Serum Thyroid-Stimulating Hormone Concentration and Morbidity from Cardiovascular Disease and Fractures in Patients on Long-Term Thyroxine Therapy

Robert W. Flynn, Sandra R. Bonellie, Roland T. Jung, Thomas M. MacDonald, Andrew D. Morris, and Graham P. Leese

Ninewells Hospital and Medical School (R.W.F., R.T.J., T.M.M., A.D.M., G.P.L.), University of Dundee, Dundee DD1 9SY, United Kingdom; and School of Accounting, Economics and Statistics (S.R.B.), Edinburgh Napier University, Edinburgh EH14 1DJ, United Kingdom

Context: For patients on T4 replacement, the dose is guided by serum TSH concentrations, but some patients request higher doses due to adverse symptoms.

Objective: The aim of the study was to determine the safety of patients having a low but not suppressed serum TSH when receiving long-term T4 replacement.

Design: We conducted an observational cohort study, using data linkage from regional datasets between 1993 and 2001.

Setting: A population-based study of all patients in Tayside, Scotland, was performed.

Patients: All patients taking T4 replacement therapy (n = 17,684) were included.

Main Outcome Measures: Fatal and nonfatal endpoints were considered for cardiovascular disease, dysrhythmias, and fractures. Patients were categorized as having a suppressed TSH (≤0.03 mU/liter), low TSH (0.04–0.4 mU/liter), normal TSH (0.4–4.0 mU/liter), or raised TSH (>4.0 mU/liter).

Results: Cardiovascular disease, dysrhythmias, and fractures were increased in patients with a high TSH: adjusted hazards ratio, 1.95 (1.73–2.21), 1.80 (1.33–2.44), and 1.83 (1.41–2.37), respectively; and patients with a suppressed TSH: 1.37 (1.17–1.60), 1.6 (1.10–2.33), and 2.02 (1.55–2.62), respectively, when compared to patients with a TSH in the laboratory reference range. Patients with a low TSH did not have an increased risk of any of these outcomes [hazards ratio: 1.1 (0.99–1.123), 1.13 (0.88–1.47), and 1.13 (0.92–1.39), respectively].

Conclusions: Patients with a high or suppressed TSH had an increased risk of cardiovascular disease, dysrhythmias, and fractures, but patients with a low but unsuppressed TSH did not. It may be safe for patients treated with T4 to have a low but not suppressed serum TSH concentration. (J Clin Endocrinol Metab 95: 186–193, 2010)

Thyroid dysfunction is one of the most common chronic diseases, with the prevalence of primary hypothyroidism exhibiting a wide geographical variation but with up to 4% of the population affected in some areas (1–3). The prevalence of thyroid disease is increasing by 6% per annum in Tayside, Scotland (2), and T4 was the sixth most commonly prescribed medicine across Scotland in 2007–2008 (4). Despite this, controversy still remains on the optimal T4 replacement. The most widely accepted current practice for treating patients with an underactive thyroid gland is to treat to achieve a biochemical objective of a “normal” TSH concentration (5–8). However, a minority of patients seem to prefer being prescribed doses of T4 that result in a low or suppressed TSH (9, 10), and in

Abbreviations: CHNo, Community Health Number(s); CI, confidence interval; HR, hazard ratio; SMR, Scottish Morbidity Records.
practice many patients taking long-term T\textsubscript{4} do have a low TSH, despite guidelines (11, 12).

Guidelines are based on studies indicating that individuals with a suppressed TSH have an increased risk of atrial fibrillation (13–15), increased mortality (16, 17), and possibly an increased fracture risk (18, 19), although this remains controversial (20). Other studies show an increased risk in subclinical hypothyroidism (21, 22). However, many of these studies have potential confounders, not the least of which is that many included patients with endogenous subclinical disease who were not on T\textsubscript{4} replacement.

In Tayside, Scotland, the Community Health Number (CHNo) has been used since 1979 as the patient identifier for all health care activities, with CHNo being assigned to every person registered with a family doctor. The multiple sources of electronic patient-specific data such as Scottish Morbidity Records (SMR) allow records to be assimilated through CHNo to create a highly detailed clinical record for each of the 400,000 residents of Tayside. Detailed information on all dispensed prescriptions for all Tayside residents is another data source, which has been active since 1993 (23). These resources have previously been combined in the creation of the Thyroid Epidemiology Audit and Research Study (TEARS) (1) and other population-based datasets.

This population-based study was designed to test the hypothesis that TSH concentration was associated with cardiovascular disease, dysrhythmias, and fractures in patients receiving long-term replacement T\textsubscript{4} therapy as well as which concentrations of TSH may be safe.

Materials and Methods

This study was conducted using data from the TEARS dataset and the anonymized record linkage methodology developed at the University of Dundee in Tayside, Scotland (1, 23). The CHNo was used to link patient-level data from this dataset with outcome data from SMR01 records, electronic death certification data, and other data relating to potential confounding variables.

Study population

This study used a dynamic study population that consisted of all patients registered with Tayside general practitioners between January 1, 1993, and December 31, 2001. The study subjects were all of those in the population who received prescriptions for T\textsubscript{4} in that period. Only subjects over 18 yr of age were included. Patients who had their T\textsubscript{4} stopped during follow-up were excluded from the study.

Data sources

CHNo master patient index

This defined the study population from which subjects were identified, providing data on registered Tayside general practi-
ICD10 S32, S52, S72, M80). These codes relate to all fractures of lower back, forearm, hip, and “other” specifically coded osteoporotic fractures. We cannot be sure that they were necessarily low trauma fractures, however, but these fracture sites are classically viewed as sites related to osteoporotic fractures. Survival analysis was used to follow up patients until an event, until they were censored, or until the end of the study. The data were modeled using the Cox proportional-hazards model. The assumption of proportional hazards was assessed by plotting log-log plots for the baseline covariates and by modeling interaction terms—in each case, the proportional hazards were found to be supported. Two-way interaction terms between all covariates were modeled and were included in the final model if they were found to be significant.

Patients were followed up from the first TSH measurement after their third T₄ prescription. This was the definition of requiring ongoing T₄ therapy reflecting stabilized therapy, but also reducing the risk of immortal time bias. Subjects whose TSH was not recorded between their third T₄ script and an event were excluded from the analysis. Some data were missing for socioeconomic status. Because these represented less than 1% of the whole dataset, it was assumed that these were missing completely at random, and a complete case analysis was performed.

### TSH control measurement

Patients were categorized into one of four groups according to their mean serum TSH concentration over the study period. A time-weighted mean was calculated that took into account differing lengths of time between tests. This was done in the expectation that there would be clustering of abnormal results because an abnormal result will likely be followed by T₄ dose adjustment and early TSH rechecks. Without time-weighting, the importance of abnormal TSH results would be overemphasized. Using a time-weighted mean that gives a greater weight to those tests when the frequency of testing is low gives a better indication of TSH control over the whole study period, akin to area under the curve. Because TSH values are known to follow a log-normal distribution (29), this was expressed as a geometric mean (i.e. expressed as the back-transformed mean of the logarithmic values of TSH value). Each patient’s time-weighted geometric mean TSH was then classified as: “suppressed” (≤0.03 mU/liter), “low” (0.04–0.4 mU/liter), “normal” (0.4–4.0 mU/liter), and “high” (>4.0 mU/liter). Because TSH values are known to be affected during episodes of severe acute illness, TSH measurements made while hospitalized for circulatory diseases were excluded (30). Free T₄ concentrations were not available, and therefore were not included in the analysis.

As an exploratory analysis, we modeled TSH concentration as a fractional polynomial in a Cox model for both the cardiovascular and osteoporotic fracture endpoints. This approach allows for the flexible nonlinear modeling of continuous variables such as TSH concentration (31). We use the %mpg8 macro in SAS described by Sauerbrei et al. (32) using the closed test procedure (RA2) to achieve a best fit for the fractional polynomial terms. The data were modeled as for the primary analysis, using a log-transformed TSH concentration and the same covariates, but with the TSH categories replaced by log TSH. The resulting functions, describing the association between TSH concentration and outcome, are displayed as plots of the log relative hazard against the TSH concentration expressed. This log relative hazard describes the partial effect of TSH concentration and is scaled by a constant to allow comparison with the “step function” calculated in the primary analysis.

### Other confounders

Other covariates adjusted for were sex, etiology of T₄ replacement (i.e. treated hyperthyroidism, primary hypothyroidism), whether the patient had diabetes mellitus, socioeconomic deprivation as measured by the Carstairs index, and history of relevant morbidity. Some covariates could not be accommodated due to lack of data, and this included smoking where reliable data were not available. For the cardiovascular disease outcomes, we considered whether or not the subject had a history of previous circulatory disease (taken from any such admissions on the SMR01 records: ICD-9 codes 390–459; ICD-10 codes 100–199), and for the osteoporotic fracture outcome we considered history of osteoporotic fracture. For the dysrhythmias endpoint we considered history of dysrhythmias and history of other circulatory disease.

### Ethical approval

In accordance with standard operating procedure, approval for the study was granted by the Tayside Medical Ethics Committee, permission to access medical records was granted by the Tayside Caldicott Guardians, and all analyses were performed on anonymized datasets (33, 34).

### Results

There were a total of 17,684 T₄-treated patients with a total follow-up of 78,518 yr after the index serum TSH measurement. The median follow-up was 1,651 d (4.5 yr). Overall, 85.9% of the subjects were female with a mean age of 60.3 yr, whereas 14.1% of the subjects were male with a mean age of 61.8 yr.

In total, there were 2,144 episodes of cardiovascular disease, 367 of dysrhythmias, and 562 osteoporotic fractures. The baseline characteristics of the cohort by TSH category are shown in Table 1. Of all patients on T₄, 61.7% had a TSH in the laboratory reference range, 11.2% had a raised TSH, 6.1% had a fully suppressed TSH, and 21.1% had a TSH that was low but not suppressed (between 0.04 and 0.4 mU/liter). At baseline, patients with a raised TSH were more likely to be male, have diabetes, have primary hypothyroid, and have preexisting morbidity, compared with patients with a TSH in the reference range. Patients with a suppressed TSH were more likely to have primary hypothyroidism but were less likely to have diabetes, preexisting cardiovascular disease, or dysrhythmias than patients with a TSH in the reference range. Overall, the median rate of TSH tests per patient was 1.35 per year. Patients in the high TSH category had the highest rate, and those in the suppressed category had the lowest rate (Table 1).

Table 2 shows the number of cardiovascular admissions, dysrhythmia admissions, and osteoporotic fractures
with adjusted and unadjusted hazard ratios (HRs). For all endpoints, there was increased risk with increased age. Females were at increased risk of fracture, and males were at increased risk of cardiovascular endpoints and dysrhythmias. Compared with a normal TSH, suppressed serum TSH was associated with a modest increase in cardiovascular morbidity and mortality, whereas high TSH was associated with a near doubling in risk. There was no significant difference between low and normal TSH levels.

Lower socioeconomic status, as measured by the Carstairs’s score, was associated with an increased risk of cardiovascular morbidity and mortality [Carstairs categories 5, 6, and 7 vs. 1 and 2; HR, 1.36; 95% confidence interval (CI), 1.21–1.53], as was the presence of diabetes at baseline (HR, 1.69; 95% CI, 1.48–1.94). A similar situation existed for dysrhythmia admissions, although the risk associated with a suppressed TSH was nearly as large as with a high TSH concentration. Again, there was no significant increased risk associated with a low TSH. Figure 1 shows adjusted survival curves derived from the Cox model for each outcome stratified by TSH concentration, which show similar results for long-term follow-up.

Figure 2 shows the estimated functional relationship between TSH concentration and outcome. For both cardiovascular disease and osteoporotic fracture, the best

### TABLE 1. Baseline characteristics stratified by TSH control for primary outcome

<table>
<thead>
<tr>
<th>Covariate†</th>
<th>Suppressed TSH</th>
<th>Low TSH</th>
<th>Normal TSH</th>
<th>High TSH</th>
<th>P value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, n (%)</td>
<td>1,070 (6.1)</td>
<td>3,731 (21.1)</td>
<td>10,908 (61.7)</td>
<td>1,975 (11.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mean age (yr)</td>
<td>61.6</td>
<td>60.8</td>
<td>60.0</td>
<td>62.1</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female, n (%)</td>
<td>959 (89.6)</td>
<td>3,313 (88.8)</td>
<td>9,377 (86.0)</td>
<td>1,547 (78.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Diabetics, n (%)</td>
<td>43 (4.0)</td>
<td>185 (5.0)</td>
<td>663 (6.1)</td>
<td>156 (7.9)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Carstairs’s category, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 and 2</td>
<td>287 (26.8)</td>
<td>1,038 (27.8)</td>
<td>2,862 (26.2)</td>
<td>475 (24.1)</td>
<td></td>
</tr>
<tr>
<td>3 and 4</td>
<td>443 (41.4)</td>
<td>1,700 (45.6)</td>
<td>4,935 (45.2)</td>
<td>901 (45.6)</td>
<td>0.0013</td>
</tr>
<tr>
<td>5, 6, and 7</td>
<td>340 (31.8)</td>
<td>993 (26.6)</td>
<td>3,111 (28.5)</td>
<td>599 (30.3)</td>
<td></td>
</tr>
<tr>
<td>Previously hyperthyroid, n (%)</td>
<td>192 (17.9)</td>
<td>812 (21.8)</td>
<td>1,458 (13.4)</td>
<td>194 (9.8)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>History of CVD, n (%)</td>
<td>111 (10.4)</td>
<td>418 (11.2)</td>
<td>1,531 (14.0)</td>
<td>428 (21.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>History of dysrhythmia, n (%)</td>
<td>18 (1.8)</td>
<td>105 (2.8)</td>
<td>385 (3.5)</td>
<td>120 (6.3)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>History of fracture, n (%)</td>
<td>41 (4.0)</td>
<td>86 (2.3)</td>
<td>321 (2.9)</td>
<td>100 (5.2)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Frequency of TSH tests per year</td>
<td>0.95</td>
<td>1.21</td>
<td>1.40</td>
<td>1.80</td>
<td></td>
</tr>
</tbody>
</table>

† Values are shown for the “cardiovascular” disease endpoint, except for “history of dysrhythmia” and “history of fracture” which relate to these specific endpoints.

‡ ANOVA for ages, χ² test for the other categorical variables. Differences between the four groups were classified as significant if P value <0.01.

### TABLE 2. Unadjusted and adjusted HRs for in-patient admissions and death due to cardiovascular disease, osteoporotic fracture, and dysrhythmias by TSH control

<table>
<thead>
<tr>
<th>Cardiovascular admission/death a</th>
<th>Population</th>
<th>Events (%)</th>
<th>Unadjusted HR (95% CI)</th>
<th>Adjusted HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suppressed TSH</td>
<td>1,070</td>
<td>180 (16.8)</td>
<td>1.21 (1.04–1.42)</td>
<td>1.37 (1.17–1.60)</td>
</tr>
<tr>
<td>Low TSH</td>
<td>3,731</td>
<td>478 (12.8)</td>
<td>0.99 (0.89–1.11)</td>
<td>1.10 (0.99–1.23)</td>
</tr>
<tr>
<td>Normal TSH</td>
<td>10,908</td>
<td>1,141 (10.5)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>High TSH</td>
<td>1,975</td>
<td>345 (17.5)</td>
<td>2.36 (2.09–2.67)</td>
<td>1.95 (1.73–2.21)</td>
</tr>
<tr>
<td>Dysrhythmia admission/death b</td>
<td>1,006</td>
<td>32 (3.2)</td>
<td>1.32 (0.91–1.91)</td>
<td>1.60 (1.10–2.33)</td>
</tr>
<tr>
<td>Low TSH</td>
<td>3,752</td>
<td>84 (2.2)</td>
<td>1.02 (0.79–1.32)</td>
<td>1.13 (0.88–1.47)</td>
</tr>
<tr>
<td>Normal TSH</td>
<td>11,014</td>
<td>196 (1.8)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>High TSH</td>
<td>1,912</td>
<td>55 (2.9)</td>
<td>2.25 (1.67–3.04)</td>
<td>1.80 (1.33–2.44)</td>
</tr>
<tr>
<td>Osteoporotic fracture admission/death c</td>
<td>1,020</td>
<td>70 (6.9)</td>
<td>1.86 (1.43–2.42)</td>
<td>2.02 (1.55–2.62)</td>
</tr>
<tr>
<td>Low TSH</td>
<td>3,741</td>
<td>135 (3.6)</td>
<td>1.08 (0.88–1.33)</td>
<td>1.13 (0.92–1.39)</td>
</tr>
<tr>
<td>Normal TSH</td>
<td>11,012</td>
<td>284 (2.6)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>High TSH</td>
<td>1,911</td>
<td>73 (3.8)</td>
<td>2.21 (1.71–2.86)</td>
<td>1.83 (1.41–2.37)</td>
</tr>
</tbody>
</table>

a Adjusted for age, sex, history of hyperthyroidism, history of cardiovascular disease, socioeconomic status, diabetic status, and age×sex interaction.

b Adjusted for age, sex, history of hyperthyroidism, history of (nondysrhythmia) cardiovascular disease, history of dysrhythmias, age×sex interaction.

c Adjusted for age, sex, history of hyperthyroidism, history of osteoporotic fracture, diabetic status.
fit was a second-degree fractional polynomial (Fig. 2), with the highest level of risk associated with extremes of TSH concentration (i.e. “suppressed” and “high”) and the lowest risk associated with TSH at the lower end of normal.

Discussion

The current study was large and population-based, including all patients in Tayside, Scotland, who were taking thyroid medication and eligible for analysis (n = 17,684). At least 70% of subjects in this cohort had received at least 6 months of treatment before commencement of follow-up, so that the effect of any previous thyroid condition could be minimized. This is because previous studies have shown an increased death rate after radioiodine therapy for hyperthyroidism, which diminishes with time and is not sustained (35, 36). We have previously shown that including only patients after they are well established on T4 eliminates any time dependent effect on cardiovascular and other outcomes (25). Thus, our results are likely to reflect the long-term effects of T4 replacement, in the main, rather than any short-term impact of previous thyroid conditions.

Our data indicate that patients taking long-term T4 replacement are at greatest risk of cardiovascular disease if they have a serum TSH greater than 4.0 mU/liter. Such patients were more likely to be male. It is possible that such patients were poorly concordant with their T4 medication, as reflected by their TSH, and by inference this might imply less concordance with other coprescribed medication such as statins and angiotensin-converting enzyme inhibitors, as well as lifestyle advice, all of which would adversely impact on their risk of cardiovascular events. Raised serum TSH is also associated with raised serum cholesterol (37), which may be another contributing factor over the long term. Patients with a suppressed TSH also had increased risk of cardiovascular events, but not to the same degree. This may be due to a direct effect of hyperthyroxinemia on cardiac myocytes resulting in dysrhythmias, such as atrial fibrillation, which can cause thromboembolic disease.

Patients with a high TSH and those with a suppressed TSH were both at increased risk of fracture, although on this occasion the risk of a suppressed TSH was greatest. This may be because hyperthyroxinemia associated with a suppressed TSH is more likely to be detrimental to bone health than poor concordance with medications, which was thought to contribute to raised TSH concentrations.

Interestingly, patients with a serum TSH below the reference range, but not suppressed (0.04–0.4 mU/liter), had no increased risk of cardiovascular disease, dysrhythmias, or fractures. It is unfortunate that we did not have access to serum free T4 concentrations in these patients to ascertain whether they were above or within the laboratory reference range. However, our data indicate that it may be safe for patients to be on a dose of T4 that results in a low serum TSH concentration, as long as it is not suppressed.
at less than 0.03 mU/liter. Many patients report that they prefer such T4 doses (9, 10). Figure 2 indicates that the best outcomes appear to be associated with having a TSH within the lower end of the reference range.

There is a long-standing debate regarding how T4 replacement should be managed. The reason for this is that there is a lack of clear evidence to guide clinicians. Most guidelines recommend trying to achieve a “normal” TSH (5–8), although some commentators recommend a greater emphasis on clinical symptoms when deciding the correct T4 dose and will lower the TSH concentration (38), whereas others adjust L-T4 replacement to achieve a higher TSH for some patients, e.g. those with dysrhythmias.

Our study has a number of strengths. The data are from a large and representative population base of 400,000 subjects with low migration, and as such the findings will have high external validity. The long-term nature of follow-up for patients routinely treated in primary care is also a key strength. The data on which this study is based are well established and are of proven high quality (1, 2, 25). All patients were taking T4, and the study used virtually all TSH measurements, and not just one or two baseline measurements.

The observational nature of the study is a weakness because the findings could be the result of residual confounding by unknown factors. However, our data have been adjusted to account for age, sex, history of previous thyroid condition, history of cardiovascular disease, socioeconomic status, and presence of diabetes. Also, some TSH tests will have been performed at times of acute severe illness. Although we excluded all TSH recordings made on patients while they were in-patients to minimize the impact of sick-euthyroid syndrome, we have not accounted for intercurrent illnesses in outpatients. Summary statistics based on periodic and possibly irregular TSH serum measures to describe TSH control were used over the period of the study. However, given the long-term nature of the mechanisms underlying both cardiovascular disease due to dyslipidemia and fracture due to reduced bone mineral density, the use of such a statistic is an appropriate measure and is likely to be more sensitive than using baseline TSH measurements.

In summary, patients on long-term T4 with either an increased serum TSH (>4 mU/liter) or a suppressed TSH (<0.03 mU/liter) have an increased risk of cardiovascular
disease, dysrhythmias, and fractures when compared with patients with a TSH within the laboratory reference range. Patients with a low, but not suppressed, TSH (0.04–0.4 mU/liter) had no increased risk of these outcomes in this study. Although most current guidelines do not generally recommend this, it may be safe for patients on T4 to have a low but nonsuppressed TSH concentration. Further work is required to confirm this finding.

Acknowledgments

We thank Chest Heart Stroke Scotland and the Anonymous Trust, University of Dundee, for providing grants to support this work.

Address all correspondence and requests for reprints to: Graham Leese, Ward 1 and 2, Ninewells Hospital and Medical School, Dundee DD1 9SY, United Kingdom. E-mail: graham.leese@nhs.net.

R.W.F. and G.P.L. had full access to the data and take responsibility for the integrity of the data and the accuracy of the data analysis. They were involved in the study design, data analysis, and manuscript writing. S.R.B. was involved in the data analysis. R.T.J. was involved in the study design and the manuscript writing. T.M.M. and A.D.M. were involved in writing the manuscript. Disclosure Summary: All of the authors have nothing to declare as a conflict of interest and have no financial interests in this work.

References

1. Flynn RW, MacDonald TM, Morris AD, Jung RT, Leese GP 2004 The thyroid epidemiology, audit, and research study: thyroid dysfunction in the general population. J Clin Endocrinol Metab 89: 3879–3884
5. Singer PA, Cooper DS, Levy EG, Ladenson PW, Braverman LE, Daniels G, Greenspan FS, McDougal IR, Nikolai TF 1995 Treatment guidelines for patients with hyperthyroidism and hypothyroidism. JAMA 273:808–812
create a diabetes register. DARTS/MEMO Collaboration. BMJ 315:524–528


38. Toft AD, Beckett GJ 2003 Thyroid function tests and hypothyroidism: authors’ reply. BMJ 326:1087