≥ 65 years with SHypo did not show a significant benefit of LT4 replacement on physical or mental fatigability after 1 year of treatment, using a more sensitive scale (10) and assessing both, physical and mental fatigability, instead of global fatigue as measured in the main TRUST trial (6).

Previous data on the impact of LT4 on fatigue are limited and conflicting. Even though fatigue is a very common symptom in SHypo and often leads to LT4 prescription (4,5), a recent systematic review showed that large randomized controlled trials investigating the benefit of LT4 replacement on fatigue in SHypo are lacking (8). The only trial, which provided quantitative data for the outcome of fatigue/tiredness, was the TRUST trial. Using the ThyPRO questionnaire to assess tiredness (seven items), the TRUST trial did not find a benefit of LT4 replacement on fatigue in SHypo (7). In their randomized, crossover trial in 100 participants with SHypo (mean age 53.8 ± 12.0 years), the authors reported that the proportion of participants with tiredness decreased from 89% to 78% under LT4 replacement, but did not provide measures of precision (such as confidence intervals) and how tiredness was assessed. However, their study differs from ours with respect to the lower mean age of participants and the higher daily LT4 dose (100 μg vs a mean dose of 47 μg at baseline in our study). It would be interesting to compare the fatigability levels of our study population to the fatigability level of a comparable population of older adults, but there are very limited population-based data currently available for comparison because our study was early in using the PFS.
At 1 year established.

Participants with higher physical fatigability (≥15 points) and 18% with higher physical fatigability (≥15 points) and 18% with higher mental fatigability (≥13 points) for baseline scores, this imbalance at baseline should not be expected to impact the results. Fourth, only participants ≥ 65 years were included. Fifth, participants with TSH levels > 10 mIU/L accounted for 4% of the study population. Thus, the findings may not be generalizable to individuals with TSH levels > 10 mIU/L. Finally, the mean TSH level in the LT4 replacement group was 2.95 mIU/L after 1 year of treatment. We could not exclude that fatigability would have decreased more under a more aggressive LT4 regimen, but potentially at the cost of harms from overtreatment, such as atrial fibrillation or fractures (25–28).

### Conclusion

Over a 1-year follow-up, normalization of TSH levels through LT4 replacement in people aged ≥ 65 years did not show a benefit on physical and mental fatigability compared to placebo. The same finding was shown for participants with higher fatigability at baseline. In line with the findings from the TRUST trial on global fatigue (6), which was the largest randomized controlled trial on the treatment of SHypo, our results do not provide evidence in favor of LT4 replacement to reduce fatigability or fatigue in older adults with SHypo.

**Table 2. Thyroid Function, Physical, and Mental Fatigability at Baseline and After 1 Year**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline</th>
<th>At 1 year</th>
<th>Adjusted Between-Group Difference (95% CI)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Levothyroxine (N = 119)</td>
<td>Placebo (N = 111)</td>
<td>Levothyroxine (N = 119)</td>
<td>Placebo (N = 111)</td>
</tr>
<tr>
<td>Thyrotropin—mIU/L</td>
<td>6.08 ± 1.80</td>
<td>5.29 ± 2.01</td>
<td>6.08 ± 1.80</td>
<td>5.50 ± 2.34</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>3.56 (5.04 to 6.62)</td>
<td>5.72 (5.06 to 6.76)</td>
<td>5.75 (2.26 to 3.77)</td>
<td>4.8 (3.62 to 6.57)</td>
</tr>
<tr>
<td>Primary Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS physical score (SD)a</td>
<td>14.7 ± 9.3</td>
<td>11.1 ± 9.1</td>
<td>14.8 ± 9.6</td>
<td>12.4 ± 9.3</td>
</tr>
<tr>
<td>PFS mental score (SD)b</td>
<td>7.4 ± 8.0</td>
<td>5.1 ± 6.9</td>
<td>6.0 ± 7.8</td>
<td>6.0 ± 8.0</td>
</tr>
<tr>
<td>Secondary Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants with higher</td>
<td>55 (46.2)</td>
<td>33 (29.7)</td>
<td>52 (43.7)</td>
<td>41 (36.9)</td>
</tr>
<tr>
<td>physical fatigability (%)c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants with higher</td>
<td>27 (22.7)</td>
<td>14 (12.6)</td>
<td>19 (16.0)</td>
<td>20 (18.0)</td>
</tr>
<tr>
<td>mental fatigability (%)d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aThe Pittsburgh Fatigability Scale (PFS) physical and mental subscores range from 0 to 50 with higher scores indicating greater fatigability. Crude means are reported.
bThe p-value generated through multiple linear regression model for the follow-up scores adjusted for PFS baseline scores, sex, country and starting levothyroxine dose.
cThe cut points for higher fatigability are set at ≥15 points for the physical PFS subscore and at ≥13 points for the mental PFS subscore (13,15), as previously established.
dThe p-value generated through multiple logistic regression model for the number of participants with higher fatigability at follow-up adjusted for number of participants with higher fatigability at baseline, sex, country and starting levothyroxine dose.

Strengths and novelty of our study were the use of a validated scale that measures the novel construct of fatigability and the consideration of both the physical and mental dimensions of fatigability (10,13). Prevalence of global fatigue varies widely depending on measurement tools, population characteristics, and cutoff points chosen to distinguish between fatigued and non-fatigued participants (21,22). It is, therefore, desirable to bring more objectivity into tiredness evaluation. Fatigability is a concept that classifies fatigue in relation to a defined activity of a specific duration and intensity (10). This conceptualization might lead to a less biased and more objective measure of fatigue than a global fatigue score (13). The classification of fatigue in relation to specific activities is especially helpful in older adults, as these individuals tend to adjust their activity level (eg, by slowing down or shortening the task duration) in order to maintain their perceived fatigue in a tolerable range (13). It is thus possible that two individuals differing in their daily activities (with one having a very active lifestyle and the other a sedentary way of living) report the same tiredness level for the last month. The construct of fatigability is adapted to take the bias of “self-pacing” into account (10,13). The PFS is furthermore the only scale taking into account the multidimensionality of fatigability by distinguishing between mental and physical fatigability (23). Mental fatigability—or self-reported cognitive tiredness related to specific activities—is rarely recognized by the medical community (24). Thus, this study is the first to test the effect of LT4 treatment on both physical and mental fatigability in older subjects with SHypo.

This study has limitations. First, participants were included independent of their baseline fatigability, leading to a proportion of 38% with higher physical fatigability (≥15 points) and 18% with higher mental fatigability (≥13 points). It is possible that findings could differ in a population of (highly) fatigued participants. A sensitivity analysis, including only these participants with higher fatigability, did not reveal a reduction in fatigability level after treatment. Second, 14% (n = 38) of participants did not have 12-month PFS measurements. However, these participants had similar characteristics to the participants who did have 12-month PFS and the proportion was balanced between the groups. In particular, these subjects had baseline scores in physical and mental fatigability that did not differ from the participants included in the main analysis. Furthermore, results were robust in a sensitivity analysis using inverse probability weighting to account for possible bias due to missing outcomes (19).

This study has limitations. First, participants were included independent of their baseline fatigability, leading to a proportion of 38% with higher physical fatigability (≥15 points) and 18% with higher mental fatigability (≥13 points). It is possible that findings could differ in a population of (highly) fatigued participants. A sensitivity analysis, including only these participants with higher fatigability, did not reveal a reduction in fatigability level after treatment. Second, 14% (n = 38) of participants did not have 12-month PFS measurements. However, these participants had similar characteristics to the participants who did have 12-month PFS and the proportion was balanced between the groups. In particular, these subjects had baseline scores in physical and mental fatigability that did not differ from the participants included in the main analysis. Furthermore, results were robust in a sensitivity analysis using inverse probability weighting to account for possible bias due to missing outcomes (19).

Third, the scores for physical and mental fatigability were unbalanced between the two groups at baseline. As analyses were adjusted for baseline scores, this imbalance at baseline should not be expected to impact the results. Fourth, only participants ≥ 65 years were included. Fifth, participants with TSH levels > 10 mIU/L accounted for 4% of the study population. Thus, the findings may not be generalizable to individuals with TSH levels > 10 mIU/L. Finally, the mean TSH level in the LT4 replacement group was 2.95 mIU/L after 1 year of treatment. We could not exclude that fatigability would have decreased more under a more aggressive LT4 regimen, but potentially at the cost of harms from overtreatment, such as atrial fibrillation or fractures (25–28).
Supplementary Material

Supplementary data are available at The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences online.

Funding

This nested study on fatigability within the TRUST trial was funded by a grant from the Swiss National Science Foundation (SNSF 320030-172676 to N.R.). The TRUST Thyroid trial was supported by research grant (278148) from the European Union FP7-HEALTH-2011 program and by grants from the Swiss National Science Foundation (SNSF 320030-150025 and 320030-172676 to N.R.) and the Swiss Heart Foundation and Velux Stiftung (grant 974a to N.R.). M.R.B’s work was supported by a grant from the Swiss National Science Foundation (P2BEP3_175289). T.H.C’s research is supported by grants from the Swiss National Science Foundation (PZ00P3-167826). The study medication (levothyroxine and matching placebo) was supplied free of charge by Merck KGaA, Darmstadt, Germany. Merck played no role in the design, analysis, or reporting of the trial. The main TRUST sponsor (NHG Greater Glasgow and Clyde Health Board) contributed to the writing of the protocol.

Conflict of Interest

None reported.

References