Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology)
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Five-year longitudinal evaluation of quality of life in a cohort of patients with differentiated thyroid carcinoma

Massimo GIUSTI[†1], Giulia MELLE[1], Monica FENOCCHIO[2], Lorenzo MORTARA[1], Francesca CECOLI[1], Valeria CAORSI[1], Diego FERONE[1], Francesco MINUTO[1], Elda RASORE[2]

([1]*Department of Endocrine and Medical Sciences, University of Genoa, I-16100 Genoa, Italy*)

([2]*Department of Psychiatric Sciences, University of Genoa, I-16100 Genoa, Italy*)

[†]E-mail: magius@unige.it

Received Nov. 3, 2010; Revision accepted Jan. 17, 2011; Crosschecked Jan. 28, 2011

Abstract: Differentiated thyroid carcinoma (DTC) generally has a favorable outcome. Thyroid disease, treatments, stress, and comorbidity can compromise health-related quality of life (QoL) and indirectly weigh upon the outcome. From 2004 to 2008, we evaluated QoL longitudinally in 128 DTC subjects. During scheduled examinations, subjects were asked to undergo a semi-structured psychiatric interview and five rated inventories. The same examination was conducted in 219 subjects after surgery for benign thyroid pathology. Low scores represent a better QoL. DTC and control subjects were similar in terms of age, male/female ratio, concomitant psychopharmacological treatments, and frequency of psychiatric diseases. In DTC subjects, Billewicz scale (BS) scores showed an increasing trend over time, especially among females. The ad hoc thyroid questionnaire (TQ) scores were similar in both groups and did not change over time, but at the end of the study ad hoc TQ and BS were significantly related. Ad hoc TQ scores were also related to age on entry to the study. In both male and female DTC subjects, Hamilton's tests for anxiety (HAM-A), but not for depression (HAM-D), showed an improving trend. At the end of the study, HAM-A and HAM-D scores were comparable to those of the control group. HAM-A and HAM-D were both positively correlated with the stage of cancer and the time between diagnosis and treatment. Only HAM-D correlated with age on entry to the study. Kellner symptom questionnaire (KSQ) item scores were higher in DTC subjects than in controls. The change over time in the items including anxiety, somatization, depression, and hostility was significant. Somatization and hostility were more significantly reduced in DTC females than in DTC males. Hostility scores were significantly lower in DTC subjects than in controls at the end of the study. Somatization and depression were significantly related to staging on diagnosis and age on entry to the study. Our study confirms a wide variation of illness perception in DTC subjects, which is generally unrelated to the favorable clinical follow-up of the disease. Psychological evaluation during long-term follow-up improved QoL scores, which reached the same levels noted in subjects with a history of thyroid surgery for benign thyroid pathology. Our data indicate that special attention should be paid to QoL in older DTC subjects and those with more severe staging on diagnosis.

Key words: Thyroid cancer, Outcome, Psychological evaluation, Quality of life, Longitudinal study
doi:10.1631/jzus.B1000382 **Document code:** A **CLC number:** R736.1

1 Introduction

Differentiated thyroid carcinoma (DTC) is generally associated with a favorable prognosis, with the ten-year survival rate reaching 90%–95% (Schlumberger and Torlontano, 2000), but the mortality rate climbs gradually from the age of 40–45 years (Haymart, 2009). Subjects with DTC remain at risk of tumor recurrence/persistence for decades after diagnosis, which highlights the importance of long-term follow-up (Pacini *et al*., 2010; Tala and Tuttle, 2010).

Quality of life (QoL) is defined as an individual's perception of his or her position in life, in the context

of the culture and value systems in which he or she lives and in relation to his or her goals, expectations, standards, and concerns (Crevenna *et al*., 2003). Cancer rehabilitation involves improving psychological symptoms and the perception of restored well-being after or during treatment (Crevenna *et al*., 2003). QoL has been extensively evaluated in DTC subjects independently of the cognitive and job-related disturbances brought about by changes in circulating thyroid hormone levels for diagnostic and therapeutic purposes (Dow *et al*., 1997; Botella-Carretero *et al*., 2003; Golger *et al*., 2003). More recently, the clinical benefits of the use of recombinant human thyroid-stimulating hormone (TSH), versus thyroid hormone withdrawal, on the preservation of QoL have been demonstrated (Schroeder *et al*., 2006; Taïeb *et al*., 2009; Lee *et al*., 2010). On the other hand, few studies have evaluated QoL in DTC subjects on stabilized levo-thyroxine (L-T4) suppressive treatment a long time after primary therapies. Moreover, these studies have been hampered by the small number of subjects involved and of QoL questionnaires administered and by the type of control group chosen (Crevenna *et al*., 2003; Schultz *et al*., 2003; Dagan *et al*., 2004; Tan *et al*., 2007).

In 2005, our group published the results of a study conducted on a cohort of 61 DTC subjects in whom QoL was tested by means of the Kellner symptom questionnaire (KSQ) and the Hamilton depression scale and compared with that of a matched cohort of subjects on L-T4 therapy after surgery for benign thyroid pathology (Giusti *et al.*, 2005). The study revealed that DTC subjects suffered a QoL impairment that correlated positively with age and prognostic risk, as evaluated by means of the metastasis age completeness invasiveness size (MACIS) score, one of the prognostic factors for papillary thyroid carcinoma (Hay *et al*., 1993). The fact that a slight improvement in the item somatization was observed in a sub-set of 31 of the same DTC subjects, retested 8–14 months later, prompted us to hypothesize that QoL improves when longitudinal follow-up comprises psychological evaluation (Giusti *et al.*, 2005).

More recently, two further studies have been published in this field. Hoftijzer *et al*. (2008) used four validated health-related questionnaires to evaluate QoL in a cohort of 153 cured DTC subjects. They recorded significantly lower QoL scores in DTC subjects than in controls, and found that prolonged follow-up in a disease-free condition was an important predictor of improvement in QoL. Using a validated perception questionnaire, Hirsch *et al*. (2009) evaluated illness perception in a study group of 110 consecutive subjects with a history of DTC under routine follow-up. They found that subjects perceived their illness on a subjective and emotional basis and that a lower level of education, short time since the last treatment, disease persistence, and the number of iodine treatments were associated with QoL impairment. A conclusion shared by all these studies is that psychological well-being requires attention, and that a trained psychologist should be included in the multidisciplinary team that manages DTC.

The aim of our five-year study was to evaluate QoL in a cohort of 128 DTC subjects by means of several instruments, both self-rated and physician-administered. Test scores obtained longitudinally were correlated with clinical data. In this study, we used self-rated questionnaires, both constructed ad hoc and previously well-validated in thyroid pathology (Giusti *et al*., 2005), and well-known instruments used extensively by thyroidologists (Billewicz *et al*., 1969) or psychiatrists (Hamilton, 1959; 1960). The objectives of the study were: (1) to confirm the wide variation of illness perception (Crevenna *et al*., 2003; Hirsch *et al*., 2009) in DTC subjects, as already documented in the whole spectrum of thyroid pathologies (Bianchi *et al.*, 2004; Miccoli *et al.*, 2007), which, however, should require a new specific thyroid questionnaire (Watt *et al*., 2010); (2) to verify our previous impression (Giusti *et al*., 2005) that, in unselected DTC, periodic follow-up examinations and the use of less aggressive diagnostic/therapeutic methods (Lee *et al*., 2010) improve health-related QoL over time; (3) to demonstrate that the prolonged follow-up of DTC subjects normalizes QoL to the levels reported in subjects with benign post-surgical thyroid pathology.

2 Subjects and methods

2.1 Subjects

The study group consisted of 128 out-patients aged 20–92 years with a history of DTC who were under routine follow-up in our Endocrine Unit from

2004 to 2008. In 95% of subjects, initial therapy consisted of near-total thyroidectomy; this was followed by one or more cycles of radioiodine in 80%. Histology of DTC was papillary, follicular variant of papillary, follicular, medullary, insular, Hurthle and cancer of unknown malignant potentiality in 89, 17, 9, 8, 2, 2, and 1 subjects, respectively. Tumor staging on diagnosis was available in 98% of subjects. A total of 78% of subjects were at a low-risk stage (stages 1–2), while the remainders were at a high-risk stage (stage 3: 13%; stage 4: 9%). Over the same period, a cohort of 219 subjects aged 24–89 years who had undergone thyroidectomy (lobectomy or near-total thyroidectomy) for benign thyroid pathologies (goiter 48%, hyperthyroidism 33%, indeterminate cytology 17%, suspicious calcitonin levels 2%) was evaluated as a control group. All DTC subjects and 90% of controls were on treatment with L-T4 at the time of the study. Informed consent was obtained yearly from all participants and the study was approved by the institutional ethics committee.

2.2 Methods

At the time of scheduled examination, all subjects were also asked to undergo psychological evaluation. Clinical examination comprised history, evaluation of current therapies, physical examination, neck sonography (AU 5 Idea equipped with linear 7.5 MHz probe; Esaote, Genoa, Italy), and blood tests. Disease-specific morbidity was evaluated by the endocrinologist by means of Billewicz scale (BS) (Billewicz *et al*., 1969) and a self-administered ad hoc thyroid questionnaire (TQ). A semi-structured clinical interview to assess psychiatric status was conducted by a psychiatrist. Psychological well-being was evaluated by means of the Hamilton scales for anxiety (HAM-A) (Hamilton, 1959) and depression (HAM-D) (Hamilton, 1960). All subjects were then asked to complete the Italian version (Fava and Kellner, 1982) of the self-rated KSQ (Kellner, 1971).

2.3 Survey measurements

2.3.1 Billewicz scale

An Italian translation of BS (Billewicz *et al*., 1969) was used to assess disease-specific morbidity (or inadequate L-T4 treatment). Seven symptoms (perspiration, hoarseness, paresthesia, skin dryness, constipation, hearing loss, and weight increase) and five signs (motor slowness of movement, lack of Achilles reflex, rough skin, eyelid edema, and cold skin) were assessed as an observer-rated set of clinical findings that have been used in studies of hypothyroidism. One point was assigned for each symptom or sign found, with a maximum score of 12. In women younger than 55 years, another point was added to the BS score recorded.

2.3.2 Ad hoc thyroid questionnaire

The ad hoc TQ was prepared as a simplified version of the medical outcome study 36-item short form survey (SF-36), which is a well-validated measure of general health status (Ware and Sherbourne, 1992). TQ comprised 39 questions grouped into eight items designed to explore changes in general-health perception, physical function, presence or absence of physical pain, social relations, mental health, osteoarticular pain, tiredness or vitality, and specific symptoms. TQ was scored by assigning one point for each negative reply (no physical or mental changes) and two points for each positive reply (physical or mental changes). The total score ranges from 39 to 78.

2.3.3 Hamilton for anxiety

The HAM-A (Hamilton, 1959) is a rating scale developed to measure the severity of anxiety symptoms. The scale consists of 14 items. Each item represents a cluster of different symptoms (from 2 to 8) which are linked to one another by nature or correlated by clinical experience. Items 7–13 indicate somatic anxiety, whereas items 1–6 and item 14 indicate psychic anxiety. Each item is scored on a scale of 0 to 4 (absent, mild, moderate, severe, very severe—the last one being rarely used), with a total score ranging from 0 to 56.

2.3.4 Hamilton for depression

The HAM-D (Hamilton, 1960) is a rating scale developed to measure the severity of depressive symptoms. HAM-D evaluates 21 items pertaining to the affective field. The presence of depressive symptoms corresponds to scores of up to a maximum of 66. HAM-D had already been validated in subjects who had undergone thyroid surgery for DTC or benign thyroid pathology (Botella-Carretero *et al*., 2003).

2.3.5 Kellner symptom questionnaire

The self-rated KSQ had already been used by us (Giusti *et al*., 2005) and others (Lee *et al*., 2010) in similar studies of DTC. KSQ comprises eight subscales. The items anxiety, depression, somatization, and hostility evaluate the degree of psychological discomfort or lack of well-being on a numerical scale ranging from 0 to 23 for each item. Higher scores indicate a lower OoL.

2.3.6 Mini-mental state

The mini-mental state examination (MMSE) was used to evaluate cognitive function during the five-year study. Normative data from the Italian population were used (Measso *et al*., 1991). Subjects with MMSE scores less than 19 (moderate or severe memory impairment) were not included in the study groups or dropped out during the study.

2.4 Laboratory evaluations

Serum TSH, free-T3 (f-T3), free-T4 (f-T4), thyroglobulin (Tg), and Tg antibodies (TgAb) were evaluated. TSH, f-T3, f-T4, and Tg were measured by means of ultra-sensitive chemiluminescence immunoassay (Roche Diagnostics, Mannheim, Germany). Normal ranges are: 0.3–4.2 mU/L for TSH, and 3.9–6.8 pmol/L and 12.0–22.0 pmol/L for f-T3 and f-T4, respectively. Normal Tg values in patients without goiter are <90 μg/L. On the basis of the functional sensitivity of the assay, in TgAb-negative patients, a Tg level <0.5 μg/L, both after thyroid ablation under L-T4 and after recombinant human-TSH, was our institutional cut-off to discriminate undetectable from detectable Tg levels. TgAb was evaluated by means of the DiaSorin assay (Saluggia, Italy); concentrations <100 mU/L were regarded as negative. In patients with medullary thyroid carcinoma (MTC), serum calcitonin (CT) was assayed by chemiluminescence immunoassay (DiaSorin); in our laboratory, the normal CT range is <10 ng/L.

2.5 Statistical analysis

The best predictor of cure in DTC patients was considered to be an undetectable Tg level combined with negative neck sonography at least one year after ablative therapies. In patients with a history of MTC, a normal baseline CT level, combined with negative imaging, was considered indicative of a disease-free condition. DTC patients studied shortly after the primary therapies (follow-up <1 year) in whom Tg levels were negative under L-T4 were regarded as probably cured. DTC patients without a history of radioiodine ablative treatment in whom Tg levels were <1 μg/L and no changes in sonographic findings over time had been recorded were regarded as cured.

The self-rated psychometric tests were deemed assessable when filled in >95% correctly.

GraphPad Prism for Windows (Version 4.0; GraphPad Software, San Diego, CA, USA) was used for all analyses. Data are expressed as mean±standard error of mean (SEM) unless indicated otherwise. To evaluate changes in experimental parameters in DTC subjects during the five-year study, the non-parametric Kruskal-Wallis analysis of variance (ANOVA), followed by Dunn's Multiple Comparison test, was used. To compare absolute and percentage data between groups, Mann-Whitney and Chi-square tests were used. Correlation analyses between variables were carried out by Spearman correlation. Data below the functional sensitivity of the assay were analyzed for statistical purposes by using the functional sensitivity value. Significance was taken as $P\leq0.05$; exact P values ranging from 0.05 to 0.0001 are given.

3 Results

3.1 Clinical data

At the time of the first study examination, the DTC group was matched for age and male/female ratio with the control group (Table 1). The median time from surgical therapy to study entry was one year in DTC subjects and three years in controls. Adiposity was a more frequent finding in DTC subjects than in controls, owing to the significantly greater number of DTC subjects with body mass index (BMI) over 35 kg/m^2 (DTC 11%, controls 1%; $P<0.0001$). Owing to the TSH-suppressive L-T4 dosage, free thyroid hormones were higher in DTC subjects than in controls. During the five-year study, L-T4 dosage was reduced, increased, and unchanged in 38%, 31%, and 31% of DTC subjects, respectively. At the time of the last examination, L-T4 dosage was unchanged on average [(872.2±17.9) μg/week, median 875 μg/week, range 525–1896 μg/week] in

Table 1 Some demographic and clinical data in DTC and control subjects

Parameter	Value*		*P*
	DTC group (*n*=128)	Control group (*n*=219)	
Age (year)	54.3±14.6	52.2±12.6	NS
Male/female ratio	26/102	33/186	NS
BMI (kg/m^2)	26.8±0.5	25.1±0.2	0.02
Time since surgery (year)	4.3±0.5	5.9±0.6	0.05
L-T4 dosage (μg/week)	864.7±15.9	674.1±15.1	<0.0001
Use of psychotropic drugs (%)	32	26	NS
TSH (mU/L)	0.80±0.20	2.29±0.27	<0.0001
f-T3 (pmol/L)	4.8±0.1	4.2±0.1	<0.0001
f-T4 (pmol/L)	20.4±0.4	17.5±0.3	<0.0001
Second cancer (%)	9	3	0.01
Major psychopathology (%)	5	3	NS

Values are expressed as mean±SEM or percentage, except values for age (mean±SD). NS: not significant

DTC subjects. During the study period, TSH was profoundly suppressed (<0.10 mU/L), below the normal range (0.10–0.29 mU/L), in the low-normal range (0.30–0.99 mU/L) or inadequately controlled (>1.00 mU/L) in 39%, 32%, 15%, and 14% of DTC subjects, respectively. A second cancer was reported more frequently in DTC subjects than in controls (Table 1); in both groups, breast cancer was the most frequent second cancer.

Many DTC (32%) and control (26%) subjects were taking psychotropic drugs for insomnia and anxiety, the most frequent being benzodiazepines. Psychoanaleptic drugs were also used in DTC (11%; as a continuous therapy in 43% of these subjects) and control (10%) subjects. A generalized anxiety disorder was diagnosed in 23% of DTC subjects and 29% of controls. A cognitive disorder was identified during the study in three DTC subjects, who were excluded from further evaluations. The main psychiatric disorders in DTC subjects were anorexia nervosa (*n*=3), major depression (*n*=2), panic disorder (*n*=1), and hallucinatory disorder (*n*=1). In the control group, major depression was found in six subjects (Table 1).

During the study, three DTC subjects died of papillary thyroid cancer (*n*=1), follicular thyroid cancer (*n*=1), and concomitant laryngeal cancer (*n*=1) after one, two, and three years, respectively. Assessment of the evolution of thyroid disease was incomplete in seven subjects, owing to poor compliance with scheduled visits. A disease-free condition was diagnosed in 80% of the 110 subjects with DTC of follicular origin, and in 50% of subjects with MTC. DTC persisted or recurred in 11 subjects. A judgment of probable freedom from disease was expressed on the basis of clinical evolution over time in some DTC subjects with positive TgAb levels (*n*=3), incomplete surgery (*n*=2), and unperformed radioiodine ablation (*n*=10).

3.2 Quality of life

QoL evaluation was stopped because of cognitive disorders in three DTC subjects in the 1st and 3rd years of the study. One subject could not complete the self-rated questionnaires owing to illiteracy.

In the control group, the average BS score was 3.2±0.1 (*n*=153, median 3.0, range 0–8). BS data were available on 95% of DTC subjects. On entry to the study, BS scores were lower in DTC subjects (*n*=116, 2.5±0.1, median 2.0, range 0–9) than in controls (*P*=0.001), but increased significantly over five years, as shown by Kruskal-Wallis ANOVA (*P*<0.0001). This increase was significant in female (*P*<0.001), but not in male, DTC subjects. At the last examination, the BS scores were seen to have increased from the baseline in the whole DTC group (3.3±0.2, *P*<0.001 by Dunn's test, median 3.0, range 0–8) and no longer differed from that recorded in controls.

The ad hoc TQ score in the control group was 51.2±0.7 (*n*=116, median 50.5, range 40–71). Ad hoc TQ data were available on 78% of DTC subjects. There was no significant change in the self-rated ad hoc TQ score in the DTC group, either as a whole (*n*=94) or divided according to sex, during the study period. Neither baseline (51.4±0.9, median 50, range 34–80) nor end-study (52.2±0.9, median 52, range 33–72) ad hoc TQ scores differed significantly from those of the control group. Among several clinical variables, the only significant (*P*=0.03) correlation was between ad hoc TQ scores and age on entry to the study (Table 2). Ad hoc TQ scores were also significantly correlated with BS scores at the end of the study (Fig. 1) but not on study entry.

In the control group, HAM-A and HAM-D scores were 16.9±1.1 (*n*=73, median 15, range 3–48) and 31.2±0.9 (*n*=83, median 30, range 9–56), respectively. In DTC subjects, HAM-A and HAM-D scores were

Table 2 BS and ad hoc TQ correlations with clinical parameters in DTC patients

Parameter	*n*	*Sr*	*P*
Staging on diagnosis			
BS	119	−0.06	NS
ad hoc TQ	94	0.06	NS
Age on study entry			
BS	120	−0.10	NS
ad hoc TQ	94	0.22	0.03
Time from diagnosis to study			
BS	119	−0.06	NS
ad hoc TQ	94	−0.03	NS
TSH levels on study entry			
BS	119	0.04	NS
ad hoc TQ	93	0.00	NS
L-T4 dosage on study entry			
BS	118	0.02	NS
ad hoc TQ	92	0.04	NS
Tg levels on study entry			
BS	111	0.07	NS
ad hoc TQ	88	−0.04	NS

n: number of pairs; *Sr*: Spearman coefficient; *P*: significance; NS: not significant

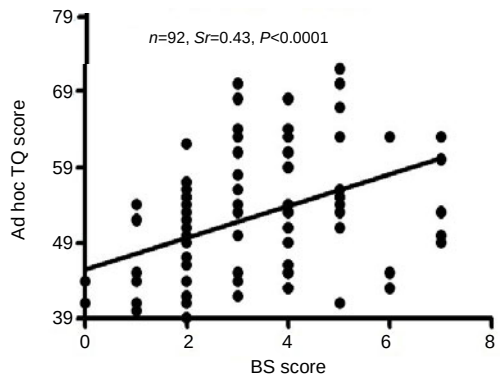

Fig. 1 Correlation between BS and ad hoc TQ scores at the end of the study in DTC patients
n: number of pairs; *Sr*: Spearman coefficient; *P*: significance

fully available during the five-year study period in 62% and 57% of subjects, respectively. During the five-year study, a significant (*P*=0.002) improvement in HAM-A scores, but not in HAM-D scores, was observed in DTC subjects on ANOVA. The improvement in HAM-A scores was noted both in male (*P*=0.006) and female (*P*=0.03) DTC subjects. However, in DTC subjects, both rating scales on entry to the study (HAM-A: *n*=99, 23.7±1.5, median 20, range 3–56; HAM-D: *n*=82, 34.6±0.8, median 34, range 22–51), but not at the end of the study (HAM-A: *n*=84, 15.0±1.0, median 12, range 0–35; HAM-D: *n*=90, 31.5±0.7, median 30, range 22–49), were significantly (HAM-A: *P*=0.005; HAM-D: *P*=0.003) higher than in controls. Table 3 reports the correlations observed between HAM-A/HAM-D and clinical parameters in DTC subjects. Scores on both tests showed a significant positive correlation with staging on diagnosis and time from diagnosis to study, and also a significant negative correlation with TSH levels. Only the HAM-D score correlated significantly with age (Table 3).

Table 3 Correlations between HAM-A/HAM-D scores and clinical parameters in DTC patients

Parameter	*n*	*Sr*	*P*
Staging on diagnosis			
HAM-A	98	0.28	0.006
HAM-D	81	0.25	0.02
Age on study entry			
HAM-A	99	0.19	NS
HAM-D	82	0.42	<0.0001
Time from diagnosis to study			
HAM-A	99	0.31	0.002
HAM-D	82	0.30	0.006
TSH levels on study entry			
HAM-A	99	−0.28	0.004
HAM-D	78	−0.25	0.03
L-T4 dosage on study entry			
HAM-A	99	−0.06	NS
HAM-D	81	0.00	NS
Tg levels on study entry			
HAM-A	96	0.08	NS
HAM-D	78	−0.03	NS

n: number of pairs; *Sr*: Spearman coefficient; *P*: significance; NS: not significant

In the control group (*n*=159), the overall scores of the items on the self-rated KSQ were: anxiety 6.6±0.3 (median 6, range 0–16), depression 5.4±0.3 (median 5, range 0–21), somatization 7.9±0.4 (median 7, range 0–22), and hostility 4.0±0.3 (median 3, range 0–22). Changes in KSQ over the five-year period were assessable in 76% of DTC subjects. Fig. 2 shows the trend in KSQ items in DTC subjects over the five years of the study. On ANOVA, significant changes over time in the items anxiety (*P*=0.006), somatization (*P*=0.02), depression (*P*< 0.05), and hostility (*P*<0.0001) were found in the

whole group of DTC subjects. Significant changes in the item anxiety were observed in both males (P=0.04) and females (P=0.04). The items somatization (P=0.05) and hostility (P<0.0001) were significantly reduced over the five-year period in females but not males, while the changes over time in the item depression were not significant when DTC subjects were divided according to sex. At the baseline, except for the item anxiety (8.2±0.7, median 7, range 1–20), all items were significantly higher in DTC subjects (somatization: 9.8±0.7, median 9, range 1–20, P=0.03; depression: 7.1±0.6, median 6, range 0–20, P=0.02; hostility: 6.2±0.6, median 5, range 0–18, P=0.002) than in controls. At the end of the study, the item hostility was lower (P=0.03) in DTC subjects (3.3±0.4, median 2, range 0–17) than in controls, while the other items were similar in both groups of subjects (Fig. 2). Among the clinical parameters considered, the items somatization and depression showed a significant correlation with the DTC stage on diagnosis and age on entry to the study (Table 4).

Table 4 Correlations between KSQ items and clinical parameters in DTC patients

Parameter	KSQ item	*Sr*	*P*
Staging on diagnosis (n=97)	Anxiety	0.14	NS
	Somatization	0.30	0.003
	Depression	0.24	0.02
	Hostility	0.14	NS
Age on study entry (n=98)	Anxiety	0.07	NS
	Somatization	0.34	0.0006
	Depression	0.32	0.001
	Hostility	0.06	NS
Time from diagnosis to study (n=98)	Anxiety	0.01	NS
	Somatization	−0.01	NS
	Depression	0.10	NS
	Hostility	0.07	NS
TSH levels on study entry (n=97)	Anxiety	−0.10	NS
	Somatization	−0.14	NS
	Depression	−0.16	NS
	Hostility	−0.03	NS
L-T4 dosage on study entry (n=95)	Anxiety	0.00	NS
	Somatization	0.00	NS
	Depression	−0.05	NS
	Hostility	0.00	NS
Tg levels on study entry (n=94)	Anxiety	−0.06	NS
	Somatization	0.17	NS
	Depression	0.07	NS
	Hostility	−0.01	NS

n: number of pairs; *Sr*: Spearman coefficient; *P*: significance; NS: not significant

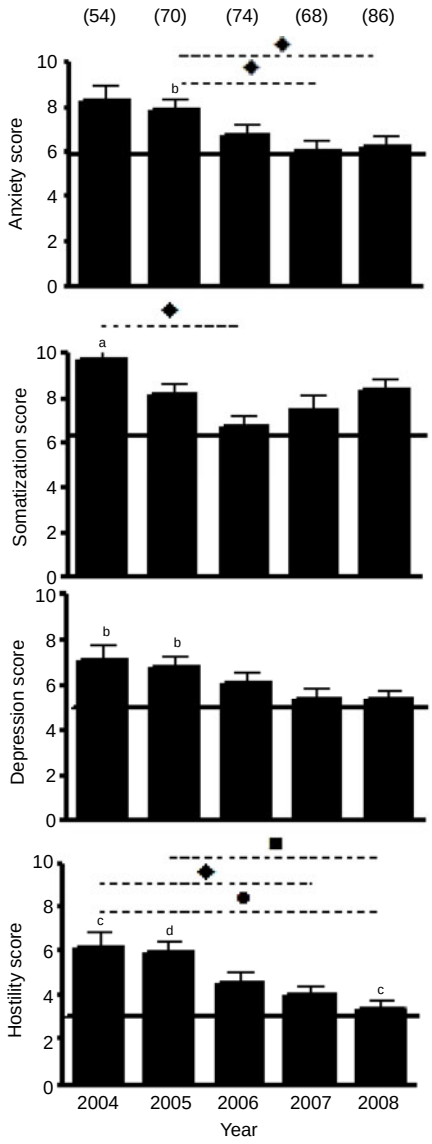

Fig. 2 Scores of KSQ items anxiety, somatization, depression, and hostility observed during the five-year study period in DTC patients

The number of self-rated KSQ tests assessable each year is reported in brackets. The median value observed in the control group is indicated by the continuous horizontal line. Post-ANOVA significance values on Dunn's multiple comparison test are reported above the dotted line: solid rhomb P<0.05, solid circle P<0.01, solid square P<0.001. Data significantly different from control group data on Mann-Whitney test are indicated: [a] P=0.03, [b] P=0.02, [c] P=0.002, [d] P=0.001. Data are expressed as mean±SEM

4 Discussion

Thyroid disorders are very common in the general population, but few subjects are overtly symptomatic. In these subjects, therapeutic decisions are

usually based much more on hormone levels than on subjects' perceptions of their health status. In an Italian population with thyroid disorders, Bianchi *et al.* (2004) used the SF-36 inventory and Nottingham health profile to evaluate health-related QoL; this was found to be poor regardless of the severity and functional type of thyroid dysfunction, as already shown by other studies (Jaeschke *et al*., 1994; Ljunggren *et al*., 1998; Biondi *et al*., 2000). In a cohort of patients scheduled for thyroid surgery for several reasons, including cancer in 17% of cases, Miccoli *et al.* (2007) assessed QoL by means of the SF-36 inventory, but not the Symptom Checklist-90; they found poorer overall scores and social functioning than in either the general population or patients scheduled for non-thyroidal surgery. Thyroidectomy was followed by a significant reduction in the severity of psychiatric symptoms, but social functioning remained markedly impaired for a long time, even after euthyroidism had been achieved (Miccoli *et al*., 2007).

One obstacle to evaluating illness perception in patients with thyroid cancer is the current lack of developed questionnaires (Watt *et al*., 2010) like that used for thyroid-associated ophthalmopathy (Terwee *et al*., 2001). In 2004, we began our study by using well-validated inventories (Hamilton, 1959; 1960; Billewicz *et al*., 1969; Kellner, 1971) which we had previously used in part of DTC subjects (Giusti *et al*., 2005). Moreover, the HAM-D and the modified version of the self-rated KSQ have recently been used by other authors (Lee *et al.*, 2010). In addition, at the end of our study, the ad hoc TQ, which had been prepared to measure general health status, was found to correlate positively with the BS used by us and others (Schroeder *et al*., 2006) to assess disease-specific morbidity. We therefore believe that the results of our study are valid, even though they will need to be confirmed by means of newly validated and reliable DTC-specific QoL inventories.

DTC generally has a good prognosis, with a long-term survival rate of 90% (Schlumberger and Torlontano, 2000; Pacini *et al*., 2010; Tala and Tuttle, 2010). In our study, five-year disease-specific mortality was less than 2% and the rate of persistence/recurrence was 9%; these rates seem to be somewhat better than those reported by Mazzaferri and Kloss (2001) in a very large DTC population studied in the former time, before sonography was widely used and when the diagnostic or therapeutic use of recombinant TSH (rhTSH) was still a long way off. For a long time after diagnosis and primary therapies, DTC survivors require medical and psychological follow-up to prevent the broad spectrum of symptoms that can impair well-being. From this perspective, psychological examination conducted by means of inventories and semi-structured interviews should be given the same importance as all the other medical examinations that are usually performed over the long term to establish or confirm DTC prognosis (Hirsch *et al*., 2009).

On analyzing 518 replies from thyroid cancer survivors from among 7000 surveys mailed to subjects who had been diagnosed with cancer, Schultz *et al*. (2003) found that DTC survivors more frequently reported that their cancer had effected their overall health than did survivors of any other cancer type; fairly severe psychological problems, memory loss, and migraine were recorded. In our pilot study (Giusti *et al*., 2005) in which QoL was evaluated in 61 DTC subjects by means of a semi-structured interview, HAM-D and KSQ, we found a greater perception of illness among DTC subjects than among 64 control subjects with a history of thyroidectomy for non-malignant thyroid pathology. In the same study, it emerged that elderly DTC subjects, especially if female, should be followed up more carefully, as they experienced the greatest impairment of well-being (Giusti *et al*., 2005). In a sub-set of 31 DTC subjects, we also observed that illness perception, and consequently QoL, improved after a second psychological examination conducted 8–14 months later. This observation prompted us to continue the study longitudinally, utilizing more inventories to cover a broad spectrum of psychological and physical symptoms.

Patient compliance with yearly psychological evaluation was fairly good. The main problems encountered were incompleteness of the self-rated questionnaires and, sometimes, refusal to wait for the semi-structured interview. Nevertheless, in our opinion, the size of the experimental sample and the length of the study allowed us to obtain valid results. The most evident phenomenon was the significant decreasing trend in the scores of almost all psychological evaluations, which declined to control group levels. Anxiety, depression, and hostility were the items which improved most over time, indicating a reduction in illness perception.

Some somatic aspects were less sensitive to yearly examinations. In fact, BS showed a slight increase over the five-year period. This may have been the result of an increased perception of symptoms due to greater attention by the physician or patient, and may be unrelated to hormone therapy, which was generally long-term TSH-suppressive. Moreover, we cannot exclude the possibility that a major comorbidity, obesity, or a second cancer may have aggravated symptoms independently of DTC. The self-rated ad hoc TQ did not display changes over time, with scores being significantly correlated only with age on entry to the study. Age is a DTC-independent indicator of well-being, and influences scores on some inventories more than other clinical parameters.

More recently, Hoftijzer *et al*. (2008) used four inventories, including the SF-36, to investigate the well-being of 153 cured DTC subjects. They found a significant decrease in QoL on several subscales in comparison with two control groups comprising 113 healthy subjects recruited directly by DTC subjects themselves and 336 healthy subjects enrolled in other QoL studies, matched for age, gender, and socio-economic status. They concluded that cured DTC subjects have impaired QoL, independently of TSH levels. QoL parameters were inversely affected by the duration of cure. In their paper, Hoftijzer *et al*. (2008) hypothesized that QoL parameters normalized in DTC after prolonged follow-up. This phenomenon is confirmed by our study, since both Hamilton and KSQ scores improved over five years in a population of DTC subjects, about 90% of whom were in a disease-free condition.

In our study, the HAM-A and HAM-D scores and the KSQ items somatization and depression were significantly related to staging of the disease; HAM-D and the items somatization and depression were also significantly linked to the age of the subject on entry to the study. Finally, both HAM-A and HAM-D scores were positively related to the time from diagnosis to study and negatively related to TSH levels on entry to the study. More recently, Hirsch *et al*. (2009) examined 110 consecutive DTC subjects by means of the self-administered illness perception questionnaire revised (IPQ-R) during routine follow-up and did not find any correlation between illness perception and its actual severity. In that study, a negative disease perception, in other words a low QoL, proved to be associated with time since last treatment, evidence of disease persistence, and the number of iodine treatments administered (Hirsch *et al*., 2009). On evaluating 150 consecutive subjects with DTC by means of the SF-36 questionnaire, Crevenna *et al*. (2003) reported a direct correlation between QoL and disease duration, in which the impairment of QoL noted in the first year after diagnosis partially improved thereafter. Their study found that physical functioning, bodily pain, and role-emotional scales were affected less in male than in female DTC subjects. Gender (Crevenna *et al*., 2003; Giusti *et al*., 2005), age (Giusti *et al*., 2005; Tan *et al*., 2007), and educational level (Tan *et al*., 2007; Hirsch *et al*., 2009) seem to modulate illness perception in DTC subjects.

Moderate hyperthyroxinemia is a common finding in patients with DTC of follicular origin, which requires TSH-suppression. Thereafter, TSH levels can be considered an indirect indicator of adequate L-T4 dosages. Both subclinical and overt hyperthyroidism (Bianchi *et al*., 2004) as well as exogenous TSH-suppressive therapy (Biondi *et al*., 2000) are, however, conditions associated with QoL impairment. In our study, only the initial evaluation by means of the Hamilton tests was negatively related to TSH levels, indicating that an effect on psychological functioning exists, but that the exogenous hormone exerts less effect on QoL than other factors do (e.g., continual medical examination, stability of disease, time from diagnosis). Moreover, L-T4 therapy did not change significantly over the five-year period and, in our study, as well as in that of Hoftijzer *et al*. (2008), several others tests were not related directly or indirectly to L-T4 therapy. In DTC subjects of Hoftijzer *et al*. (2008) and the present study, psychological impairment is less marked and less long-lasting than that produced by subclinical hyperthyroidism (Bianchi *et al*., 2004).

Our study was not designed to evaluate the effect of radioiodine ablation on QoL, and only a few DTC subjects were examined shortly after this procedure, when TSH levels are still uncontrolled and transient subclinical hypothyroidism could impair QoL scores. We therefore assumed that higher scores in DTC subjects than in the control group were not linked to this treatment. However, in 154 DTC subjects evaluated by means of the Portuguese version of the University of Washington QoL questionnaire from four

months to ten years after radioiodine ablation, Almeida *et al*. (2009) reported a relationship between QoL and therapy when activities above 5.55 GBq were administered. On the other hand, in a previous study involving the same inventory test, Dagan *et al.* (2004) found that age was more important than treatments in determining answers to the University of Washington QoL questionnaire.

In conclusion, our study confirms, in agreement with previous researches (Crevenna *et al*., 2003; Schultz *et al*., 2003; Giusti *et al*., 2005; Hoftijzer *et al*., 2008; Hirsch *et al.*, 2009), a wide variation in illness perception in DTC subjects, which is generally unrelated to the favorable clinical follow-up of the disease. Psychometric measurement of QoL by means of several inventories documented higher scores at the baseline in our DTC subjects than in controls; these were partly correlated with staging on diagnosis and age on entry to the study. Our study shows that full compliance with repeated measurements and psychiatric interviews is not easy to obtain. However, longitudinal evaluations in DTC subjects over a five-year period strongly indicated an improvement in the majority of tests and items. Moreover, this is the first study of DTC subjects which confirms previous advice (Giusti *et al*., 2005; Hirsch *et al.*, 2009) to introduce psychological evaluation into long-term follow-up in order to improve QoL. Although most of our DTC subjects were female, some gender differences seem to emerge from the study, such as higher scores in somatic symptoms and signs rated by means of the BS in DTC females but a better trend in the psychological items hostility and somatization on the self-rated KSQ. The long-lasting follow-up of DTC subjects improved psychometric QoL test scores, which reached the same levels as those recorded in the group of subjects with a history of thyroid surgery and under L-T4 therapy for benign thyroid pathology, regardless of the different intervals from surgery and differences in the amount of L-T4 administered and in some clinical characteristics between the groups. Reports in the literature recommend that special attention should be paid to QoL in DTC subjects heavily treated with radioiodine (Almeida *et al*., 2009; Hirsch *et al.*, 2009) and with a more recent history of DTC (Hoftijzer *et al*., 2008). Our data indicate that older DTC subjects and those with more severe staging on diagnosis also need such attention. A future prospective study using newly validated tests for thyroid cancer will be useful in order to confirm our present data.

References

Almeida, J., Vartanian, J.G., Kowalshi, L.P., 2009. Clinical predictors of quality of life in patients with initial differentiated thyroid cancers. *Arch. Otolaryngol. Head Neck Surg.*, **135**(4):342-346. [doi:10.1001/archoto.2009.16]

Bianchi, G.P., Zaccheroni, V., Solaroli, E., Viscini, F., Cerutti, R., Zoli, M., Marchesini, G., 2004. Health-related quality of life in patients with thyroid disorders. *Qual. Life Res.*, **13**(1):45-54. [doi:10.1023/B:QURE.0000015315.35184.66]

Billewicz, W.Z., Chapman, R.S., Crooks, J., Day, M.E., Gossage, J., Wayne, E., Young, J.A., 1969. Statistical methods applied to the diagnosis of hypothyroidism. *Q. J. Med.*, **38**(150):255-266.

Biondi, B., Palmieri, E.A., Fazio, S., Cosco, C., Nocera, M., Saccà, L., Filetti, S., Lombardi, G., Perticone, F., 2000. Endogenous subclinical hyperthyroidism affects quality of life and cardiac morphology and function in young and middle-aged patients. *J. Clin. Endocrinol. Metab.*, **85**(12): 4701-4705. [doi:10.1210/jc.85.12.4701]

Botella-Carretero, J.I., Galan, J.M., Caballlero, C., Sancho, C., Escobar-Morreale, H.F., 2003. Quality of life and psychometric functionality in patients with differentiated thyroid carcinoma. *Endocr. Relat. Cancer*, **10(**4):601-610. [doi:10.1677/erc.0.0100601]

Crevenna, R., Zettinig, G., Keilani, M., Posch, M., Schmidinger, M., Pirich, C., Nuhr, M., Wolzt, M., Quittan, M., Fialka-Moser, V., *et al.*, 2003. Quality of life in patients with non-metastatic differentiated thyroid cancer under thyroxine supplementation therapy. *Support. Care Cancer*, **11**(9):597-603. [doi:10.1007/s00520-003-0474-4]

Dagan, T., Bedrin, L., Horowitz, Z., Chaushu, G., Wolf, M., Kroenberg, J., Talmi, Y.P., 2004. Quality of life of well-differentiated thyroid carcinoma patients. *J. Laryngol. Otol.*, **118**(7):537-542. [doi:10.1258/0022215041615137]

Dow, K.H., Ferrell, B.R., Anello, C., 1997. Quality-of-life changes in patients with thyroid cancer after withdrawal of thyroid hormone therapy. *Thyroid*, **7**(4):613-619. [doi:10.1089/thy.1997.7.613]

Fava, G.A., Kellner, R., 1982. Versione Italiana del Symptoms Questionnaire di Kellner. *In*: Canestrari, R. (Ed.), Nuovi metodi in Psicometria. Organizzazioni Speciali, Firenze, p.51-64 (in Italian).

Giusti, M., Sibilla, F., Cappi, C., Dellepiane, M., Tombesi, F., Ceresola, E., Augeri, C., Rasore, E., Minuto, F., 2005. A case-controlled study of quality of life in a cohort of patients with history of differentiated thyroid carcinoma. *J. Endocrinol. Invest.*, **28**(7):599-608.

Golger, A., Fridman, T.R., Eski, S., Witterick, I.J., Freeman, J.L., Walfish, P.G., 2003. Three-week thyroxine withdrawal thyroglobulin stimulating screening test to detect low-risk residual/recurrent well-differentiated thyroid

carcinoma. *J. Endocrinol. Invest.*, **26**(10):1023-1031.

Hamilton, M., 1959. The assessment of anxiety states by rating. *Br. J. Med. Psychol.*, **32**(1):50-55.

Hamilton, M., 1960. A rating scale for depression. *J. Neurol. Neurosurg. Psychiat.*, **23**(1):56-59.

Hay, I., Bergstralh, E., Goellner, J., Ebersold, J., Grant, C., 1993. Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at one institution during 1940 through 1989. *Surgery*, **114**(6): 1050-1058.

Haymart, M.R., 2009. Understanding the relationship between age and thyroid cancer. *The Oncologist*, **14**(3):216-221. [doi:10.1634/theoncologist.2008-0194]

Hirsch, D., Ginat, M., Levy, S., Banbassat, C., Weinstein, R., Tsvetov, G., Singer, J., Shraga-Slutzky, I., Grozinski-Glasberg, S., Mansiterski, Y., *et al.*, 2009. Illness perception in patients with differentiated epithelial cell thyroid cancer. *Thyroid*, **19**(5):459-465. [doi:10.1089/thy.2008.0360]

Hoftijzer, H.C., Heemstra, K.A., Corssmit, E.P., van der Klaauw, A.A., Romijn, J.A., Smit, J.W.A., 2008. Quality of life in cured patients with differentiated thyroid carcinoma. *J. Clin. Endocrinol. Metab.*, **93**(1):200-203. [doi:10.1210/jc.2007-1203]

Jaeschke, R., Guyatt, G., Cook, D., Harper, S., Gerstein, H.C., 1994. Spectrum of quality of life in hypothyroidism. *Qual. Life Res.*, **3**(5):323-327. [doi:10.1007/BF00451724]

Kellner, R., 1971. Part 1. Improvement criteria drug trial with neurotic patients. *Psychol. Med.*, **1**(5):416-421. [doi:10.1017/S0033291700044809]

Lee, J., Yun, M.J., Nam, K.H., Choung, W.Y., Soh, E.Y., Park, C.S., 2010. Quality of life and effectiveness comparisons of thyroxine withdrawal, triiodothyronine withdrawal, and recombinant thyroid-stimulating hormone administration for low-dose radioiodine remnant ablation of differentiated thyroid carcinoma. *Thyroid*, **20**(2):173-179. [doi:10.1089/thy.2009.0187]

Ljunggren, J.G., Torring, O., Wallin, G., Taube, A., Tallstedt, L., Hamberger, B., Lundell, G., 1998. Quality of life aspects and costs in treatment of Grave's hyperthyroidism with anti-thyroid drugs, surgery, or radioiodine: results from a prospective, randomized study. *Thyroid*, **8**(8): 653-659. [doi:10.1089/thy.1998.8.653]

Mazzaferri, E.L., Kloss, R.T., 2001. Clinical review 128: current approaches to primary therapy for papillary and follicular thyroid cancer. *J. Clin. Endocrinol. Metab.*, **86**(4):1447-1463. [doi:10.1210/jc.86.4.1447]

Measso, G., Cavarzeran, F., Zappalà, G., Lebowitz, B.D., Crook, T.H., Pirozzolo, F.J., Amaducci, L.A., Massari, D., Grigoletto, F., 1991. The mini-mental state examination: normative study of an Italian random sample. *Dev. Neuropsychol.*, **9**(2):77-85. [doi:10.1080/87565649109540545]

Miccoli, P., Minuto, M.N., Saggini, R., Rucci, P., Oppo, A., Donatini, G., Goliam, F., Novelli, L., Carlini, M., Dell'Osso, L., 2007. The impact of thyroidectomy on psychiatric symptoms and quality of life. *J. Endocrinol. Invest.*, **30**(10):853-859.

Pacini, F., Castagna, M.G., Cipri, C., Schlumberger, M., 2010. Medullary thyroid carcinoma. *Clin. Oncol.*, **22**(6):475-485. [doi:10.1016/j.clon.2010.05.002]

Schlumberger, M.J., Torlontano, M., 2000. Papillary and follicular thyroid carcinoma. *Best Pract. Res. Clin. Endocrinol. Metab.*, **14**(4):601-613. [doi:10.1053/beem.2000.0105]

Schroeder, P.R., Haugen, B.R., Pacini, F., Reiners, C.R., Schlumberger, M., Sherman, S.I., Cooper, D.S., Schuff, K.G., Braverman, L.E., Skarulis, M.C., *et al*., 2006. A comparison of short-term changes in health-related quality of life in thyroid carcinoma patients undergoing diagnostic evaluation with recombinant human thyrotropin compared with thyroid hormone withdrawal. *J. Clin. Endocrinol. Metab.*, **91**(3):878-884. [doi:10.1210/jc.2005-2064]

Schultz, P.N., Stava, C., Vassiloppoulou-Sellin, R., 2003. Health profiles and quality of life of 518 survivors of thyroid cancer. *Head Neck*, **25**(5):349-356. [doi:10.1002/hed.10217]

Taïeb, D., Sebag, F., Chrenko, M., Baumstarck-Barrau, K., Fortanier, C., Farman-Ara, B., de Micco, C., Vaillant, J., Thomas, S., Conte-Devolx, B., *et al*., 2009. Quality of life changes and clinical outcomes in thyroid cancer patients undergoing radioiodine remnant ablation (RRA) with recombinant TSH (rhTSH): a randomized controlled study. *Clin. Endocrinol.*, **71**(1):115-123. [doi:10.1111/j.1365-2265.2008.03424.x]

Tala, H., Tuttle, R.M., 2010. Contemporary post surgical management of differentiated thyroid carcinoma. *Clin. Oncol.*, **22**(6):419-429. [doi:10.1016/j.clon.2010.04.005]

Tan, L.G., Nan, L., Thumboo, J., Sundram, F., Tan, L.K., 2007. Health-related quality of life in thyroid cancer survivors. *Laryngoscope*, **117**(3):507-510. [doi:10.1097/MLG.0b013e31802e3739]

Terwee, C.B., Dekker, F.W., Mourits, M.P., Gerding, M.N., Baldeschi, L., Kalmann, R., Prummel, M.F., Wiersinga, W.M., 2001. Interpretation and validity of changes in scores on the Grave's ophthalmopathy quality of life questionnaire (GO-QOL) after different treatments. *Clin. Endocrinol.*, **54**(3):391-398. [doi:10.1046/j.1365-2265.2001.01241.x]

Ware, J.E., Sherbourne, C.D., 1992. The MOS 36-item short-form healthy survey (SF-36). I. Conceptual framework and item selection. *Med. Care*, **30**(6):473-483. [doi:10.1097/00005650-199206000-00002]

Watt, T., Hegedus, L., Groenvold, M., Bjornerm, J.B., Rasmussen, A.K., Bonemma, S.J., Feldt-Rasmussen, U., 2010. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *Eur. J. Endocrinol.*, **162**(1):161-167. [doi:10.1530/EJE-09-0521]