A Strategy for Thyroid Function Tests

K. E. BRITTON, VALERIE QUINN, B. L. BROWN, R. P. EKINS

British Medical Journal, 1975, 3, 350-352

Summary

The concept of decision-aiding ranges was introduced to improve diagnostic efficiency. The clinical uncertainty in distinguishing borderline cases of hyperthyroidism and hypothyroidism among 1569 consecutive requests for thyroid function tests was 47%. This was reduced to 22% by using the free thyroxine index and to under 2% by using specific tests determined by the decision-aiding ranges.

Introduction

Relevant diagnostic tests are those whose results alter clinical management. The interaction between clinician and laboratory is becoming inefficient owing to the increase in laboratory requests and the number and complexity of tests applicable to a specific diagnostic problem. Strategies are required to relate each diagnostic problem to the appropriate tests. Decision-aiding ranges were introduced in the Middlesex Hospital to help the laboratory to decide automatically from the result of a screening test which specific test should be undertaken next. Our aims in introducing the strategy were twofold: to improve accuracy in answering the diagnostic problem and to improve efficiency by reducing the time, money, and effort wasted by patients, clinicians, and diagnostic departments in obtaining that answer.

Rationale and Method

The accuracy of the answer depends on biological variations between patients, test variations, interaction errors where external or internal factors acting in the patient affect the test, and errors such as mislabelling the specimen tube. The efficiency of the diagnostic test is affected by the number of consultations with the patient, the relevance and accuracy of the request, and the organization (from the initiation of a request to the decision on behalf of the patient by the requestor) and time required for the assay. Given a particular state of knowledge about a patient's disorder and the need to run an efficient and accurate diagnostic service the emphasis in reporting results should be on "decision-aiding ranges" rather than "normal values." With these objectives we re-evaluated thyroid function tests.

The old ^{131}I uptake test is inaccurate, because of interaction errors, and inefficient, because patients have to attend several times. The ^{99m}Tc pertechnetate 20-minute uptake test and other early uptake tests do not readily cope with hypothyroidism or with the usual "screening" request (87% of all requests in this study), but their immediacy may help in a thyroid treatment follow-up clinic which the patient is already attending.[1 2] The protein-bound iodine test, though economic and automated, suffers severely from interaction errors due, for example, to iodide ingestion and from variations in thyroxine-binding globulin (T.B.G.), as in pregnancy; there remains the worry of obtaining a normal result from a patient who is hypothyroid but has had slight iodine contamination.

Department of Nuclear Medicine, Middlesex Hospital Medical School, London W1N 8AA

K. E. BRITTON, M.D., M.R.C.P., Senior Lecturer in Nuclear Medicine
VALERIE QUINN, M.SC., Research Assistant
B. L. BROWN, B.SC., PH.D., Lecturer
R. P. EKINS, M.A., PH.D., Professor of Biophysics

As medicine becomes more oriented towards community practice tests that do not require hospital attendance become more attractive. To assess thyroid function blood tests such as the effective thyroxine ratio,[3 4] the normalized thyroxine ratio,[5] and the free thyroxine index (F.T.I.) meet most of the requirements. The F.T.I. is estimated by dividing the total serum thyroxine determined using analysis[6] by the thyopac-3 value,[7] a representation of T.B.G. binding sites. Hospital attendance is not needed and errors due to iodine contamination and T.B.G. variation are avoided. We therefore devised a strategy using only blood tests to assess thyroid function.

DECISION-AIDING RANGES

Many doctors demand a "normal range" for each test. Thus, we give a normal range for the F.T.I. of 58-135 nmol/l (45-105 μg/l) while hoping that doctors appreciate that, for example, values of 38 μg/l or 115 μg/l are not diagnostic of hypothyroid or hyperthyroid function respectively. In routine clinical practice these tests produce many equivocal results (22% in this study) owing to the natural overlap of hyperthyroid and hypothyroid values into the "normal" range, which is caused by the compounding of biological and assay variations. Such equivocal results tend to lead to further consultations, and the efficiency of the diagnostic process is thus reduced.

To deal with this problem we divided the F.T.I. results into five decision-aiding ranges. *Definitely hypothyroid* (S1) and *definitely hyperthyroid* (S5) were ranges into which no euthyroid patients' results should ever fall and *definitely euthyroid* (S3) was the range into which no hypothyroid or hyperthyroid patients' results should ever fall. Two further categories—*borderline raised* (S4) and *borderline low* (S3)—were introduced. We tried to regroup the equivocal F.T.I. results (S4) into the hyperthyroid or euthyroid ranges by estimating the serum total triiodothyronine (T-3) in patients with borderline raised results (S4) and the serum thyroid stimulating hormone (TSH) in those with borderline low results (S2). In either case the F.T.I. results were returned to the requester with a comment indicating that T-3 or TSH, as appropriate, would be measured on the original sample. Thus the patient did not have to attend again. Hence, the F.T.I. was the "screening" test and total T-3[8] and serum TSH[9] assays were more specific tests.

OVERALL STRATEGY

In a routine service laboratory effort is wasted by requesting a TSH assay in a patient with suspected hyperthyroidism or a T-3 estimation in one with suspected hypothyroidism; T-3 is a poor discriminator of hypothyroidism;[10] and a single value of serum TSH is a poor discriminator of hyperthyroidism[9 11]—the TSH assay is less precise at low hormone concentrations. Thus, under our strategy a direct request for T-3 or TSH assay was not indicated except on the basis of F.T.I. result. Since this is not true in cases of, for example, T-3 toxicosis[12 13] the simple strategy was incorporated into an overall strategy in which clinical and treatment status of the patient was recognized. Thus, even if the result of a F.T.I. estimation was quite "normal" (S3) a total T-3 was automatically done if the request form indicated "definite clinical hyperthyroidism," but not if it stated just "? thyrotoxic." Similarly, "? endocrine exophthalmos" elicited both TSH and T-3. If, for example, the patient was taking carbimazole then an equivocal S4 result on screening would not automatically lead to a T-3 test nor an equivocal S2 result to a TSH assay since in this context the benefits of T-3 or TSH over F.T.I. measurements in routine practice have not been effectively evaluated.

The clinical and treatment categories into which all request forms were divided are shown in table I. We accepted combinations of treatment categories but only one clinical category. The decision-aiding ranges for S1-S5 were initially chosen on the basis of previous experience, but it became clear that even wider boundaries for borderline decision-aiding ranges were necessary (table II). The strategy incorporating clinical, screening, and treatment categories is given in table III.

TABLE I—*Clinical and Treatment Classification from 1569 Consecutive Request Forms*

Code	Clinical Category	% Frequency	Code	Treatment Category	% Frequency
C0	Definitely hypothyroid	0·45	D0	Not in any category D1-D9	56·09
C1	Probably hypothyroid	0·76	D1	Thyroxine	14·40
C2	Possibly hypothyroid	19·57	D2	Triiodothyronine	0·12
C3	Probably euthyroid	44·61	D3	Carbimazole	7·05
C4	Possibly hyperthyroid	23·45	D4	Propylthiouracil or perchlorate	0·36
C5	Probably hyperthyroid	2·04	D5	Cortisone-like steroids	3·64
C6	Definitely hyperthyroid	0·77	D6	Pregnant or on oral contraceptives	1·61
C7	Probably T-3 toxicosis	0·32	D7	Previous ^{131}Iodine therapy	6·57
C8	Endocrine exophthalmos	2·10	D8	Previous thyroid surgery	3·47
C9	Pituitary disorder	5·93	D9	Other drugs	6·69

TABLE II—*Normal and Decision-aiding Ranges*

		Values Chosen at Start of Study				Final Values
		F.T.I.		T-3 (pmol/l)	TSH (mU/l)	F.T.I. (nmol/l)
		(nmol/l)	% Frequency			
Normal range:		58–135		922–2765	<2–8	
Hypothyroid	S1	<45	5·16		⩾20	<38
Borderline low	S2	45–67	14·47		12–19	38–77
Euthyroid	S3	68–128	61·89	922–2765	<1·11	78–118
Borderline high	S4	129–155	7·78	2766–4148		119–176
Hyperthyroid	S5	>155	10·70	>4148		>176

Conversion: SI to Traditional Units—
F.T.I.: 1 nmol/l ≈ 0·78 μg/l. T-3: 1 pmol/l ≈ 0·65 pg/ml.

TABLE III—*Modification of Automatic Strategy Response by Treatment Categories*

Code	Automatic Strategy Response	Code	Modification of Automatic Response
C0	TSH for S2 and S3	D0	No effect
C1	TSH for S2 and S3	D1	No automatic T-3 or TSH
C2	TSH for S2 and S3	D2	T-3 assayed not T-4
C3	TSH for S2; T-3 for S4	D3	No automatic T-3 or TSH
C4	TSH for S2; T-3 for S4	D4	No automatic T-3 or TSH
C5	T-3 for S2 and S3	D5	Combination C9 and D5 excludes TSH
C6	T-3 for S2 and S3	D6	No effect
C7	T-3 for S2, S3, and S4	D7	Combination C4, D7, and S3 elicits T-3
C8	TSH for S2; T-3 and TSH for S3; T-3 for S4	D8	Combination C4, D8, and S3 elicits T-3
C9	TSH for S1-S5	D9	No effect

Results

From a total of 1569 consecutive request forms we excluded the endocrine exophthalmos (C8) and pituitary (C9) categories and added the probable T-3 toxicosis category (C7) to the hyperthyroid category. We were left with 1443 requests, which was normalized down to 1000 and the figures rounded for clarity in table IV. The frequency distribution of the S categories is given in table II. In table IV the TSH frequency distribution was derived from 343 results and the T-3 frequency distribution from 359 results taken only from clinical categories C2, 3, and 4 and treatment categories D0, 5, and 9. The results shown in table IV therefore concern patients not suspected of having pituitary disorders and free of any specific treatment who presented with a possible thyroid problem. The clinical uncertainty, as recorded on the request forms, was 47%. By using the F.T.I. screening test this was reduced to 22%, and the specific tests further reduced this uncertainty to under 2%.

The distribution of results for T-3 versus F.T.I. is shown in table V for all F.T.I. categories. The S4 categories gave 109 results in the "normal" T-3 range, 20 borderline results, and six hyperthyroid values. Out of 116 values in the definitely euthyroid (S3) range, 17 lay in the borderline raised range for T-3. Whereas only one case of T-3 toxicosis was found where the F.T.I. result was in the borderline raised category, eight cases of biochemical T-4 toxicosis were noted where the T-3 result was in the normal range.

TABLE V—*T-3 versus F.T.I.: Distribution of Results in each Decision-aiding Range*

T-3 (pmol/l)	F.T.I. (nmol/l)					
	S1 (<38)	S2 (–77)	S3 (–118)	S4 (–176)	S5 (>176)	Total
0–2765	9	67	99	109	8	292
–4148		8	17	20	4	49
–5530				5	3	8
>5530				1	9	10
Total	9	75	116	135	24	359

The relation of these results to the previous normal range for T-3 and for the F.T.I. is shown in table VI. There was much overlap. The T-3 results in three of the eight T-4 toxicoses were below the normal range for T-3 and the "T-3 toxic" fell in the "normal" T-4 range as identified in table VI but in the "borderline raised" T-4 range of the decision-aiding range strategy shown in table V. The clinical features of patients with biochemical T-4 toxicosis will be presented separately.[14]

TABLE VI—*F.T.I. versus T-3: Distribution of Results into Normal Ranges*

F.T.I. (nmol/l)	T-3 (pmol/l)			
	⩽922	–2765	–4148	>4148
<58	17	21	5	0
–135	9	183	28	1
–176	9	46	12	5
>176	3	5	4	12

TSH versus F.T.I. for all F.T.I. values is shown in table VII. The S2 category gave 216 results in the "normal" TSH range, 14 in the borderline, and 28 in the hypothyroid range. As the specific test TSH discriminated well; only one borderline TSH value was found in the euthyroid S3 category. The seven normal TSH values in the S1 category were due to misclassification on the request forms of "pituitary" cases as primary hypothyroid cases.

TABLE IV—*Simplified Decision-aiding Range Strategy: Results Normalized to 1000 Requests*

Class	Clinical Assessment	No. of Requests in Each Class	Class	F.T.I. Screening Assessment	No. of Assays in Each Class	TSH Special Test	No. of Assays	T-3 Special Test	No. of Assays	Diagnosis	Total (%)
C0	Hypothyroid	10	S1	Hypothyroid	50	Hypothyroid	18			Hypothyroid	68 (7)
C1	Probably hypothyroid										
C2	Possibly hypothyroid	210	S2	Borderline low	140	Borderline	9			Borderline	9 (1)
C3	Euthyroid	490	S3	Euthyroid	620	Euthyroid	113	Euthyroid	70	Euthyroid	803 (80)
C4	Possibly hyperthyroid	260	S4	Borderline high	80			Borderline	7	Borderline	7 (1)
C5	Probably hyperthyroid	30	S5	Hyperthyroid	110			Hyperthyroid	3	Hyperthyroid	113 (11)
C6	Hyperthyroid										
	Total	1000			1000		140		80		1000 (100)

TABLE VII—*TSH versus F.T.I.: Distribution of Results in each Decision-aiding Range*

TSH (mU/l)	F.T.I. (nmol/l)					
	S1 (<38)	S2 (-77)	S3 (-118)	S4 (-176)	S5 (>176)	Total
0-11·5	7	216	60	6		289
-19·5	1	14	1			16
>19·5	10	28				38
Total	18	258	61	6	0	343

Tables V, VI, and VII show the futility of trying to widen or narrow the normal F.T.I. ranges to encompass all results so as to place them into their appropriate categories. This also highlights the need for replacing normal ranges with decision-aiding ranges.

Discussion

Let us consider two extremes: the doctor who thinks that thyrotoxicosis and hypothyroidism are clinical entities diagnosable on clinical grounds and that tests are a waste of effort; and the doctor who feels that thyrotoxicosis and hypothyroidism are states in which biochemical definition is essential because the clinical criteria are so often misleading. The former extreme is unlikely since only 13% of the request forms we saw contained a firm clinical diagnosis. The latter extreme is unlikely since attempts to define a range for normality and for disordered states founder on the continuous spectrum of thyroid disorder.

Since the essence of diagnosis is to determine the management that is best for the patient we devised a compromise strategy of defining "decision aiding ranges." Unlike computer-aided diagnosis,[15] the general principle of decision-aiding ranges relies on the current view of methods of diagnosis rather than the frequency distribution of disease which is needed for the Bayesian approach. Any doctor can set up his own strategy based on his own view of the diagnostic problem and the services offered by his laboratory.

The weakness in our study was the poor quality of the information on the request forms, but this may be overcome by designing request forms with coded clinical data. Thus, the numerical result may be correlated with the clinical state encoded in the request form through a set of rules for a technician or a computer program.[16] Also, the clinical acumen of the requester should be sharpened by having to record the clinical and treatment input accurately. The decision-aiding range strategy prevents all thyroid function tests to being requested at once. Our strategy is only one view of the state of thyroid function tests. Other strategies are equally valid and we are searching for one that will give the best results from the clinical, laboratory, and cost-effective view points.

We shall consider the effects of antithyroid and thyroid replacement treatment and pituitary disorders and endocrine exophthalmos on the strategy elsewhere. But, clearly, more complex decision-aiding ranges than those outlined here are needed for the appropriate evaluation of treatment. Nevertheless, our simple strategy for thyroid function assays reduced the clinical uncertainty in routine hospital practice from 47% to under 2%. For borderline raised F.T.I. results T-3 estimation was moderately successful but for borderline low results TSH estimation discriminated well between the euthyroid and the hypothyroid state.

We thank Professor E. S. Williams for his encouragement.

References

1 Goolden, A. W. G., Glass, H. I., and Williams, E. D., *British Medical Journal*, 1971, **4**, 396.
2 Ell, P. J., and Keeling, D. H., *Clinical Radiology*, 1974, **25**, 217.
3 Mincey, E. K., Thorson, S. C., and Brown, J. L., *Clinical Biochemistry*, 1971, **4**, 216.
4 Murray, I. P. C., Parkin, J., and Gubanyi, M., *Medical Journal of Australia*, 1972, **1**, 1190.
5 Ashkar, F. S., and Bezjian, A. A., *Journal of the American Medical Association*, 1972, **221**, 1483.
6 Ekins, R. P., Williams, E. S., and Ellis, S., *Clinical Biochemistry*, 1969, **2**, 253.
7 Clark, F., and Brown, J., *British Medical Journal*, 1970, **1**, 713.
8 Hall, R., *Clinical Endocrinology*, 1972, **1**, 115.
9 Hoffenburg, R., *Clinical Endocrinology*, 1973, **2**, 75.
10 Patel, Y. C., and Burger, H. G., *Clinical Endocrinology*, 1973, **2**, 339.
11 Toft, A. D., *et al.*, *Clinical Endocrinology*, 1972, **2**, 127.
12 Hollander, C. S., *et al.*, *Lancet*, 1972, **1**, 609.
13 Wahner, H. W., *Mayo Clinic Proceedings*, 1972, **47**, 938.
14 Britton, K. E., *et al.*, *Lancet*, in press.
15 Oddie, T. H., *et al.*, *Journal of Clinical Endocrinology and Metabolism*, 1974, **38**, 876.
16 Ekins, R. P., proposal to the Supra-regional Assay Service, 1973.
17 Ellis, S. M., and Ekins, R. P., *Journal of Endocrinology*, 1973, **59**, 43.
18 Ekins, R. P., and Ellis, S. M., 7th International Thyroid Conference, Boston 1975, in press.

SHORT REPORTS

Hypervitaminosis A Accompanying Advanced Chronic Renal Failure

Many symptoms and lesions appearing in both experimental and clinical vitamin A intoxication (such as anorexia, nausea, vomiting, skin dryness, headache, pruritus, muscle fasciculation, peripheral paraesthesias, bleeding, and bone changes) are also common in severe uraemia. This prompted us to investigate serum vitamin A levels in patients with advanced renal failure.

Patients, Methods, and Results

We studied four groups of patients: group 1 consisted of 100 normal people (35 women and 65 men) aged between 20 and 61: group 2 of 12 patients (five women and seven men) aged between 20 and 68 with severe acute renal failure and serum creatinine over 884 μmol/l (10 mg/100 ml); group 3 of 100 patients (40 women and 60 men) aged between 18 and 52 with advanced chronic renal failure and serum creatinine constantly over 707 μmol/l (8 mg/100 ml); and group 4 of 100 patients (32 women and 68 men) aged between 16 and 64 undergoing regular chronic haemodialysis in five centres in Athens. Dialysis had lasted from six to 50 months using Ultra Flo 100, EX-01, or SP 75 R coil at a blood flow rate 250-350 ml/min three times a week (15-18 hours). The dialysate concentration of sodium was 135·0, acetate 35·0, chloride 102·0, calcium 1·75, magnesium 0·5, and potassium 1·5 mmol/l.

No subject had taken any hormones or vitamins for at least four months. Blood samples were taken after an overnight fast at 08.00 hours. All measurements were made in duplicate by the trifluoroacetic acid method (macroprocedure).[1]

The serum vitamin A results are summarized in the diagram. There was no significant difference between mean serum vitamin A in the controls (0·94±0·28 (0·03) μmol/l (27·75±8·10 (0·8) μg/100 ml)) and that in patients with acute renal failure (0·96±0·30 (0·08) μmol/l (27·50±8·57 (2·4) μg/100 ml) (t=0·09). In contrast the mean value was high in patients with advanced chronic renal failure (2·13±0·73 (0·07) μmol/l (60·93±21·0 (2·1) μg/100 ml) (t=15·0), and higher in patients undergoing chronic haemodialysis (3·00±0·72 (0·07) μmol/l (85·37±20·49 (2·0) μg/100 ml) (t=26·2). The difference in both groups was significant ($P<0·001$, and $P<0·001$, respectively).