Journal of Nutritional & Environmental Medicine (2000) 10, 115–124

CLINICAL RESEARCH

Clinical Response to Thyroxine Sodium in Clinically Hypothyroid but Biochemically Euthyroid Patients

G. R. B. SKINNER MD DSc FRCPath FRCOG, D. HOLMES, A. AHMAD PhD, J. A. DAVIES BSc and J. BENITEZ MSc

Vaccine Research Trust, 22 Alcester Road, Moseley, Birmingham B13 8BE, UK

Abstract

Purpose: *To examine clinical response to thyroid replacement therapy in patients considered to be clinically hypothyroid but with normal thyroid biochemistry.*

Design: *Practice-based open intervention study; control group used for baseline laboratory values only.*

Materials and Methods: *Clinical response to thyroxine (T4 only) was examined in 139 patients who were considered hypothyroid by 16 recognized criteria but whose free thyroxine (FT4) and thyroid stimulating hormone (TSH) fell within 95% laboratory reference intervals (133 patients) or whose FT4 or TSH fell within these intervals (6 patients). Patients were treated with 25 μg day^{-1} thyroxine sodium for 1 week followed by 50 μg day^{-1} for 6 weeks and an increase thereafter of 25 μg at 6 week intervals until the patient was clinically euthyroid. Clinical response was adjudged by improvement or disappearance of clinical features of hypothyroidism; thyroid chemistry was estimated in 41 patients at 6–12 months following institution of thyroid replacement.*

Results: *There was improvement or disappearance of all 16 clinical features in 30 patients (22%) and in over 12 features in 106 patients (76%), with a decrease in the mean number of clinical features from 13.3 ± 0.18 before treatment to 3.0 ± 0.23 following treatment over a minimum follow-up period of 6 months. Energy loss and poor memory and concentration were most responsive to treatment while reduction in tongue size and weight gain improved in 57% and 24% of patients respectively. Clinical response correlated with the level of thyroid replacement but not significantly with pre-treatment or post-treatment levels of FT4 and TSH nor with duration of illness or treatment.*

Conclusions: *Clinically hypothyroid but biochemically euthyroid patients had favourable clinical response to thyroid replacement which correlated with the level of thyroid replacement. It is suggested that these findings be examined in a prospective placebo controlled clinical trial.*

Key words: hypothyroidism, thyroxine, thyroid biochemistry.

INTRODUCTION

Notwithstanding informal communication from the World Health Organisation [1], opinion of the College of General Practitioners UK [2] and traditional wisdom culled over one hundred years of clinical practice [3–7] it is increasingly unusual for family practitioners, physicians and endocrinologists to diagnose and treat hypothyroidism in clinically hypothyroid patients unless the free thyroxine (FT4) and/or thyroid stimulating hormone (TSH) falls outside 95% reference intervals. This arises from a current misconception that evidence-

ISSN 1359-0847 print/ISSN 1364-6907 online/00/020115-10

based medicine means laboratory-based medicine wherein clinical observation—albeit of an objective nature, for example pulse rate or blood pressure—is accorded lower evidential weight than laboratory measurements; this represents an example of a more general misconception in medical science which equates "scientificness" with "molecularity". Secondly, while there have been extensive studies on the complementary problem—the efficacy of thyroid replacement in sub-clinical hypothyroidism, which is most usually (and properly) defined as a patient with no clinical features of hypothyroidism but with one or more (usually TSH) abnormal laboratory measurements [8–12]—there has been no prospective controlled trial of the clinical response to thyroid replacement in clinically hypothyroid but biochemically euthyroid patients.

METHODS

The study comprises the last 139 patients who were managed either solely by the principal author G. R. B. Skinner (GRBS) or whose care was shared with their family practitioner and who were considered clinically hypothyroid but their thyroid chemistry indicated FT4 (1.4%) or TSH (2.8%) or both measurements (95.8%) within the 95% reference interval. The patients were assessed at first visit (and at each subsequent visit) by GRBS and interviewed following their completion of a questionnaire which allowed unbiased report by the patient of her/his clinical features: it is acknowledged that certain criteria, for example size of tongue and thyroid gland, are unavoidably subjective measurements. Patients were entered into the study if there were more than six of the 16 "pivotal" clinical criteria referred to in Table 2; these criteria would be considered by most authorities as reasonable indices of hypothyroidism although in retrospect it might have been useful to have included headache, menstrual disturbance, side vision or floor hallucinations, which—in the author's experience—are characteristic features of moderately severe hypothyroidism and may relate to opacities in the eye which move with traversation of the eye ball. Patients were excluded from the study if they were already receiving thyroid or systemic adrenocorticoid medication.

Sixty-four patients had been diagnosed with myalgic encephalopathy (ME), chronic fatigue syndrome (CFS) or post viral syndrome (PVS), 32 with depression and 43 with other diagnoses; it is emphasized that these diagnoses were provided by the family practitioner or consultant physician and are not defined by any strict criteria of (for example) ME or depression; this stratification is irrelevant to the study and is included for interest but not to posit any differences between these groups or statements of their relative frequency in patients who are hypothyroid or have other attributes in the study; it is indeed apparent that there were no significant differences in clinical features, clinical response to thyroid replacement or thyroid chemistry between patients referred with ME, depression or other diagnosis. Patients with ME, CFS or PVS are collectively referred to as ME in this paper for the purpose of brevity. A control group of 20 healthy volunteers who were randomly selected and had not consulted a family practitioner for any related reason during the last 6 months were included to allow comparison of thyroid chemistry.

The patients were treated with 25 μg thyroxine sodium for 1 week followed by 50 μg day^{-1} for 6 weeks and thereafter increasing by 25 μg integers at 6-week intervals until the patient was clinically euthyroid. The decision to terminate further increase in thyroxine (T4) replacement was based on the clinical findings although follow-up FT4 and TSH readings were obtained in 38 and 41 patients respectively if adjunctive guidance was required on dosage schedule or if the patient requested information on his/her thyroid biochemistry; this was not unusual as most patients seemed well informed and took an active interest in their treatment.

Thyroid biochemistry was estimated on blood samples obtained by venesection from the anterior cubital fossa but occasionally from the back of the hand in patients with inaccessible veins; serum was separated and the blood samples documented and collated by

TABLE 1. Demographic features

	ME	Depression	Other diagnoses	Total
Mean age (years)	39.2 ± 1.7	43.4 ± 2.0	46.8 ± 2.2	
Mean height (m)	1.7 ± 0.0	1.7 ± 0.0	1.7 ± 0.0	
Mean weight (kg)	79.4 ± 3.1	85.3 ± 4.4	80.1 ± 3.2	
Gender				
Female	54	30	42	126
Male	10	2	1	13
Marital status				
Married	32	22	30	84
Marriage dissolution	3	2	7	12
Single	23	5	6	34
Not known	6	3	0	9
Ethnic background				
Caucasian	62	32	42	136
African Caribbean	2	0	1	3
Socioeconomic group				
Group I	2	1	3	6 (4%)
Group II	20	6	8	34 (24%)
Group III	34	21	27	82 (59%)
Group IV	7	3	2	12 (9%)
Group V	2	1	2	5 (4%)
Total	65	32	42	139 (100%)

Micropathology Ltd., University Campus, University of Birmingham and estimations were carried out at the Biochemistry Department, Selly Oak Hospital, Birmingham. Free thyroxine and TSH were estimated by chemiluminescent assay using the automatic chemiluminescent system (ACS 180) supplied by Chiron Diagnostic Ltd, Colchester Road, Halstead, Essex.

Clinical response was adjudged by the number of patients in whom a given clinical feature improved or was no longer present and by the number of clinical features which were no longer present at follow-up. Mean values were compared by Mann–Whitney U test for non-parametric data—for example, clinical features—and by student's t test for parametric data, for example serum FT4 levels. Correlations were examined by inspection of standard graphical display and the significance of correlations by Spearman rank correlation co-efficient (R) for disparate variables—for example, clinical response versus biochemical measurement T4 dose level or duration of illness or treatment—and Pearson's product moment of correlation (r) for compatible variables, for example, FT4 versus TSH levels.

RESULTS

Demographic Features

The study comprised a homogeneous population of middle aged, middle class white female patients with only 12 patients in socioeconomic class 1 and 5. There was no difference in any feature between the ME, depression or other diagnostic group (Table 1).

Features of Disease Prior to Thyroid Replacement

Clinical features. The frequency distribution of the 16 criterion clinical features is shown in Fig. 1(a). Ninety-nine patients (71%) had more than twelve clinical features and none of

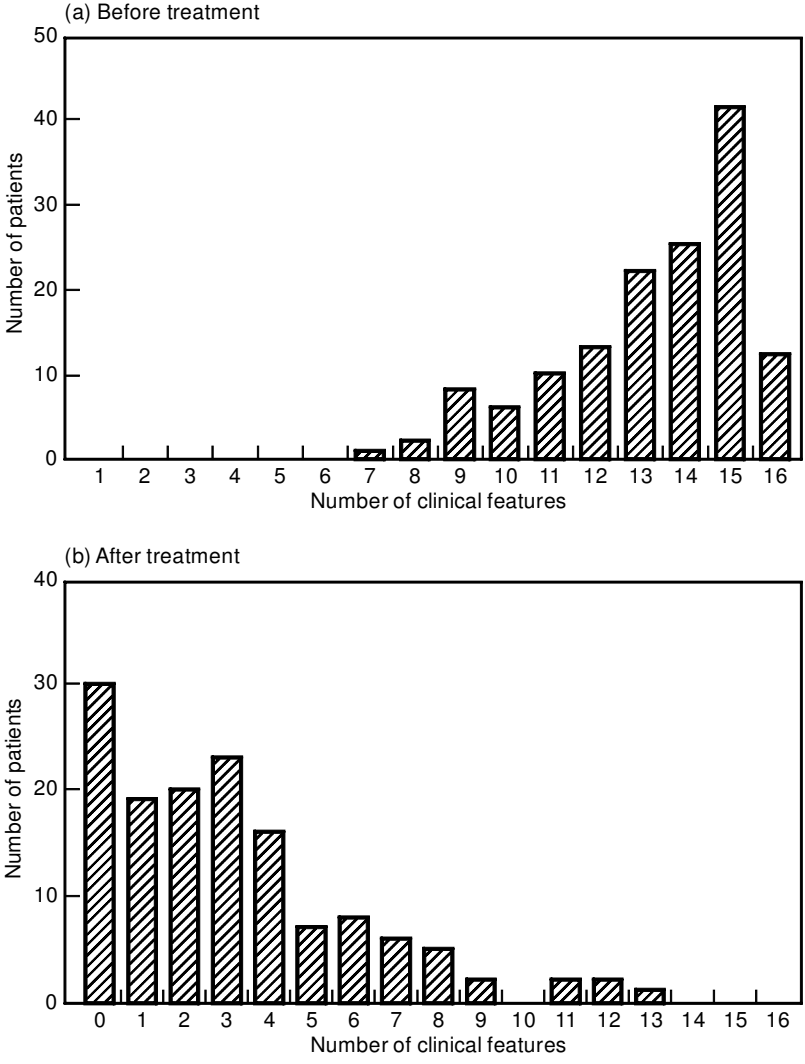

FIG. 1. Distribution of clinical features before and after treatment.

the patients had less than seven features with an average of 13.3 ± 0.18 clinical features for the 139 patients in the study.

The average duration of illness was 7.5 ± 1.2 years. Certain clinical features were consistently represented at high frequency, for example loss of energy, poor memory and concentration, change in general appearance or texture of hair and skin and generalized aches and pains in muscles and joints. Other features were less frequently represented, for example increase in weight was noted in *only 23%* of patients (see Table 2).

Thyroid chemistry. The distribution of FT4, TSH and TSH/FT4 ratio values are indicated in Fig. 2 and Table 3. FT4 values followed an unusual distribution with a cut-off at 10.0 which was artefactual as patients with low FT4 values were excluded from the study unless the patient had a TSH value within the reference interval; the mean FT4 value of 13.5 ± 0.2

TABLE 2. Frequency of clinical features before and after treatment

Clinical features	Before treatment Total	After treatment No longer present	Improved (%)	Unchanged (%)	Worse (%)	Total
Loss of energy, tiredness, lethargy or exhaustion	139	12 (8.6)	114 (82.0)	13 (9.4)		139
Poor memory or concentration or "fuzzy head"	136	17 (12.5)	109 (80.1)	10 (7.4)		136
Changes in appearance	136	10 (7.4)	106 (77.9)	20 (14.7)		136
Change in skin or hair	135	7 (5.2)	94 (71.8)	30 (22.9)		131
Change in voice	132	26 (20.6)	85 (67.5)	14 (11.1)	1 (0.8)	126
Reduced libido	127	17 (13.8)	80 (65.0)	26 (21.1)		123
Muscle or joint pain	126	17 (13.8)	74 (60.1)	31 (25.2)	1 (0.8)	123
Constipation or bloated stomach	122	9 (7.6)	84 (67.8)	25 (21.2)		118
Cold or heat intolerance	118	37 (26.3)	65 (57.0)	12 (10.5)		114
Enlarged thyroid gland	117	17 (15.3)	74 (66.7)	20 (18.0)		111
Loss or blurring of vision	110	26 (23.6)	67 (60.9)	14 (12.7)	3 (2.7)	110
Difficulty in swallowing	108	4 (3.8)	38 (36.9)	61 (59.2)		103
Puffiness of face or extremities	106	5 (4.9)	59 (57.8)	36 (35.3)	2 (1.9)	102
Bradycardia < 65 per minute	104	3 (3.1)	78 (79.6)	17 (17.3)		98
Enlarged tongue	90	22 (24.4)	52 (57.8)	15 (17.7)	1 (1.1)	90
Increase in weight	32	17 (58.6)	7 (24.1)	5 (17.2)		29
Total	1838	246 (13.8)	1186 (66.3)	349 (19.5)	8 (0.4)	1789

The numbers represent patients with each clinical feature.

was significantly below the mean FT4 value for the 20 healthy subjects, namely 16.8 ± 0.53 ($p < 0.0001$, Table 3). Similarly TSH values were not normally distributed but had a skewed distribution to the lower end of the range where, again, patients whose TSH was above the upper limit of the 95% TSH reference interval would be excluded from the study unless their FT4 value was within the interval; the mean TSH value of 2.00 ± 0.10 was significantly higher than the control average value of 1.0 ± 0.14 ($p < 0.0001$, Table 3). The distribution of TSH/FT4 values was again skewed to the lower end of the ratio values with a mean ratio value of 16 ± 1.0, which was significantly higher than the mean ratio value of 6.3 ± 0.8 ($p < 0.0001$) for healthy subjects. There was no significant difference in any biochemical parameter between the three groups of patients, no correlation between FT4 and TSH values for any group of patients, nor correlation between clinical disease as adjudged by these criteria and levels of FT4 or TSH or with TSH/FT4 ratios (data not shown).

Results of Thyroid Replacement

Clinical features. The frequency distribution of clinical features following treatment expressed in terms of improvement or disappearance of each feature is indicated in Table 2 and Fig. 1(b). Thirty patients had improvement or disappearance of every clinical feature (76%) while 129 (93%) patients improved or lost at least ten or more features following treatment (Fig. 1(b)). The mean number of post-treatment clinical features was 3.0 ± 0.23, which was significantly less than the pre-treatment level of 13.3 ± 0.18, $p < 0.0001$ (Fig. 1(a)). Every clinical feature excepting puffiness of the face or extremities (61%) and difficulty in swallowing was improved by at least 70% and certain features, namely loss of energy and poor memory and concentration, were improved or no longer present in at least 80% of patients (Table 2).

These analyses might be complicated by unavoidable differences in the number of initial clinical features prior to treatment where there is less "room for improvement" if a patient

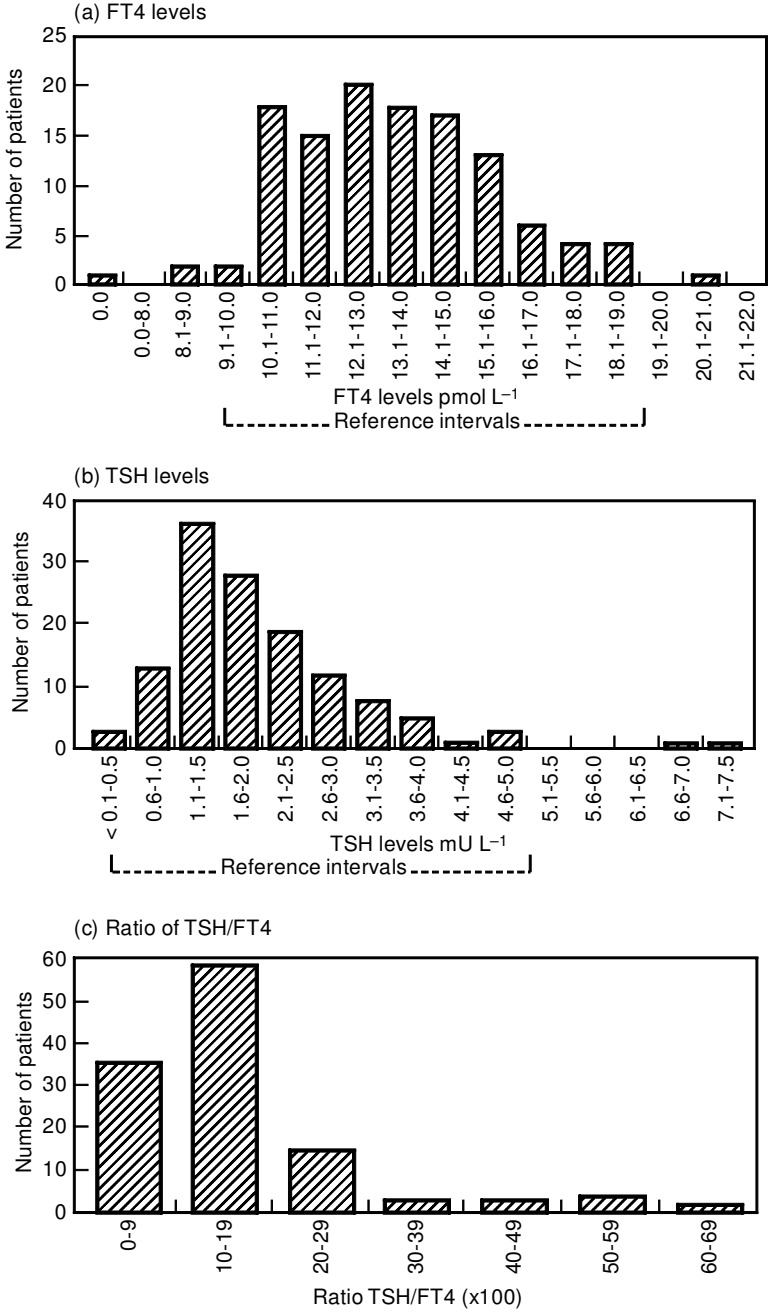

FIG. 2. Distribution of FT4 and TSH levels and ratio of TSH/FT4 before treatment.

presents with only 9 of the 16 features compared to a patient who has all 16 features. However, when clinical response was examined in relation to initial level of disease, there were no significant differences (data not shown).

TABLE 3. Blochemical parameters prior to treatment

	Number	Median	Mean	Range
Free thyroxine (FT4)				
ME	64	13.7	13.6 ± 0.4	8.6–20.6
Depression	32	13.0	13.6 ± 0.4	9.0–17.7
Other diagnoses	43	12.7	13.5 ± 0.4	10.2–12.7
Total patients	139	13.2	13.5 ± 0.2	8.6–20.6
Healthy control subjects	20	16.0	16.8 ± 0.5	15.1–23.6
Thyroid stimulating hormone (TSH)				
ME	64	1.8	2.0 ± 0.2	0.6–6.6
Depression	32	1.7	2.1 ± 0.3	0.5–7.1
Other diagnoses	43	1.7	1.9 ± 0.1	0.1–4.9
Total patients	139	1.7	2.0 ± 0.1	0.1–7.1
Healthy control subjects	20	1.1	1.0 ± 0.1	0.1–2.6
Ratio TSH/FT4 (X100)				
ME	64	14.0	15.0 ± 1.0	4.0–54.0
Depression	32	13.0	17.0 ± 2.7	0.7–67.0
Other diagnoses	43	13.0	16.0 ± 2.0	0.5–59.0
Total patients	139	14.0	16.0 ± 1.0	0.5–67.0
Healthy control subjects	20	6.9	6.3 ± 0.8	0.6–15.8

Thyroid chemistry. Thyroid chemistry before and after thyroid replacement was available for FT4 in 38 and for TSH in 41 patients. Following thyroid replacement, FT4 levels increased from a mean level of 13.2 ± 0.2 to 18.4 ± 0.8 and TSH levels declined from a mean level of 2.3 ± 0.2 to 0.35 ± 0.1; the TSH/FT4 ratio declined from mean 18.0 ± 3.0 to 3.0 ± 1.0 following treatment ($p < 0.001$ for all cases).

Response to thyroid replacement in relation to other parameters. There was no correlation between clinical response and pretreatment FT4, TSH, or TSH/FT4 values, change in these measurements following treatment, or duration of illness or treatment (data not shown). However, the mean FT4 from patients who recovered fully in terms of clinical features was 24.1 ± 2.9, which was significantly higher than the FT4 value of 18.7 ± 1.3 and 18.1 ± 0.9 in patients with poorer clinical response ($p < 0.01$, Table 4). Similarly, TSH levels and TSH/FT4 ratios indices were lower in fully recovered patients although the difference did not reach levels of statistical significance on account of the unrecordable TSH readings in patients receiving T4 replacement (Table 4). Thus the level of FT4 would seem the better—if imperfect—prognosticator of clinical response. There was significant correlation between clinical response and dosage of T4 replacement at follow-up (R 0.5, $p < 0.001$, Fig. 3).

DISCUSSION

This study indicates that patients considered to be hypothyroid on clinical features but with either FT4 or TSH or—as in most cases—both FT4 and TSH within the 95% reference interval had favourable clinical response to thyroid replacement; this was adjudged by improvement or disappearance of a proportion of the 16 clinical features.

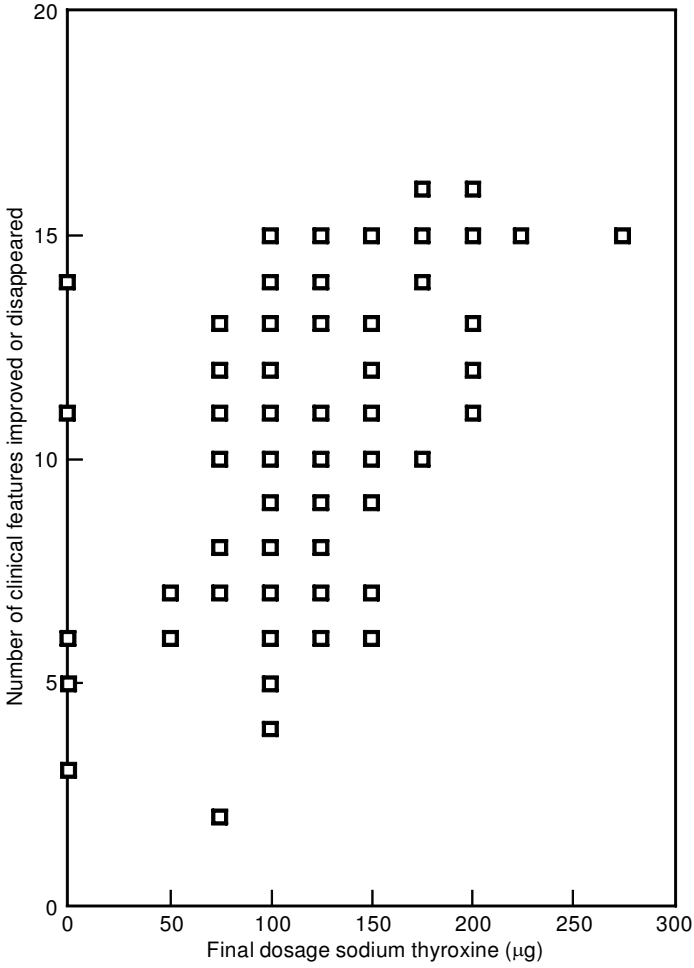

FIG. 3. Relationship between clinical improvement and final dosage of sodium thyroxine.

Our results may under-, or over-, represent purported clinical improvement. There was no attempt to quantify the extent of improvement in any given clinical feature and it is accepted that the record of "improvement" may represent marginal or total improvement where both situations will be recorded as "improvement". Conversely the finding that (for example) 129 of 139 patients ($\sim 93\%$) have improvement or disappearance in at least 10 of the 16 pivotal clinical features under-emphasizes the fact that improvement or disappearance in even one critical feature represents a significant and dramatic improvement in clinical well-being for an individual patient.

In the absence of a control group, a placebo effect cannot be excluded in this or any study. However, the average duration of illness was 7.5 years in patients who had usually undergone an alarming array of traditional and alternative medications without significant improvement as evidenced by their wish to seek further medical advice. Secondly, certain clinical features allowed objective assessment, namely change in appearance, hair or skin texture, reduction in size of tongue and thyroid gland and increase in pulse rate. Therefore, while acknowledging the criticality of formal clinical trial—and indeed we are pleased to report that such a trial has been instituted at Stobhill Hospital, Glasgow, UK—the

TABLE 4. FT4 and TSH levels following thyroid replacement

	Number of clinical features which have improved or are no longer present					
	< 11		11–13		14–16	
	Before treatment	After treatment	Before treatment	After treatment	Before treatment	After treatment
FT4						
Mean	13.4	18.1	13.6	18.7	13.4	24.1
SE	0.33	0.89	0.38	1.33	0.53	2.88
Number	56	21	45	17	18	9
TSH						
Mean	2.15	0.39	1.74	0.36	2.20	0.19
SE	0.16	0.16	0.12	0.12	0.27	0.16
Number	59	21	51	17	19	8
TSH/FT4 (X100)						
Mean	17.0	3.0	13.0	3.0	20.0	1.0
SE	2.0	1.0	1.0	1.0	3.0	1.0
Number	54	21	44	17	18	8

proposition of placebo effect to explain sustained clinical response in 135 of 139 patients is unlikely.

The absence of any significant correlation between clinical response and FT4 or TSH levels questions the sometime over-reliance of practitioners on these criteria of therapeutic benefit particularly if considered without due cognizance of improvement in clinical features. However, the correlation of clinical response with final dose level of thyroid replacement (Fig. 3) confirms a long-held belief of the author (GRBS) that many patients are chronically under-treated in terms of thyroid replacement for reasons that are unclear but appear to relate to ill-founded fears of thyroid gland suppression, osteoporosis or cardiac catastrophe on evidence which is nonexistent in patients who are well monitored and are receiving sensible and incremental T4 replacement. In summary, many patients who were clinically hypothyroid by established criteria but had normal thyroid biochemistry had dramatic and sustained response to thyroid replacement.

A number of hypotheses might explain this finding. These would include the simplistic mathematical argument that the skewed distribution of FT4 and TSH values negates application of Gaussian theory to (for example) TSH values in excess of 3 in the usual 0.5 to 5.5 scale; this aspect has been explored in interesting studies by *Billewicz et al.* and Meler *et al*. [13, 14] which emphasize the insecurity of FT4 or TSH values as pivotal diagnostic criteria. Secondly, the level of serum FT4 and FT3 adjudged by an immunological estimation does not measure biological activity; FT4 or FT3 could be neutralized by enzymatic or immunological mechanisms, yet retain the epitope which is operative in a laboratory immune assay and indeed might still, albeit inactivated, depress pituitary TSH levels. Thirdly, the "correct" or physiological level of FT4 or TSH for optimal health in a given patient or indeed the global population might encourage speculation that certain populations—including the UK—have somehow become deficient in FT4, perhaps from nuclear or chemical environmental insult to the thyroid gland or from chlorination or fluoridation of water supplies or from increased immunological or enzymatic deactivation of circulating T4 or tri-iodothyronine; discrimination of these hypotheses by appropriate epidemiological or immunological research would be a feasible and fruitful research programme.

However, irrespective of aetiological considerations, we suggest that the rigid concept of hypothyroidism as a disease—versus euthyroidism as a state of optimal health—should be

reconsidered towards the concept that there may be a spectrum of optimal health in terms of thyroid related body functions which relates approximately to levels of FT4 where (for example) patients whose FT4 levels are 17 or above feel better than patients with levels between 17–11, while levels 10 or 11 are associated with sub-optimal health. It is ridiculous—although one author (GRBS) has observed this on many occasions during the last decade—that a patient whose FT4 level is 0.1 below the lower limit of a reference interval is considered "hypothyroid" but if it is 0.1 above the lower limit the patient is considered "not hypothyroid" and thus should not receive appropriate thyroid replacement.

In summary we feel that the results of this study should be subjected to the scrutiny of formal clinical trial to examine not the relationship of abnormal thyroid chemistry to disease, which is well established, but the relationship between "normal" thyroid chemistry and ill health.

REFERENCES

[1] Holmes D. Tears behind closed doors. Avon Books. London: 1998.
[2] Chronic Fatigue Syndrome. Report of a joint working group of the Royal Colleges of Physicians, Psychiatrists and General practitioners. Cathedral Print Services, Salisbury 1996.
[3] Bayliss RIS. Thyroid Disease: the Facts. Oxford: Oxford University Press, 1982.
[4] DeGroot LJ, Larsen PR, Refetoff S, Stanbury JB. The Thyroid and its Diseases. Wiley Medical, Chichester, NY 1984.
[5] Ingbar SH, Braverman LE. Werner's The Thyroid, 5th edn. JB Lippincott Company, 1986.
[6] Burrow GN, Oppenheimer JH, Volpé R. Thyroid Function and Disease. Sanders, Philadelphia 1989.
[7] McDougall IR. Thyroid Disease in Clinical Practice. Chapman and Hall Medical, London 1992.
[8] Staub J-J, Althaus BU, Engler H, et al. Spectrum of subclinical and overt hypothyroidism: effect on thyrotropin, prolactin and thyroid reserve, and metabolic impact on peripheral target tissues. Am J Med 1992; 92: 631–42.
[9] Surks MI, Ocampo E. Subclinical thyroid disease. Am J Med 1996; 100: 217–23.
[10] Cooper DS, Halpern R, Wood LC, Levin AA, Ridgway EC. L-thyroxine therapy in subclinical hypothyroidism. A double-blind, placebo-controlled trial. Ann Intern Med 1984; 101: 18–24.
[11] Weetman, AP. Hypothyroidism: screening and subclinical disease. BMJ 1997; 314: 1175–8.
[12] Nystrom E, Caidahl K, Fager G, Wikkelso C, Lunberg PA, Lindstedt G.A. Double-blind cross-over 12-month study of L-thyroxine treatment of women with "subclinical" hypothyroidism. Clin Endocrinol 1988; 29: 63–76.
[13] Billewicz WZ, Chapman RS, Crooks J, Day ME, Gossage J, Wayne Sir E, Young JA. Statistical methods applied to the diagnosis of hypothyroidism. Q J Med, 1968; New Series XXXVIII, no. 150: 255–66.
[14] Meler CA, Malsey MN, Muller J, Smith MA. Inter individual differences in the pituitary-thyroid axis influence the interpretation of thyroid function tests. Clin endocrinol 1993; 39: 101–7.