

A 2-YEAR PRELIMINARY LONGITUDINAL STUDY OF NEUROPSYCHOLOGICAL FUNCTIONING IN HASHIMOTO'S THYROIDITIS UNDER LEVOTHYROXINE TREATMENT: ONLY TRAIL MAKING TEST IS MAKING A DIFFERENCE

Vaitsa Giannouli

Bulgarian Academy of Sciences, Bulgaria
Email: giannouliv@hotmail.com

Nikolaos Syrmos

Aristotle University of Thessaloniki, Greece

Abstract

Few studies have examined longitudinally depression, stress and cognition in Hashimoto's Thyroiditis (HT) under levothyroxine (LT4) treatment. Although HT is claimed to cause several mood changes and cognitive deficits in patients who are left untreated, we still know little about the longitudinal profile of these patients when they are under stable treatment. Sixteen middle-aged euthyroid HT women and sixteen healthy control women were examined with a detailed neuropsychological battery covering different cognitive functions (attention, visual and verbal memory, executive functions and mood), and with thyroid functions tests at three time points at one and two years post-baseline. Depression and perceived stress, along with verbal and visual memory were found to be stable at a two-year retesting in euthyroid HT women. Only TMT-Parts A and B deteriorated at one and two years in euthyroid HT. Although impaired performance in attention and executive deficits are linked with heightened levels of depression and stress, here they were not accompanied by changes in mood. The present study failed to detect any noticeable changes longitudinally in the cognitive and emotional function in euthyroid HT women under stable and continuous LT4 treatment. Patients' restricted deficits may indicate that the decline in attention and executive functioning performance might be related to the effects of a specific progressive degenerative process associated with the disorder, rather than the effects of aging.

Keywords: Hashimoto's Thyroiditis, depression, stress, longitudinal neuropsychological assessment, LT4 treatment.

Introduction

Although it is generally accepted that thyroid function has an impact on various aspects of cognition and mood, only in the recent years an increasing number of studies report on the association of Hashimoto's Thyroiditis (HT) with cognitive and emotional disturbances also in the euthyroid state (Leyhe & Müssig, 2014). Thyroid hormones (within and outside normal ranges) and cognition are closely linked differentially by sex, race, and elevated depressive symptoms status (Beydoun et al., 2013).

More specifically, «within reference ranges elevated free thyroxine (fT4) is associated with better performance on tests of visuospatial/visuo-construction ability (overall, women, and African Americans) and learning/memory (women and African Americans), whereas a higher total thyroxine (tT4) level is associated with better performance in the domain of psychomotor speed (individuals without elevated depressive symptoms) and higher levels of both fT4 and tT4 are linked to better language/verbal test performance among men. In contrast, higher T3 is related to better performance on tests of visuo-spatial/visuo-construction ability and psychomotor speed among whites. Thus, a within reference range fT4 is linked to better performance on visuo-spatial/visuo-construction ability and psychomotor speed, whereas a below normal range TSH level (compared with the reference range) is linked to better performance in domains of psychomotor speed and attention» (Beydoun et al., 2013). In addition to that, in the case of Subclinical hypothyroidism (SCH), also called mild thyroid failure, which is diagnosed when peripheral thyroid hormone levels are within normal reference laboratory range but serum thyroid-stimulating hormone (TSH) levels are mildly elevated, a recent meta-analysis demonstrates a relationship between sHT and cognitive impairment only in individuals younger than 75 years of age and those with higher TSH concentrations (Pasqualetti et al., 2015).

Nevertheless, cognitive performance in subclinical hypothyroidism and euthyroid HT, whether under levothyroxine (LT4) treatment or not, still remains a vastly unexplored topic, while emotional factors that directly influence performance during neuropsychological assessment, such as stress and depression, are disregarded (Giannouli, Toulis, & Syrmos, 2014).

Some researchers claim poor performance on various domains of neurocognitive functioning (attention and verbal memory tests) as well as lower levels of well-being in euthyroid HT patients compared with those of the general population (Toudou Daouda et al., 2016), thus supporting that neurocognitive functioning and psychological well-being may not be completely restored in patients with hypothyroidism despite T4 treatment (Wekking et al., 2005). In addition, even slightly higher levels of free T3 and thyroglobulin antibodies levels within the normal range are claimed to associate with slower cognitive performance and more errors on neuropsychological tests of executive function (Grigorova & Shewin, 2012), while other studies find no such detrimental influences on cognition when baseline measures of neurocognitive functioning is normal (Giannouli, Toulis, & Syrmos, 2014).

Therefore, still little is known about the different aspects of cognition in specific homogeneous groups, such as middle-aged women, who are more likely to be found suffering from HT than other groups from the general population (Canaris et al., 2000; Caturegli, De Remigis & Rose, 2014) and who are under treatment for this disease. It still remains unanswered which is their neuropsychological profile in concurrent assessment of demographically equated patients and healthy participants, but also when taking into account their cognitive performance in time. Thus, this study aims to describe, compare and document the progression of cognitive abilities in euthyroid HT women who have normal cognitive functioning at baseline over a time period of two consecutive years, while controlling for the influence of emotional variables that can have a confounding influence, such as depression and perceived stress.

Methodology of Research

General Background of Research

This study was motivated by the need for data regarding the cognitive capacities of euthyroid HT women under treatment. The two groups of participants (healthy women and women with a diagnosis of euthyroid HT), were compared with the use of a neuropsychological battery consisting of several questionnaires and tests that covered a wide range of cognitive capacities and emotional variables.

Sample of Research

Sixteen middle-aged women without a diagnosis of a physical health and/or psychiatric-neurological problem, and not taking dietary or any other medical supplements (baseline:

$M_{\text{age}} = 58.00$, $SD_{\text{age}} = 6.66$; $M_{\text{education}} = 12.00$, $SD_{\text{education}} = 1.46$; $M_{\text{Geriatric Depression Scale}} = 3.00$, $SD_{\text{Geriatric Depression Scale}} = 1.54$; $M_{\text{Perceived Stress Scale}} = 8.06$, $SD_{\text{Perceived Stress Scale}} = 6.80$) and sixteen women with similar demographics (baseline: $M_{\text{age}} = 56.93$, $SD_{\text{age}} = 7.70$; $M_{\text{education}} = 12.50$, $SD_{\text{education}} = 1.36$; $M_{\text{Geriatric Depression Scale}} = 3.06$, $SD_{\text{Geriatric Depression Scale}} = 2.20$; $M_{\text{Perceived Stress Scale}} = 8.66$, $SD_{\text{Perceived Stress Scale}} = 5.24$), but with a diagnosis of HT under levothyroxine sodium treatment ($M_{\text{years of diagnosis-treatment}} = 6.68$, $SD_{\text{years of diagnosis-treatment}} = 2.54$) were included in this research.

Diagnosis of Hashimoto's disease was based on signs and symptoms and the results of blood tests that measure levels of thyroid hormone and thyroid-stimulating hormone (TSH) produced in the pituitary gland, including hormone tests, as well as antibody tests. The women with the HT diagnosis did not have any other medical diagnosis (e.g. cancer).

The two groups were equated regarding age ($t(30) = .417$, $p = .680$), education years ($t(30) = 1.000$, $p = 3.25$), perceived stress ($t(29) = .275$, $p = .785$), and depression reported symptoms ($t(30) = .093$, $p = .927$).

Instrument and Procedures

Data were collected using a data sheet containing socio-demographic, clinical, and therapeutic information of the patient. Informed consent was obtained from all individual participants included in the study.

Women with a diagnosis of HT were screened with thyroid function tests and thyroid antibodies including measures of thyroglobulin antibodies (Anti-TG) (> 150 IU/ml at all three testings), thyroid peroxidase antibodies (Anti-TPO) (> 150 IU/ml at all three testings), vitamin B12 (200-870), free thyroxine (fT4) (0.8-2), free triiodothyronine (fT3) (2-4.8), thyroxine (T4) (5-12), and thyroxine, thyrotropin (TSH) (0.3-5 μ IU/ml). These scores did not differentiate between the three time points and were within the normal range. In addition to that, they were under synthetic T4 (levothyroxine sodium) everyday treatment (dose range 75mcg-100mcg that had not been interrupted or changed during the last six months from the baseline and the other testings). All the women were recruited during a study examining attitudes (Giannouli & Syrmos, 2017) and after that in a study on cognitive functioning (Giannouli, 2017). Two women were excluded from the initial sample of the study on cognitive functioning, because these patients had a change in their treatment doses one year after.

Several cognitive domains were assessed with the following neuropsychological tests: attention [Trail Making Test-Part A (TMT-A, time to completion), Digit Span (WAIS-III) forward (greatest span)], executive functioning [Trail Making Test-Part B (TMT-B, time to completion), Digit Span (WAIS-III) backward (greatest span)], fluency (total number of words generated on semantic and on phonological fluency), verbal memory [word list learning (number of words on immediate and delayed recall), story learning (number of words on immediate and delayed recall)], and visual memory [Rey-Osterrieth Complex Figure Test (number of correct components on immediate and delayed recall trials)]. Mood was also examined with Geriatric Depression Scale (GDS) and Perceived Stress Scale (PSS).

Data Analysis

Data were processed statistically by means of SPSS 21 using descriptive statistics for presenting the baseline, one-year and two-year cognitive performance of the two groups of participants in a number of neuropsychological tests. A mixed between-within subjects analysis of variance (ANOVA) was performed with one categorical independent between-subjects variable with two levels (healthy women and HT euthyroid patients), one categorical independent within-subjects variable with three levels (baseline, first re-examination at one-year interval and second re-examination at two-year interval), and as continuous dependent variables the neuropsychological test scores.

Results of Research

For the depression levels there was no statistically significant effect in GDS for time $F(2,60) = 3.019$, $p = .060$, no statistically significant interaction between time and group $F(2,60) = 1.698$, $p = .192$, and no statistically significant difference for the two groups ($F(1,30) = .091$, $p = .765$). In the

same line, no statistically significant effect in stress, as measured with the self-report PSS, was found for time $F(2,52) = .976, p = .383$), no statistically significant interaction between time and group $F(2,52) = .435, p = .649$), and no statistically significant difference for the two groups ($F(1,26) = .086, p = .771$).

Therefore, in the statistical analyses depression and stress were not included as possible influencing factors as they did not differentiate. The mixed between-within ANOVA with one categorical independent between-subjects variable with two levels (healthy women and HT euthyroid patients), one categorical independent within-subjects variable with three levels (baseline, first re-examination at one-year interval and second re-examination at two-year interval), and as continuous dependent variables the neuropsychological test scores revealed a number of interesting findings (presented in Table 1).

Table 1. Cognitive assessment for healthy participants and euthyroid HT participants at baseline, one year and two years follow-up.

Neuropsychological test	Group	Baseline		One-year re-examination		Two-year re-examination	
		Mean	SD	Mean	SD	Mean	SD
<i>Visual memory</i>							
Rey-Osterrieth Complex Figure Test design copy	Healthy	30.56	4.35	32.00	2.89	31.81	2.74
	HT patients	28.34	7.24	28.15	7.36	28.00	7.29
Rey-Osterrieth Complex Figure Test immediate design recall	Healthy	15.21	4.94	15.84	4.75	15.71	4.68
	HT patients	14.31	7.75	14.18	7.60	14.06	7.55
Rey-Osterrieth Complex Figure Test delayed design recall	Healthy	14.37	4.25	12.87	6.12	13.06	6.24
	HT patients	13.68	4.60	10.84	7.16	10.68	7.01
<i>Verbal memory</i>							
Word list learning	Healthy	29.00	5.49	29.81	5.00	29.50	5.52
	HT patients	26.87	6.40	26.68	5.94	26.12	4.99
Immediate memory of story	Healthy	20.68	4.72	17.62	4.74	18.25	5.42
	HT patients	17.87	6.48	18.37	5.40	17.56	4.80
Delayed memory of words	Healthy	6.62	1.45	6.81	1.22	7.37	1.08
	HT patients	6.31	1.66	5.87	2.02	6.12	2.12
Word recognition	Healthy	19.56	1.03	19.00	2.52	19.00	2.55
	HT patients	19.18	1.32	17.87	4.27	18.06	3.97
Delayed memory of story	Healthy	11.06	3.41	10.12	4.14	10.12	4.45
	HT patients	8.75	5.44	9.75	4.35	9.87	4.31
<i>Executive functioning</i>							
Trail Making Test-Part B	Healthy	90.50	50.07	87.31	42.28	88.87	48.09
	HT patients	111.68	48.65	115.12	50.37	118.18	52.57
Fluency semantic	Healthy	9.12	3.53	9.18	4.08	9.06	3.66
	HT patients	8.87	3.28	9.06	3.43	9.06	3.56
Fluency phonological	Healthy	9.18	3.44	9.62	2.72	9.81	2.40
	HT patients	8.87	2.82	9.00	2.80	9.42	2.70
Digit span backward	Healthy	5.18	1.64	4.87	1.62	4.93	1.61
	HT patients	4.56	1.71	4.81	1.51	4.62	1.54

Neuropsychological test	Group	Baseline		One-year re-examination		Two-year re-examination	
		Mean	SD	Mean	SD	Mean	SD
<i>Attention</i>							
Trail Making Test-Part A	Healthy	63.56	26.45	62.93	26.43	62.37	25.89
	HT patients	66.43	23.59	69.93	23.51	71.75	24.54
Digit span forward	Healthy	7.12	1.20	6.31	2.15	6.25	2.20
	HT patients	6.56	2.09	6.06	1.76	5.93	1.61
<i>Mood</i>							
Geriatric Depression Scale	Healthy	3.00	1.54	2.75	2.23	2.56	1.26
	HT patients	3.06	2.20	2.81	2.19	3.00	2.03
Perceived Stress Scale	Healthy	8.06	6.80	8.00	6.79	8.50	4.92
	HT patients	8.75	4.71	8.41	4.39	7.62	6.44

More specifically, for the Rey design-Copy condition there was no statistically significant effect for time $F(2,60) = .551, p = .579$, no statistically significant interaction between time and group $F(2,60) = 1.148, p = .292$, and no statistically significant difference for the two groups ($F(1,30) = 3.043, p = .092$). For the Rey design-Immediate design recall condition there was no statistically significant effect for time $F(2,60) = .340, p = .713$, no statistically significant interaction between time and group $F(2,60) = 1.019, p = .367$, and no statistically significant difference for the two groups ($F(1,30) = 3.98, p = .053$). For the word list learning there was no statistically significant effect for time $F(2,60) = 1.625, p = .201$, no statistically significant interaction between time and group $F(2,60) = .224, p = .800$, and no statistically significant difference for the two groups ($F(1,30) = 3.183, p = .085$). For the immediate memory of story there was no statistically significant effect for time $F(2,60) = 2.069, p = .035$, no statistically significant interaction between time and group $F(2,60) = 3.644, p = .065$, and no statistically significant difference for the two groups ($F(1,30) = 0.76, p = .784$). For the delayed memory of words no statistically significant effect for time was found $F(2,60) = 1.496, p = .232$, no statistically significant interaction between time and group $F(2,60) = 1.968, p = .149$, and no statistically significant difference for the two groups ($F(1,30) = 2.670, p = .113$). For the word recognition there was no statistically significant effect for time $F(2,60) = .2247, p = .115$, no statistically significant interaction between time and group $F(2,60) = .321, p = .727$, and no statistically significant difference for the two groups ($F(1,30) = .915, p = .346$). For the Rey design-delayed recall condition word recognition there was a statistically significant effect for time $F(2,60) = 3.372, p = .041$, no statistically significant interaction between time and group $F(2,60) = 6.362, p = .043$, and no statistically significant difference for the two groups ($F(1,30) = .881, p = .355$). For the delayed memory of story there was no statistically significant effect for time $F(2,60) = .011, p = .989$, no statistically significant interaction between time and group $F(2,60) = 1.672, p = .196$, and no statistically significant difference for the two groups ($F(1,30) = .509, p = .481$).

For the digit span backward there was no statistically significant effect for time $F(2,60) = .120, p = .887$, no statistically significant interaction between time and group $F(2,60) = 1.042, p = .359$, and no statistically significant difference for the two groups ($F(1,30) = .407, p = .528$). For the digit span forward there was a statistically significant effect for time $F(2,60) = 7.604, p = .001$, no statistically significant interaction between time and group $F(2,60) = 2.712, p = .075$, and no statistically significant difference for the two groups ($F(1,30) = .000, p = 1.000$).

For the semantic fluency test there was no statistically significant effect for time $F(2,60) = .175, p = .874$, no statistically significant interaction between time and group $F(2,60) = .175, p = .873$, and no statistically significant difference for the two groups ($F(1,30) = .010, p = .921$). For the phonological fluency test there was no statistically significant effect for time $F(2,60) = 3.029, p = .061$, no statistically significant interaction between time and group $F(2,60) = .235, p = .792$, and no statistically significant

difference for the two groups ($F(1,30) = .206, p = .653$).

Statistically significant differences were found for executive functioning (only for the Trail Making Test-Part B) and attention (Trail Making Test-Part A). For the Trail Making Test-Part A there was a statistically significant effect for time ($F(2,60) = 8.260, p = .001$), and a statistically significant interaction between time and group ($F(2,60) = 19.971, p = .000$). For the Trail Making Test-Part B there was a marginally statistically significant effect for time ($F(2,60) = 3.102, p = .052$), and a statistically significant interaction between time and group ($F(2,60) = 7.693, p = .001$). In Trail Making Test Part A and B, it is clear that women with HT need more time to complete these two tasks at baseline, and although under treatment they deteriorate (have longer completion times) for both tests. This deterioration trend differentiates the HT women scores from healthy women who respond faster to the tasks and thus have reduced their reaction times at one year retesting and at two years (see Table 1).

Discussion

The present study did not detect any baseline visual and verbal memory deficits in middle-aged euthyroid women with HT who are under chronic LT4 treatment and no statistically significant differences were found at one and two years retesting. This is in line with previous results from a case study of a 62-year-old woman (Giannouli, Toulis & Syrmos, 2014), but not in line with results from a research in euthyroid patients with primary hypothyroidism on adequate thyroxine (T4) treatment, which supports that neurocognitive functioning as well as psychological well-being may not be completely restored in patients with hypothyroidism, despite T4 treatment (Wekking et al., 2005). In this study, although a slightly better cognitive performance for healthy women is reported, it does not reach statistical significance. This seems to support that any possible cognitive deficits due to prior hypothyroidism are largely reversible with the treatment if it is followed uninterrupted (Samuels, 2014). Of course, in this sample there was no pre-treatment neuropsychological testing, but this new view of memory functions from a long-term perspective supports that LT4 medication may be an effective way to protect against memory decline.

Although visual and verbal memory as measured with several tests were not affected, only executive functioning (Trail Making Test-Part B) and attention (Trail Making Test-Part A) deteriorated in HT women under treatment. These specific tests seem to be sensitive to changes, that are not detected in other cognitive domains. The above findings are in contrast to research that supports that in euthyroid women with low T4 levels (within the normal range) there is a greater risk of cognitive decline over a 3-year period (Volpato et al., 2002), as in this case cognitive impairment is found only in two tests and not in overall cognitive functioning. The cognitive deterioration in Trail Making completion times cannot be linked to depression and stress or to changes in treatment or thyroid function as no changes were recorded for these variables.

The methodology used demonstrates the potential to increase our understanding of how neuropsychological functioning changes and interacts with treatment over time, but some methodological limitations of this study should be considered in the interpretation of the results: the small sample size of the groups studied and the possible existence of practice effects (for the retest conditions), due to the learning that may result from the repeated exposure to the testing materials. Although a strength of the present study is that we used for the first time a wide range of measures, which are commonly used in neuropsychological assessment in Greece, nonetheless, the results are constrained by these measures.

Conclusions

While this study only provided data for some of the relationships between euthyroid HT parameters, emotional and cognitive functioning in middle-aged women, it appears that there are remaining gaps in knowledge and issues for further research.

Future research should further investigate if and how HT diagnosis without treatment can cause morphological or functional changes in brain areas related to attention and executive functions that are evident later in the functioning of the patients, and should further clarify if these findings are applicable across age groups and across the two genders, while the reasons for the wide deviation in

neuropsychological scores of HT euthyroid patients which are evident at all examinations, should be also taken into account.

Ethical Approval

All procedures performed in this study involving human participants were in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments.

References

- Beydoun, M. A., Beydoun, H. A., Kitner-Triolo, M. H., Kaufman, J. S., Evans, M. K., & Zonderman, A. B. (2013). Thyroid hormones are associated with cognitive function: moderation by sex, race, and depressive symptoms. *The Journal of Clinical Endocrinology & Metabolism*, *98* (8), 3470-3481.
- Canaris, G. J., Manowitz, N. R., Mayor, G., & Ridgway, E. C. (2000). The Colorado thyroid disease prevalence study. *Archives of Internal Medicine*, *160* (4), 526-534.
- Caturegli, P., De Remigis, A., & Rose, N. R. (2014). Hashimoto thyroiditis: Clinical and diagnostic criteria. *Autoimmunity Reviews*, *13* (4), 391-397.
- Giannouli, V. (2017). Memory and Hashimoto's thyroiditis under levothyroxine treatment: Is there a placebo or nocebo effect? *Endocrine*, *57* (1), 183-186. doi: 10.1007/s12020-017-1262-0.
- Giannouli, V., & Syrmos, N. (2017). Exploring attitudes towards endocrine diseases in Greece. *Endocrine*, *55* (2), 655-657.
- Giannouli, V., Toulis, K. A., & Syrmos, N. (2014). Cognitive function in Hashimoto's thyroiditis under levothyroxine treatment. *Hormones (Athens)*, *13* (3), 430-433.
- Grigorova, M., & Sherwin, B. B. (2012). Thyroid hormones and cognitive functioning in healthy, euthyroid women: A correlational study. *Hormones and Behavior*, *61* (4), 617-622.
- Leyhe, T., & Müssig, K. (2014). Cognitive and affective dysfunctions in autoimmune thyroiditis. *Brain, Behavior, and Immunity*, *41*, 261-266.
- Pasqualetti, G., Pagano, G., Rengo, G., Ferrara, N., & Monzani, F. (2015). Subclinical hypothyroidism and cognitive impairment: systematic review and meta-analysis. *The Journal of Clinical Endocrinology & Metabolism*, *100* (11), 4240-4248.
- Samuels, M. H. (2014). Psychiatric and cognitive manifestations of hypothyroidism. *Current Opinion in Endocrinology, Diabetes, and Obesity*, *21* (5), 377-383.
- Toudou Taouda, M., Samuel Obenda, N., Maazou, L., Camara, D., Assadeck, H., & Souirti, Z. (2016). Neurological manifestations in Hashimoto's Thyroiditis. *IOSR Journal of Dental and Medical Sciences*, *15* (5), 62-68.
- Volpato, S., Guralnik, J. M., Fried, L. P., Remaley, A. T., Cappola, A. R., & Launer, L. J. (2002). Serum thyroxine level and cognitive decline in euthyroid older women. *Neurology*, *58* (7), 1055-1061.
- Wekking, E. M., Appelhof, B. C., Fliers, E., Schene, A. H., Huyser, J., Tijssen, J. G., & Wiersinga, W. M. (2005). Cognitive functioning and well-being in euthyroid patients on thyroxine replacement therapy for primary hypothyroidism. *European Journal of Endocrinology*, *153* (6), 747-753.

Received: October 15, 2017

Accepted: December 24, 2017

Vaitsa Giannouli PhD, Researcher, Bulgarian Academy of Sciences, Bulgaria.
Email: giannouliv@hotmail.com

Nikolaos Syrmos PhD, Department of the Physical Education and Health,
Aristotle University of Thessaloniki, Greece.
E-mail: milanako76@yahoo.gr
