

## CLINICAL STUDY

# Cognitive functioning and well-being in euthyroid patients on thyroxine replacement therapy for primary hypothyroidism

Ellie M Wekking, Bente C Appelhof<sup>1</sup>, Eric Fliers<sup>1</sup>, Aart H Schene, Jochanan Huyser, Jan G P Tijssen<sup>2</sup> and Wilmar M Wiersinga<sup>1</sup>

Department of Psychiatry, <sup>1</sup>Department of Endocrinology and Metabolism, and <sup>2</sup>Department of Cardiology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

(Correspondence should be addressed to E M Wekking; Email: ewekking@fsw.leidenuniv.nl)

E M Wekking is now at Leiden University, Departments of Psychology, Neurology and Neuropsychology, PO Box 9555, 2300 RB Leiden, The Netherlands

## Abstract

**Objective:** Hypothyroidism is associated with neurocognitive impairment. Sparse data suggest that treatment of hypothyroidism, resulting in a return to euthyroidism, may be associated with only partial recovery of overall neurocognitive functioning. The aim of this study was to assess neurocognitive functioning and well-being in euthyroid patients with primary hypothyroidism on adequate thyroxine (T4) treatment. We also investigated whether serum TSH and thyroid antibodies are determinants of neurocognitive functioning and well-being.

**Design:** We assessed neurocognitive functioning and well-being in 141 patients with primary hypothyroidism.

**Methods:** Neurocognitive test results and scores on questionnaires measuring well-being of 141 patients were compared with the reference values for these tests as published and used in Dutch clinical neuropsychological practice. Assessment of neurocognitive functioning included tests for cognitive or psychomotor speed, attention, working memory as well as learning and memory. Well-being was measured with the Symptom Check List-90 total score and the Rand 36-item Health Survey subscales for 'mental health' and 'vitality'.

**Results:** Patients showed poor performance on various domains of neurocognitive functioning compared with mean standard reference values, especially on a complex attention task and on verbal memory tests. Levels of well-being were significantly lower for patients compared with those of the general population. Neither serum TSH nor thyroid antibodies were determinants of neurocognitive functioning and well-being.

**Conclusion:** The results of this study suggest that neurocognitive functioning as well as psychological well-being may not be completely restored in patients with hypothyroidism, despite T4 treatment.

*European Journal of Endocrinology* 153 747–753

## Introduction

Overt hypothyroidism is associated with deficits in general intelligence, memory, attention, psychomotor speed, visuoperceptual and constructional skills (1). Thyroxine (T4) replacement therapy is the standard treatment of hypothyroidism that appears effective in restoring biochemical euthyroidism as evidenced by serum thyrotropin (TSH), and free T4 and triiodothyronine (T3) concentrations within the normal range. Although hormonal substitution therapy has been very successful in reducing morbidity and mortality of primary hypothyroidism, it is well known in clinical practice that a minority of patients have persistent symptoms despite adequate T4 replacement therapy. These are often vague, aspecific complaints about

fatigue, muscle aches, depressed mood or decreased memory function. A recent community study (2) provided the first evidence to indicate that patients on thyroxine replacement even those with a normal serum TSH display significant impairment in psychological well-being compared with controls of similar age and sex. These findings raise the question whether neurocognitive functioning in T4-treated hypothyroid patients might also be impaired. A review of neurocognitive aspects of hypothyroidism suggested that successful treatment of hypothyroidism might be associated with only partial and typically inconsistent patterns of recovery of overall neurocognitive functioning (1). However, this conclusion was reached from only a few studies on neurocognitive functioning carried out

shortly after re-establishing euthyroidism with thyroid replacement therapy in hypothyroid patients. In 1969, Whybrow *et al.* reported on four such patients, and concluded that when the hypothyroidism had been long-standing, brain impairment persisted after thyroid replacement therapy (3). Osterweil *et al.* reported on neurocognitive functioning of 54 patients with overt or minimal hypothyroidism (4). A subset of 11 to 34 patients was available for neurocognitive retesting. It was concluded that the deleterious effect of hypothyroidism is at least partly reversible, but possibly not completely, as some test results remained significantly impaired compared with euthyroid controls without hypothyroidism (4). A single case report by Mennemeier *et al.* reported similar conclusions (5).

As almost no data are available concerning neurocognitive functioning of patients with T4-treated hypothyroidism, we investigated the level of cognitive functioning in a large group of Dutch patients, all adequately treated for primary autoimmune hypothyroidism, and compared the results with the normal reference values for these tests. This is the first study to evaluate neurocognitive functioning after long-standing T4 therapy (median duration of substitution therapy 5.5 years). Secondly, we investigated whether psychological well-being in this patient group differs from healthy subjects. Finally, we examined whether serum TSH or thyroid peroxidase antibodies (TPO-ab) are determinants of neurocognitive performance and well-being.

## Subjects and methods

### Subjects

Patients were recruited from 13 general practices in the cities of Amsterdam and Almere. At these practices, prescribing records were checked to identify all patients receiving levothyroxine (T4) treatment. After checking for exclusion criteria, patients eligible for participation were invited by letter to participate in the study. All patients were between 18 and 70 years of age and had been on T4 replacement therapy for primary autoimmune hypothyroidism for at least 6 months. At inclusion, all patients were on an adequate dose of T4 as evidenced by a serum TSH concentration between 0.11 and 4.0 mU/l, as measured in the morning before levothyroxine intake. The lower end of this TSH interval was chosen because TSH levels of 0.10 mU/l or lower are associated with a higher risk of atrial fibrillation and bone loss, suggesting a too high T4 replacement dose. One hundred and forty-one patients participated in the study. This study was carried out between October 2001 and December 2003 at the Academic Medical Center of the University of Amsterdam. The protocol was approved by the institutional ethics review committee and all patients provided written informed consent. For more details see Appelhof *et al.* (6).

### Study design

Patients arrived in the morning in a fasting state. After blood sampling, patients took their usual dose of levothyroxine and had breakfast at the Academic Medical Center. About one hour later they proceeded with the neurocognitive testing session that lasted for approximately 1.5 h, and handed over the baseline set of questionnaires that they had filled out on the day prior to the visit.

### Neurocognitive tests

Assessment of neurocognitive functioning included neuropsychological domains that are believed to be affected in hypothyroid patients (7) and consisted of tests for cognitive or psychomotor speed, attention, working memory as well as learning and memory.

Cognitive or psychomotor speed was measured by the following tests. (1) The Dutch adaptation of the Digit Symbol subtest (8, 9) in which patients are presented with a key that pairs digits to geometrical symbols. Subsequently, patients are presented with a string of symbols and are required to pair each symbol with the correct digit, using the key, during 90 s. The raw score consists of the total number of correctly completed entries. The score is transformed into a standard value based on a normative population ( $n = 745$ ) with a score range from 1 to 19, with a mean of 10 and a standard deviation (s.d.) of 3. (2) The Memory Comparison Task, paper and pencil version (MCT) (10). During this task patients have to memorize visually presented target letters. Subsequently, they have to detect these on a paper amongst other non-target letters and strike out the target letter(s). The test consists of an example-task (%) and 4 subtasks with increasing memory load (1, 2, 3 or 4 letters to memorize). The time in seconds needed to complete the respective subtasks is measured.

Attention was measured by the Dutch adaptation of the Paced Auditory Serial Addition Task (PASAT) (11, 12). In this task, tape-recorded one digit numbers are presented in different blocks of 60 items at different rates. The test consists of 5 blocks with decreasing interstimulus intervals. Patients are required to add every pair of successive numbers and immediately give the addition sum. The total number of correct answers, calculated over all 5 blocks, is scored (13).

Working memory was measured by the Digit Span subtest (WAIS-III) (8, 9) in which the patient is required to repeat spoken digit series in the same order (forward recall) or in reverse order (backward recall). With each series, the number of digits increases. The total score consists of the total number of correctly repeated digits of both the forward and backward tests. The total score is converted into a standard score in the same manner as for the Digit Symbol test (8).

Learning and memory were tested by the following tests. (1) The Dutch adaptation of the California Verbal Learning Test (CVLT) (14). A shopping list is presented orally 5 successive times. The outcome measures include (a) the total number of items correctly recalled after 5 successive trials (immediate recall), (b) the number of items recalled after a 20-min delay (delayed recall) and finally (c) the number of items correctly identified amongst distracting items (recognition). These three scores are transformed into a standard value with a mean of zero and a s.d. of 2 based on the Dutch CVLT-reference group (14). (2) The Dutch version of the Rivermead Behavioral Memory Test, Story Recall (Rivermead, stories) (15, 16). Patients have to repeat two short stories immediately after oral presentation and after a 20-min delay. The sum of the correctly recalled elements of the two stories is calculated, both for the immediate recall as well as for the delayed recall. In addition, the ratio of the delayed recall relative to the immediate recall is calculated (relative delayed recall). The raw scores are converted into standard scores (T-scores, mean = 50, s.d. = 10) (17).

All tests are widely used and internationally accepted (18), or have previously been used in earlier research on neurocognitive functioning in hypothyroid patients before and after combined therapy with T4 and T3 (MCT, paper and pencil version, also known as visual scanning test) (19). The tests were administered by a trained psychometrician, under the supervision of a clinical neuropsychologist (E W). All investigations were carried out once in each patient. Raw test scores, except for the PASAT, are corrected for education, age and gender if these variables in the reference groups proved to be relevant, and converted into standard scores. The PASAT score is corrected only for education. Finally, all test scores of the T4-treated hypothyroid patients were compared with the scores in the reference groups, thus following the standard procedure in Dutch clinical neuropsychological practice.

### Questionnaires

Well-being was measured by means of self-report questionnaires, filled out on the day prior to the study visit and included the (Dutch version of) the Symptom Checklist (SCL-90) (20) and the 'mental health' and 'vitality' subscales of the Rand 36-item Health Survey (Rand-36) (21). The SCL-90 is a 90 item self report scale containing 8 subscales, measuring multidimensional psychopathology. In this study, the total score (the sum-score of all 90 items) is reported. This score can be considered as a measure of psychoneuroticism or general level of psychological and somatic well-being. The higher the total score, the higher the level of psychoneuroticism and the lower the level of well-being. The SCL-90 is a widely used international instrument in studies on psychological consequences of somatic diseases (22). The scores are compared with

the reference values for the Dutch general population, based on a sample of 2368 individuals with a mean age of 41 years (20). The Rand-36 is the Dutch version of the MOS Short-Form General Health Survey (23) and measures health-related quality of life. It is known for its high standards of reliability and validity (24, 25). In this study, only the subscales 'mental health' and 'vitality' were used, measuring feelings of depression and nervousness, and feelings of energy and fatigue respectively. Scores are compared with reference values of a representative sample of the Dutch population ( $n = 1063$ ) with a mean age of 44 years.

### Biochemical measurements

All blood samples were collected in the morning before T4 medication was ingested. Serum TSH and free T4 were measured by time-resolved fluoroimmunoassay (hTSH Ultra and Delfia FT4 respectively; Wallac Oy, Turku, Finland), with a detection limit of 0.01 mU/l and 2 pmol/l respectively; serum T3 was measured by an in-house RIA method (26), and anti-thyroid peroxidase antibodies (TPO-ab) by chemiluminescence immunoassay LUMI-test anti-TPO (Brahms, Berlin, Germany) with a detection limit of 30 kU/l. The cut-off value for a positive anti-TPO titer was 60 kU/l.

### Statistical methods

All data were analyzed with SPSS version 11.5 (SPSS, Chicago, IL, USA). The mean scores of the neurocognitive tests were compared with the mean scores of the reference population by means of a one-sample *t*-test. Pearson correlation coefficients were used in order to study the relationship between neurocognitive functioning, well-being and serum TSH values. Serum TSH was log-transformed to normalize distribution before statistical analysis. Statistical significance was defined as a 2-tailed  $P < 0.05$ . Test results were considered impaired if below the 5th percentile of the reference groups.

### Results

In total, 141 patients were tested. The patient characteristics are shown in Table 1. Patients were predominantly middle-aged females, with a mid-level educational standard.

Table 2 shows the mean scores of the neurocognitive tests and well-being questionnaires of the patients and the reference populations.

### Neurocognitive tests

**Cognitive speed/psychomotor speed** Digit symbol scores of the patients were not significantly different from those of the reference population. However, patients performed significantly slower on the easier

**Table 1** Patient characteristics ( $n = 141$ ).

Female ( $n, \%$ )	121 (85.8%)
Age (mean $\pm$ s.d.)	47.8 $\pm$ 9.8
Duration of substitution therapy (yr; median, range)	5.5 (0.5 – 25)
Serum TSH at inclusion (mU/l median, range)	1.4 (0.11–4.0)
Serum free T4 (pmol/l; mean $\pm$ s.d.)	14.8 $\pm$ 2.9
Serum T3 (nmol/l; mean $\pm$ s.d.)	1.7 $\pm$ 0.33
Educational level ( $n, \%$ ) <sup>1</sup>	
Primary education	15 (10.6%)
Education lower than mid-level	20 (14.2%)
Mid-level education	61 (43.3%)
College level	45 (31.5%)

<sup>1</sup>The level of education is not based on years of education but on level of education attained.

parts of the MCT paper and pencil tasks (% and 1 letter) compared with the reference population. In fact, for these tests 26.4% and 16.4% of the patients respectively had impaired scores, i.e. scores below the 5th percentile of the reference values. By contrast, patients performed significantly better on the most complex task (4 letters) compared with the reference group.

**Attention** PASAT scores of the patients were significantly lower compared with the reference group, indicating impaired performance on this complex attention task. Impaired scores (below the 5th percentile of

the reference values) were observed in 14.9% of the patients. PASAT scores were corrected for educational level but not for age. However, correlation analyses revealed no significant relationship between age and the education-corrected PASAT score ( $r = -0.12, P \leq 0.24$ ).

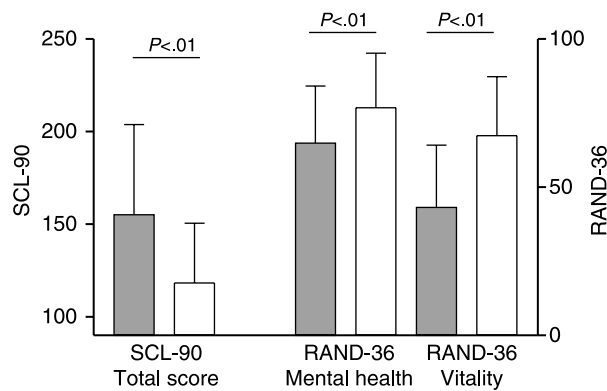
**Working memory, and learning and memory** For Digit Span there were no significant differences between patients and the reference population. Patients had significantly lower CLVT scores for immediate recall, and 10.6% of the patients scored more than 2 standard deviations below the mean of the reference population (normally about 2.5%). For the Rivermead stories patients had significantly lower scores on all indices, indicating worse performance on this memory test as compared with the reference population. Both immediate and delayed recall scores were impaired in 25% of the patients – for the relative delayed recall this was 18.7%.

### Questionnaires

Figure 1 shows the scores on the SCL-90 and Rand-36 questionnaires. T4-treated patients had significantly higher SCL-90 total scores compared with the Dutch general reference population. Using SCL-90, the mean total score of the patients would be classified as 'high'

**Table 2** Scores on neurocognitive tests and well-being questionnaires. Higher scores indicate better performance/higher level of well-being except for MCT and SCL-90.

	<i>n</i>	Patients	<i>n</i>	Reference population	<i>P</i> value
<b>Cognitive psychomotor speed</b>					
Digit symbol	140		745		
Total score		9.9 $\pm$ 3.8		10 $\pm$ 3	0.69
MCT (paper & pencil version)	140		1778		
% symbol (s)		21.6 $\pm$ 7.0		17.7 $\pm$ 3.7	0.001
1 letter (s)		26.2 $\pm$ 6.6		23.8 $\pm$ 4.7	0.001
2 letters (s)		36.6 $\pm$ 9.3		35.2 $\pm$ 7.6	0.07
3 letters (s)		45.0 $\pm$ 12.0		45.9 $\pm$ 11.2	0.38
4 letters (s)		57.0 $\pm$ 15.9		61.3 $\pm$ 17.4	0.002
<b>Attention</b>					
PASAT	134		124		
Total score		146.5 $\pm$ 45.5		170.9 $\pm$ 32.7	0.001
<b>Working memory</b>					
Digit span	140		745		
Total score		10 $\pm$ 3		10 $\pm$ 3	0.73
<b>Learning and memory</b>					
CVLT	141		492		
Immediate recall		-1.1 $\pm$ 2.2		0 $\pm$ 2	0.001
Delayed recall		0.19 $\pm$ 2.6		0 $\pm$ 2	0.38
Recognition		-0.11 $\pm$ 1.9		-0.14 $\pm$ 1.9	0.52
Rivermead (stories)	139		284		
Immediate recall		42.5 $\pm$ 11		50 $\pm$ 10	0.001
Absolute delayed recall		41.7 $\pm$ 11.6		50 $\pm$ 10	0.001
Relative delayed recall		44.5 $\pm$ 15.6		50 $\pm$ 10	0.001
<b>Well-being</b>					
SCL-90 total score	140	155.1 $\pm$ 48.7	2368	118.3 $\pm$ 32.3	0.001
RAND-36 Mental health	140	64.9 $\pm$ 19.2	1063	76.8 $\pm$ 18.4	0.001
RAND-36 Vitality	140	43.2 $\pm$ 21.0	1063	67.4 $\pm$ 19.9	0.001



**Figure 1** Questionnaire scores for patients (shaded bars) and the reference group (open bars). Lower levels of well-being are indicated by higher SCL-90 scores and lower RAND-36 scores.

(a score between the 80th and 95th percentile of the reference values for the general population), indicating a lower mean level of general psychological and somatic well-being. Likewise, scores on the RAND-36 mental health and vitality subscales were significantly lower as compared with healthy controls of roughly the same age, indicating a lower level of health-related quality of life. The RAND-36 score for mental health lies within 1 standard deviation below the mean of the reference population, whereas the score for vitality was between 1 and 2 standard deviations below the mean. For the RAND-36, these figures indicate that the score for vitality (in particular), which measures feelings of energy and fatigue, is impaired in hypothyroid patients.

### **Serum TSH, TPO-ab and their relationship with neurocognitive functioning and well-being**

The patient group was divided into patients with a mean TSH lower than 2 mU/l ( $n = 101$ ) and patients with a mean TSH higher than 2 mU/l ( $n = 40$ ). There appeared to be no significant differences between these groups on either the neurocognitive tests or the SCL-90 and RAND-36 scores. Furthermore, scores on neurocognitive tests and well-being questionnaires were not significantly different for patients with a positive ( $> 60$  kU/l) TPO-ab titer ( $n = 29$ ) compared with patients with a negative ( $< 60$  kU/l) TPO-ab titer ( $n = 112$ ).

### **Discussion**

In this study, we investigated several domains of neurocognitive functioning in a large group of patients with treated hypothyroidism. The results were compared with standard reference values as applied in Dutch clinical neuropsychological practice. In addition, we

evaluated whether psychological well-being in this patient group differed from healthy subjects and whether serum TSH was a determinant of neurocognitive performance and well-being. Patients had poor performance on various domains of neurocognitive functioning compared with the mean standard reference values, especially on a complex attention task and on verbal memory tests. No less than a quarter of the patients had impaired scores (below the 5th percentile) on the memory test story recall, both for immediate reproduction and for delayed recall. These findings support the sparse findings from earlier studies reporting a significant decrease in memory function in hypothyroid patients after adequate thyroid replacement treatment as compared with normal controls (4). Similarly, Burmeister *et al.* reported on the neurocognitive functioning of 13 athyrotic patients while on T4 therapy and again after discontinuing thyroid hormone treatment (7). It was reported that discontinuing T4 therapy led to a decrement in delayed recall of verbal information, but not in other objective measures of cognition. Performance on tests for cognitive speed seemed contradictory: patients had significantly slower performance at the beginning of the tests (MCT paper and pencil), but could compare favourably with the reference group, or perform even better, when the task progressed and increased in difficulty. These findings could indicate some 'starting slowness'. This could also explain why Rivermead delayed recall but not CLVT delayed recall was impaired. In the former test, the stories to be memorized are presented only once, whereas in the latter, a shopping list is repeated 5 times, giving patients the opportunity to compensate for 'starting slowness'. This starting slowness could also account for part of the attention impairment.

Levels of well-being, as measured with the SCL-90 total score and Rand-36 subscales for mental health and for vitality, were significantly lower for patients compared with those of the general population. This corroborates the findings of the community-based study of Saravanan *et al.* (2), who found that patients on thyroxine replacement display significant impairment in psychological well-being as compared with controls of similar age and sex. One could hypothesize that a decreased level of well-being is the cause of decreased neurocognitive performance in this population. Indeed, correlation analyses revealed that the SCL-90 total score was weakly, but significantly, correlated with about half, but not all of the neurocognitive test results (correlation coefficients between  $-0.19$  and  $-0.24$ ). Serum TSH concentrations were not related to outcome on neurocognitive function tests nor with scores on questionnaires concerning well-being. Apparently, neurocognitive functioning and well-being in patients with serum TSH values above 2 mU/l (but below 4.0 mU/l) is not different from patients with TSH values below 2 mU/l, but this might well be

different for patients with abnormal serum TSH concentrations. However, Osterweil *et al.* found no correlation between serum TSH levels and neurocognitive functioning in patients during hypothyroidism (mean serum TSH: 66 mU/l) either (4). They suggested that the duration of untreated hypothyroidism might be a more significant determinant. Furthermore, dividing the group into patients with positive TPO antibodies (>60) versus patients with negative TPO antibodies ( $\leq$ 60) did not result in significant differences in the results of tests and questionnaires, possibly indicating that a high level of thyroid antibodies is not related to neurocognitive dysfunctions and a lower level of well-being.

A major limitation of this study is the lack of a proper control group, with controls individually matched for age, sex and educational level. Results of neurocognitive tests and questionnaires were compared with their reference values for the general population. Therefore, we cannot exclude the possibility that these results are due to differences in patient characteristics. On the other hand, the reference values of the neurocognitive tests have already been corrected for factors relevant to the specific test, such as age and education, and are commonly used to evaluate individual scores in Dutch clinical neuropsychological practice.

In summary, despite its limitations, this study is the first to provide some evidence for cognitive impairment in a large group of adequately treated patients with hypothyroidism, especially concerning the domains of attention and memory. Results of tests for cognitive speed revealed some 'starting slowness', but otherwise normal performance. Levels of well-being, as measured with the SCL-90 total score and Rand-36 subscales for mental health and for vitality, were significantly lower for patients compared with those of the general population. Our findings corroborate earlier sparse reports, as well as the widespread clinical notion that neurocognitive function and psychological well-being may not be completely restored in patients with hypothyroidism, despite adequate T4 treatment. Whether these impairments in neurocognitive functioning and well-being could be due to imperfections of the substitution therapy, because cognitive and affective functioning is not fully restored, or to the autoimmune process that caused the hypothyroidism, or reflect 'having a chronic disease' in general, remains to be clarified. Future studies might first attempt to confirm the presence of neurocognitive impairment in T4-treated patients in a controlled study, comparing patients with healthy individuals, closely matched for sex, age and educational level. In order to investigate whether symptoms may be due to chronic disease in general, it would be interesting to compare levels of well-being and neurocognitive function of T4-treated patients with hypothyroidism, with those of a group of patients who suffer from an equally chronic disease

that can be treated with a single portion of medication a day, for example patients with essential hypertension.

## Acknowledgements

This study was supported by a grant from the Academic Medical Center Anton Meelmeijer Fund (fund number: SAG 05002).

We would like to thank Leonie van der Heuvel, Diane Jungerman, Tea Elezovic, Tanja de Krijger, Famke Stegenga and Aram Tefsen for the assessment of the neuropsychological tests.

## References

- 1 Dugbartey AT. Neurocognitive aspects of hypothyroidism. *Archives of Internal Medicine* 1998 **158** 1413–1418.
- 2 Saravanan P, Chau WF, Roberts N, Vedhara K, Greenwood R & Dayan CM. Psychological well-being in patients on 'adequate' doses of l-thyroxine: results of a large, controlled community-based questionnaire study. *Clinical Endocrinology* 2002 **57** 577–585.
- 3 Whybrow PC, Prange AJ Jr & Treadway CR. Mental changes accompanying thyroid gland dysfunction. A reappraisal using objective psychological measurement. *Archives of General Psychiatry* 1969 **20** 48–63.
- 4 Osterweil D, Syndulko K, Cohen SN, Pettler-Jennings PD, Hershman JM, Cummings JL, Tourtellotte WW & Solomon DH. Cognitive function in non-demented older adults with hypothyroidism. *Journal of the American Geriatrics Society* 1992 **40** 325–335.
- 5 Mennemeier M, Garner RD & Heilman KM. Memory, mood and measurement in hypothyroidism. *Journal of Clinical and Experimental Neuropsychology* 1993 **15** 822–831.
- 6 Appelhof BC, Fliers E, Wekking EM, Schene AH, Huyser J, Tijssen JG, Endert E, Van Weert HCPM & Wiersinga WM. Combined therapy with levothyroxine and liothyronine in two ratios compared with levothyroxine monotherapy in primary hypothyroidism; a double blind, randomized controlled clinical trial. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 2666–2674.
- 7 Burmeister LA, Ganguli M, Dodge HH, Toczek T, DeKosky ST & Nebes RD. Hypothyroidism and cognition: preliminary evidence for a specific defect in memory. *Thyroid* 2001 **11** 1177–1185.
- 8 Uterwijk J. *WAIS-III. Nederlandstalige bewerking. Technische Handleiding*. Lisse: Swets & Zeitlinger, 2000.
- 9 Wechsler D. *Wechsler Adult Intelligence Scale III: Technical Manual*. San Antonio, TX: The Psychological Corporation, 1997.
- 10 Brand AN & Jolles J. Information processing in depression and anxiety. *Psychological Medicine* 1987 **17** 145–153.
- 11 Gronwall DMA & Sampson H. *The Psychological Effects of Concussion*. Auckland, New Zealand: Auckland University Press, 1974.
- 12 Spikman JM, van Zomeren AH & Deelman BG. Deficits of attention after closed-head injury: slowness only? *Journal of Clinical and Experimental Neuropsychology* 1996 **18** 755–767.
- 13 Aarnoudse CC, van den Burg W & Saan R. *De Paced Auditory Serial Addition Task (PASAT) in Een Steekproef van Gezonde Personen. Betrouwbaarheden en Normeringen*. Internal report, Department of Neuropsychology, Academic Hospital, University of Groningen, The Netherlands. Ref Type: Report, 1995.
- 14 Mulder JL, Dekker R & Dekker PH. *Verbale Leer en Geheugen Test: Handleiding*. Lisse: Swets & Zeitlinger, 1996.
- 15 Wilson B, Cockburn J & Baddeley AD. *The Rivermead Behavioural Memory Test*. Bury St Edmunds, UK: Thames Valley Test Company, 1985.

- 16 Wimmers WFHG & van Balen HGG. *Rivermead Behavioural Memory Test, Nederlandstalige Bewerking*. Lisse: Swets & Zeitlinger, 1987.
  - 17 Schmand B, Houx P & de Koning I. *Normen voor Stroop Kleurwoord Tests, Trail Making Test en Story Recall van de Rivermead Behavioral Memory Test*. Sectie Neuropsychologie, Nederlands Centrum voor Psychologen, Amsterdam, Ref Type: Report, 2005.
  - 18 Lezak MD, Howieson DB & Loring DW. *Neuropsychological Assessment*, 4th ed., Oxford, UK: Oxford University Press, 2004.
  - 19 Bunevicius R, Kazanavicius G, Zalinkevicius R & Prange AJ Jr. Effects of thyroxine as compared with thyroxine plus triiodothyronine in patients with hypothyroidism. *New England Journal of Medicine* 1999 **340** 424–429.
  - 20 Ettema JHM & Arrindell WA. Symptom Checklist-90. In *Handleiding Bij een Multi-dimensionele Psychopathologie-indicator*. Amsterdam: Harcourt Test Publisher, 2003.
  - 21 Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, Sprangers MA, te Velde A & Verrips E. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. *Journal of Clinical Epidemiology* 1998 **51** 1055–1068.
  - 22 Derogatis LR & Savitz KL. The SCL-90, brief symptom inventory, and matching clinical rating scales. In *The use of Psychological Testing for Treatment Planning and Outcome Assessment*. pp 679–724. Ed. ME Maruish. New Jersey: Erlbaum, 1999.
  - 23 Ware JE Jr & Sherbourne CD. The MOS 36-item short-form health survey (SF-36) I. Conceptual framework and item selection. *Medical Care* 1992 **30** 473–483.
  - 24 van der Zee KI, Sanderman R. *Het Meten van de Algemene Gezondheidstoestand Met de RAND-36. Een Handleiding*. Groningen: Rijksuniversiteit, 1993.
  - 25 van der Zee KI & Sanderman R, Heyink J & de Haes H. The psychometric qualities of the RAND 36-item Health Survey 1.0: a multidimensional measure of general health status. *International Journal of Behavioral Medicine* 1996 **3** 104–122.
  - 26 Wiersinga WM & Chopra IJ. Radioimmunoassay of thyroxine (T<sub>4</sub>), 3,5,3'-triiodothyronine (T<sub>3</sub>), 3,3',5'-triiodothyronine (reverse T<sub>3</sub>, rT<sub>3</sub>), and 3,3'-diiodothyronine (T<sub>2</sub>). *Methods in Enzymology* 1982 **84** 272–303.
- 

Received 27 April 2005

Accepted 1 September 2005