

CLINICAL STUDY

From guidelines to hospital practice: reducing inappropriate ordering of thyroid hormone and antibody tests

Marie-Elisabeth Toubert, Sylvie Chevret¹, Bruno Cassinat, Marie-Hélène Schlageter, Jean-Paul Beressi² and Jean-Didier Rain

Service de Médecine Nucléaire, ¹Département de Biostatistique et Informatique Médicale, and ²Service d'Endocrinologie, Hôpital Saint-Louis, Paris, France

(Correspondence should be addressed to Marie-Elisabeth Toubert, Service de Médecine Nucléaire, Hôpital Saint-Louis, 1, avenue Claude Vellefaux, 75475 Paris Cedex 10, France; Email: marie-elisabeth.toubert@sls.ap-hop-paris.fr)

Abstract

Objective: Because of major technical improvements and conscious care about cost effectiveness, limiting the inadequate use of thyroid biological tests appears to be a major issue.

Design: To (i) estimate the ordering prevalence of each thyroid test, (ii) assess the prevalence of relevant thyroid tests, and (iii) evaluate the impact of expressing justification for tests during a 2-month intervention period on these prevalences.

Methods: During a prospective 2-month survey (June–July 1997), all the request forms were divided into four groups of prescription: (1) investigation of thyroid function, (2) taking drugs affecting the thyroid, (3) monitoring of nodule and cancer, and (4) investigation of thyroid autoimmunity. Their appropriateness was thus determined according to consensus in our hospital and previously published recommendations. Results were compared with those of retrospective similar 2-month periods in 1996 and 1998. Combinations of thyroid function tests and thyroid antibodies were analyzed during the 1996, 1997 and 1998 periods.

Results: The overall estimated rate of appropriate ordering between 1996 and 1997 increased from 42.5% to 72.4% ($P < 10^{-4}$), with a significant improvement in each group of main diagnosis referral, except in group 3 where suitability was always over 85%. However, in group 4, appropriateness remained low (36%). Combinations of thyroid tests revealed an increase in single TSH order forms and single autoantibodies to thyroperoxidase (TPOAb) ones, while TSH + free thyroxine + free tri-iodothyronine and TPOAb + autoantibodies to thyroglobulin ones decreased significantly. Interestingly, all these changes were maintained 1 year later (June–July 1998) even though physicians were not aware of this new study.

Conclusions: Persistent change in medical practice was thus assessed.

European Journal of Endocrinology 142 605–610

Introduction

The last decade has seen major advances in thyroid laboratory testing, particularly related to improvements in thyroid function tests, mainly thyrotropin (TSH) assays (1, 2), and in methods for measuring thyroid autoantibodies with the identification of the microsomal antigen as the thyroperoxidase (3). Thyroid biological testing is an essential tool for disease screening, diagnosis, and monitoring. Although algorithms and guidelines for thyroid hormone and antibody testing have been published in order to optimize the practice of physicians (4–9), the appropriate use of these tests remains controversial and medical practice may vary. Pressure to use tests in a more cost-effective manner continues however. Improving the appropriateness of

testing behavior is therefore a major focus of quality improvement, particularly since the ordering is not restricted to specialists in endocrinology, but is usually done by residents and interns, and even occasionally by nurses.

Except in iodine-deficient areas where free tri-iodothyronine (FT3) and free thyroxine (FT4) should both be regarded as second-line tests (10) because T3 toxicosis is fairly common, the ordering of T3 is unquestionably a third-level measurement that should be limited to restrictive indications, such as discordance between TSH and FT4 levels and/or discordance between TSH measurement and clinical symptoms (4, 11, 12). The Mayo Clinic group has thus estimated that only 2% of patients require T3 or FT3 measurements (11). In the Assistance Publique-Hôpitaux de Paris, a hospital

	TSH	FT4	FT3	TPOAb	TgAb	Tg
(1) Investigation of thyroid function		a	b			c
if hypothalamic-pituitary disease						
(2) Taking drugs affecting the thyroid						
if acute dose adjustment (LT4*, ADT**), or after immediate surgery ^o or ¹³¹ I-therapy						
if amiodarone		d	d			
(3) Monitoring of nodule and cancer		e	e		f	f
(4) Investigation of thyroid autoimmunity					g	

Figure 1 Criteria of relevant biological thyroid tests, according to local consensus and the literature. *LT4, L-thyroxine; **ADT, antithyroid drug therapy; surgery^o, total or partial thyroidectomy; TSH, third-generation TSH assay. First-line tests are shown by open boxes, follow-up or complementary tests by shaded boxes and no-use tests by solid boxes. a, if inpatients (23, 29); b, if discordance between TSH measurement and clinical thyroid evaluation or FT4 results (4, 11) or if suspicion of T3 toxicosis in iodine-deficient areas (10); c, if suspicion of thyrotoxicosis factitia (30); d, if suspicion of amiodarone-induced hyperthyroidism (31); e, if LT4-suppressive therapy, FT4 or FT3 can be prescribed (32); f, in the follow-up of differentiated cancer (32), and g, when TPOAb is negative (9).

network of 50 hospitals located in the Paris region, we have previously found an overall abnormally high proportion of T3 requests (22% of the overall thyroid function tests assayed during a cross-sectional survey conducted in June 1995) (12). In the Saint-Louis Hospital, during the same month, this proportion was even higher (29%).

We therefore decided to conduct a prospective study in our hospital to (i) estimate the ordering prevalence of each thyroid test (thyroid function tests and thyroid antibodies) with an analysis of their eventual combinations, (ii) assess the prevalence of relevant thyroid tests based on the literature and on the basis of expert opinion and (iii) evaluate the impact of expressing justification for the test required, spread over a 2-month intervention period (June–July 1997), on these prevalences. The long-term effect of this intervention was tested 1 year later (June–July 1998), while physicians were not aware of this new study.

Materials and methods

Design and intervention

Our institution is an 800-bed teaching hospital which registers an average of 25 266 inpatients and 242 013 outpatients a year. The treatment setting of our hospital is a large multispeciality out-patient clinic combined with acute care departments and an endocrinology department. The biochemistry laboratory workload reaches an average of 2026 137 tests per year, 5% being immunoanalyses. In 1997, biological thyroid tests accounted for 16% of all immunoanalyses.

The order sheet used formerly quoted all available thyroid tests (TSH, FT4, FT3, autoantibodies to thyroperoxidase (TPOAb), autoantibodies to thyroglobin (TgAb), autoantibodies to TSH-receptor (TRAb), and

thyroglobin (Tg)) and provided information on the cost of each test. Physicians were expected to justify the test(s) they needed; in fact, clinical information as well as justification of the requested tests was given in about 40% of the cases. During an intervention period of 2 months (June–July 1997) and after a local information campaign, all the physicians in our hospital were asked to provide a proper clinical justification if a parameter other than TSH alone was ordered. Furthermore, as many hospital laboratories are now using algorithms, we also proposed a TSH-based strategy for the investigation of thyroid function with automatic subsequent FT4 measurement in cases of abnormal TSH results. When no justification was given, even though all the assays ordered were performed, a repeat letter was sent to the prescriber.

Order forms collected during the intervention period were subgrouped according to the clinical conditions of the prescription, as listed in Fig. 1. The appropriateness of each prescription was thus determined according to local consensus with the Endocrine Department of our hospital and previous published recommendations, as shown in Fig. 1.

To assess the impact of the intervention, the former order sheets during a past 2-month period of June and July 1996 were retrospectively analyzed and split up into the four defined groups of prescription: ordering rates and relevance according to Fig. 1 were thus compared. Finally, 1 year later in June–July 1998, similar analyses were performed without informing the physicians.

Statistical methods

Comparisons of appropriate use rate of specific thyroid function tests between pre-intervention, intervention, and post-intervention periods were based on the

Table 1 Number and nature of thyroid function test requests during the 1996, 1997 and 1998 2-month study periods.

Nature of test	Pre-intervention period (1996) (no. of tests)	Intervention period (1997) (no. of tests)	Post-intervention period (1998) (no. of tests)
Total TSH tests	1329	1119	1062
Total FT4 tests	1011	535	539
Total FT3 tests	715	247	226
Total number of thyroid function tests	3145	1901	1827
Single TSH	305	563	512
TSH+FT4	319	313	333
TSH+FT3	23	25	20
TSH+FT4+FT3	682	218	197
FT4+FT3	10	4	9
Total number of thyroid function order forms	1339	1123	1071

The total number of thyroid tests (thyroid function tests) is indicated in the upper part. As thyroid tests may be prescribed in association, the combinations are noted and numbered in the lower part.

chi-square test and the exact Fisher test when required. Analysis was then stratified according to the prescription group as detailed above. Estimated appropriate rates are reported with 95% confidence intervals.

Statistical analysis was performed with the SAS software system (Statistical Analysis System, Inc., Cary, NC, USA).

Results

Number and nature of thyroid laboratory test requests

Our study focused on the request forms for specific thyroid tests filled in by physicians in our hospital, including specialists in endocrinology, but also residents and interns. Patients' demographic characteristics did not vary between 1996 and 1998 (69% female, median age = 48 years).

Thyroid function tests were always the main cause of ordering: 1339 of the 1357 request forms in 1996 (99%), 1123 of the 1154 request forms in 1997 (97%), and 1071 of the 1118 request forms in 1998 (96%). During the 2-month intervention in 1997, the laboratory burden related to thyroid function tests fell markedly (Table 1). This was due to a significant decrease in the number of FT3 and FT4 tests as compared with those for TSH tests. Furthermore, a clear decrease in the relative proportion of FT3 tests, from 23% to 13%, was observed between 1996 and 1997. Interestingly, all these changes persisted 1 year later.

Analysis of prescription showed a change in the nature of thyroid function tests prescribed. The proportion of combined TSH+FT4+FT3 order forms fell between 1996 and 1997, from 51% to 19% of the total number of thyroid function order forms, while the

number of single TSH order forms increased respectively from 23% to 50% (Table 1). Conversely, the proportion of TSH+FT4 order forms did not vary between 1996 and 1997 (24% vs 28%). One year later, the same pattern of prescription as in 1997 was observed: TSH was the only test ordered in 512 specimens (48%), while TSH, FT4 and FT3 tests were ordered together in 197 cases (18%) (Table 1).

The intervention period (June–July 1997) was also marked by a decrease in the overall number of thyroid antibody tests performed from 625 in 1996 to 388 tests in 1997 (Table 2). Most of the prescriptions for anti-thyroid antibodies (85.6% in 1996, 83.5% in 1997, 81.1% in 1998) were concerned with TPOAb and TgAb, whereas the number of TRAb tests did not vary between these three periods (14.4% in 1996, 16.5% in 1997 ($P=0.57$) and 16.8% in 1998). We therefore focused our study on TPOAb and TgAb prescriptions.

Between 1996 and 1997, as shown in Table 2, the number of single TPOAb prescriptions increased significantly from 7.9% to 38.7% of the total number of TPOAb and TgAb order forms ($P<0.0001$). However, a combined TPOAb+TgAb test was still mainly ordered even though their relative number decreased significantly (87.3% and 50.6% respectively). Further analysis of these TPOAb+TgAb request forms showed unquestionably high levels of TPOAb (>500 U/ml), which resulted in unnecessary TgAb tests in about 25% of them. In addition, when patients with differentiated thyroid cancer (where TgAb is ordered to validate Tg assays) (13, 14) are excluded, TgAb alone was still ordered in 56 out of 592 (9.5%) of the overall TPOAb+TgAb order forms (1996–1998) (Table 2). Interestingly, similar patterns of antibody prescription were found 1 year later (Table 2).

Table 2 Number and nature of thyroid antibody requests during the 1996, 1997 and 1998 2-month study periods.

Nature of test	Pre-intervention period (1996) (no. of tests)	Intervention period (1997) (no. of tests)	Post-intervention period (1998) (no. of tests)
Total TPOAb tests	240	150	146
Total TgAb tests	295	174	141
Total TRAb tests	90	64	58
Total number of thyroid antibody tests	625	388	345
Single TPOAb	20	65	68
TPOAb + TgAb	220	85	78
Single TgAb*	12	18	26
Total number of TPOAb/TgAb* order forms	252	168	172

* Patients with differentiated thyroid cancer were excluded.

The total number of thyroid tests (thyroid antibodies) is indicated in the upper part. As thyroid tests may be prescribed in association, the combinations are noted and numbered in the lower part.

Appropriate use of thyroid laboratory tests

Table 3 shows an evaluation of the appropriateness of the thyroid laboratory test requests during the 2-month periods of 1996, 1997 and 1998. In 1997, during the intervention period, the rate of appropriate ordering was estimated at 72.4% (95% confidence interval: 69.8–75%) compared with 42.5% (95% confidence interval: 39.9–45.1%) in 1996 during the pre-intervention period ($P < 10^{-4}$). Notably, similar improvements were found 1 year later. The frequency of appropriate ordering depended on the main diagnosis referral whatever the observation period, with overall estimates in 1996 ranging from 5% (95% confidence interval: 2.2–8%) in cases of investigation of thyroid autoimmunity to 89% (95% confidence interval: 81.8–96.2%) in cases of cancer. In 1997 and 1998, these estimates significantly increased for all main diagnoses except nodule and cancer for which the frequency of appropriate ordering was unchanged ($P = 0.51$ 1996 vs 1997, and $P = 0.22$ 1996 vs 1998).

Discussion

Laboratory test relevance is crucial in medical practice because of its relation to optimal patient care and cost containment (15). Because of major technical improvements and concern for cost-effectiveness, considerable debate has focused on limiting the unnecessary laboratory use of thyroid function tests that are clinically irrelevant, as suggested in a recent report (16). Indeed, during a cross-sectional survey conducted in June 1995, we have shown an abnormally high proportion of T3 requests in the Paris hospital setting (12).

Many attempts have been made to change physicians' testing practices, with varying degrees of success. The wide variation in test-ordering behavior may be related

to the inability of physicians to estimate test performance characteristics, their inaccurate interpretation of diagnostic test results, and/or rapid advances in diagnostic technology (17). To optimize the practice of physicians, algorithms and guidelines for thyroid hormone and antibody testing have been published (4–9), including proposals from various medical societies (18–22). Some authors (4, 5, 7) recommend to evaluate thyroid function following a TSH based strategy algorithm. However, consecutive measurement of laboratory parameters and their added cost should be controlled. Therefore, like others (4, 5), we also propose a TSH-based strategy for the investigation of thyroid function with an automatic measurement of FT4 when TSH results are abnormal: this notion is still poorly understood since it was ordered in only 45 out of the 1157 request forms collected during the intervention period in 1997. The American Thyroid Association (18, 19) suggests that both TSH and FT4 tests are necessary, while others have proposed the assessment of thyroid hormone levels as the first test of thyroid function (23). Nevertheless, those recommendations appear to be insufficient, as reported by retrospective studies (24). However, simple administrative intervention, such as altering the order forms, have been shown to be successful in other areas (25–27).

To ensure the usefulness of the tests prescribed, i.e. to detail the nature of change, we assessed and compared the relevance of each given prescription over three periods of 2 months (1996, 1997 and 1998) on the basis of published data and local consensus (Fig. 1). A significant improvement was evidenced in each group of main diagnosis referral, except in group 3 (Table 3) where suitability was always over 80%. Since thyroid nodule and cancer patients are usually referred to specialists in endocrinology, the appropriateness of thyroid tests therefore remained high. In the other three prescription groups, as irrelevant tests were

Table 3 Estimation of the frequency of appropriate use of specific thyroid laboratory tests during the 1996, 1997 and 1998 2-month study periods, according to the main reason for medical referral.

	Pre-intervention period (1996)	Intervention period (1997)	Post-intervention period (1998)	<i>P</i> value* 1996 vs 1997	<i>P</i> value* 1996 vs 1998
Overall	577/1357 (42.5%)	836/1154 (72.4%)	790/1118 (70.7%)	$< 10^{-4}$	$< 10^{-4}$
(1) Thyroid function	398/876 (45%)	586/788 (74%)	571/853 (67%)	$< 10^{-4}$	$< 10^{-4}$
(2) Treatment	103/191 (54%)	119/141 (84%)	151/156 (97%)	$< 10^{-4}$	$< 10^{-4}$
(3) Nodule and cancer	65/73 (89%)	87/102 (85%)	50/62 (81%)	0.51	0.22
(4) Thyroid autoimmunity	11/217 (5%)	44/123 (36%)	18/47 (38%)	$< 10^{-4}$	$< 10^{-4}$

* Exact Fisher test.

documented during the pre-intervention period, we were able to demonstrate a significant improvement during the intervention period which was confirmed 1 year later (Table 3). Improvement was especially noticed on thyroid function tests, which represent the main test ordered. We aimed at reducing unnecessary thyroid tests, especially FT3 ones which should only be a second-line test if there is suspicion of T3 toxicosis in iodine-deficient areas (10) such as our region, or even a third-level measurement (4, 11, 12). We observed a clear decrease in the number of FT3 tests, especially because of a great reduction in the number of combined TSH + FT4 + FT3 requests (Table 1). Interestingly, improvement was confirmed 1 year later during a similar post-intervention period (June–July 1998) even though we did not change the design of the pre-intervention (1996) and post-intervention slips (1998). Thus, persistent change in medical practice could therefore be assessed (Table 1).

An overall improvement was also observed in the prescription of anti-thyroid antibodies, especially TPOAb and TgAb tests. Although still insufficient, this improvement was still observed 1 year later (Table 3). TRAb measurements, which only accounted for about 15% of all anti-thyroid antibody tests ordered, did not vary between each period (Table 2): actually, the TRAb test is not a first-line measurement and is usually ordered by specialists in endocrinology. Furthermore, we had not previously planned to look at the medical records and too few indications on the course of hyperthyroidism were given on the order sheets: therefore, we could not perform a pertinent analysis of their suitability. We thus focused our analysis on how TPOAb and TgAb tests were ordered.

Despite an increase in the ordering of single TPOAb tests between the 1996 and 1997 periods, we noticed a persistently high proportion of combined TPOAb + TgAb tests (Table 2), of which about 25% of the TgAb tests could have been avoided because of high TPOAb levels. Interestingly, 1 year later, the single TPOAb test was still similarly prescribed (Table 2): persistent change was also thus assessed. However, TPOAb, known for its highest sensitivity and specificity in autoimmune thyroid disease (9, 28), was still not prescribed in

about 9.5% (56/592) of our global cohort where TgAb was the only request (Table 2). We could therefore assess that these inappropriate request forms were related to the misuse of proper justifications.

This phase of implementation of prescription guidelines considered the appropriateness of test ordering and did not measure the costs and potential benefits of the intervention. The reduction of the total number of tests ordered results in a reduction in reagent costs, but the relationship may not be straightforward. In certain cases, reduction of laboratory activity does not result in proportional savings because of fixed costs or because of waste of reagent with a limited lifetime. In other cases, additional costs may be due to diagnoses missed or to repeat tests. As far as thyroid function tests are concerned, repeat testing may be done on frozen samples without patient recall or a prolonged length of stay. The economic assessment of practice guidelines, however, requires a full study which was outside the scope of this paper.

In conclusion, although the ordering of thyroid tests is not restricted to specialists in endocrinology but is usually done by residents and interns, we were able to observe a change in medical practice during the intervention period. Interestingly, this change persisted 1 year after the intervention period. Physicians, who accepted the fact that they were asked to justify their requests during a 2-month period, were therefore aware of cost containment without a noticeable change in patient care. Inadequately ordered thyroid biological tests were reduced in most instances. However, unsuitable justifications remained, related in particular to excessive thyroid hormone or anti-thyroid antibody requests, mainly FT3 and TgAb ones. An appropriate thyroid ordering laboratory cascade, as a help to the prescription, will be clearly stated in a new order form to avoid unnecessary and unhelpful expenses. This is under study in our hospital.

Acknowledgements

The authors are indebted to Professor J Orgiazzi for very helpful comments. They thank Mrs N Roman (Financial Manager of Saint-Louis Hospital) and Professor R Jian

(President of the 'Commission d' Evaluation' of Saint-Louis Hospital) for substantial contributions to the conception of this study. This work was partly presented in abstract form at the 25th Meeting of the European Thyroid Association.

References

- Nicoloff JT & Spencer CA. The use and misuse of the sensitive thyrotropin assays. *Journal of Clinical Endocrinology and Metabolism* 1990 **7** 553–558.
- Klee GG & Hay ID. Role of thyrotropin measurements in the diagnosis and management of thyroid disease. *Clinical Laboratory Medicine* 1993 **13** 673–682.
- Czarnocka B, Ruf J, Ferrand M, Carayon P & Lissitzky S. Purification of the human thyroid peroxidase and its identification as the microsomal antigen involved in autoimmune thyroid diseases. *FEBS Letters* 1985 **190** 147–152.
- Klee GG & Hay ID. Biochemical testing of thyroid function. *Endocrinology and Metabolism Clinics of North America* 1997 **26** 763–775.
- LoPresti JS. Laboratory tests for thyroid disorders. *Otolaryngologic Clinics of North America* 1996 **29** 557–575.
- Davey RX, Clarke MI & Webster AR. Thyroid function testing based on assay of thyroid-stimulating hormone: assessing an algorithm's reliability. *Medical Journal of Australia* 1996 **164** 329–332.
- Feldkamp CS & Carey JL. An algorithm approach to thyroid function testing in a managed care setting. *American Journal of Clinical Pathology* 1996 **105** 11–16.
- Volpe R. Rational use of thyroid function tests. *Critical Reviews in Clinical Laboratory Science* 1997 **34** 405–438.
- Feldt-Rasmussen U. Analytical and clinical performance goals for testing autoantibodies to thyroperoxidase, thyroglobulin, and thyrotropin receptor. *Clinical Chemistry* 1996 **42** 160–163.
- Knudsen K, Jorgensen T, Rasmussen S, Christiansen E & Perrild H. The prevalence of thyroid dysfunction in a population with borderline iodine deficiency. *Clinical Endocrinology* 1999 **51** 361–367.
- Klee GG. Clinical usage recommendations and analytic performance goals for total and free triiodothyronine measurements. *Clinical Chemistry* 1996 **42** 155–159.
- Vidal-Trecan G, Toubert M-E, Paycha F, Fulla Y, Fior R, Abella A *et al.* Les dosages hormonaux thyroïdiens à l'Assistance-Publique-Hôpitaux de Paris: prescriptions, coût et opinions des médecins qualifiés en endocrinologie. *Annales d'Endocrinologie* 1999 **60** 48–55.
- Spencer CA, Takeuchi M & Kazarosyan M. Current status and performance goals for serum thyroglobulin assays. *Clinical Chemistry* 1996 **42** 164–173.
- Spencer CA, Takeuchi M, Kazarosyan M, Wang CC, Guttler RB, Singer PA *et al.* Serum thyroglobulin autoantibodies: prevalence, influence on serum thyroglobulin measurement, and prognostic significance in patients with differentiated thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* 1998 **83** 1121–1127.
- van Walraven C & Naylor CD. Toward optimal laboratory use. Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits. *Journal of the American Medical Association* 1998 **280** 550–558.
- Roti E, Gardini E, Magotti MG, Pilla S, Minelli R, Salvi M *et al.* Are thyroid function tests too frequently and inappropriately requested? *Journal of Endocrinological Investigation* 1999 **3** 184–190.
- Solomon DH, Hashimoto H, Daltroy L & Liang MH. Techniques to improve physicians' use of diagnostic tests. *Journal of the American Medical Association* 1998 **280** 2020–2027.
- Surks MI, Chopra IJ, Mariash CN, Nicoloff JT & Solomon DH. American Thyroid Association guidelines for use of laboratory tests in thyroid disorders. *Journal of the American Medical Association* 1990 **263** 1529–1532.
- Becker DV, Bigos ST, Gaitan E, Morris JC, Rallison ML, Spencer CA *et al.* Optimal use of blood tests for assessment of thyroid function. *Journal of the American Medical Association* 1993 **269** 2736–2737.
- National Academy of Clinical Biochemistry (NCAB). Standards of laboratory practice symposium on thyroid-function testing. *Clinical Chemistry* 1996 **42** 119–192.
- Singer PA, Cooper DS, Levy EG, Ladenson PW, Braverman LE, Daniels G *et al.* Treatment guidelines for patients with hyperthyroidism and hypothyroidism. Standards of Care Committee, American Thyroid Association. *Journal of the American Medical Association* 1995 **273** 808–812.
- Vanderpump MP & Ahlquist JA. Consensus statement for good practice and audit measures in the management of hypothyroidism and hyperthyroidism. *British Medical Journal* 1996 **313** 539–544.
- De Groot LJ & Mayor G. Admission screening by thyroid function tests in an acute general care teaching hospital. *American Journal of Medicine* 1992 **93** 558–563.
- Bauer DC & Brown AN. Sensitive thyrotropin and free thyroxine testing in outpatients. *Archives of Internal Medicine* 1996 **116** 2333–2337.
- Tierney WM, Miller ME & McDonald CJ. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. *New England Journal of Medicine* 1990 **322** 1499–1504.
- Zaat JO, van Eijk JT & Bonte HA. Laboratory test form design influences test ordering by general practitioners in The Netherlands. *Medical Care* 1992 **30** 189–198.
- Durand-Zaleski I, Rymer JC, Roudot-Thoraval F, Revuz J & Rosa J. Reducing unnecessary laboratory use with new test request form: example of tumour markers. *Lancet* 1993 **342** 150–153.
- Mariotti S, Caturegli P, Piccolo P, Barbesino G & Pinchera A. Antithyroid peroxidase autoantibodies in thyroid disease. *Journal of Clinical Endocrinology and Metabolism* 1990 **71** 661–669.
- Ehrmann DA, Weinberg M & Sarne DH. Limitations to the use of a sensitive assay for serum thyrotropin in the assessment of thyroid status. *Archives of Internal Medicine* 1989 **149** 369–372.
- Braverman LE. Evaluation of thyroid status in patients with thyrotoxicosis. *Clinical Chemistry* 1996 **42** 174–178.
- Stockigt JR. Serum thyrotropin and thyroid hormone measurements and assessment of thyroid hormone transport. In *Werner and Ingbar's The Thyroid, A Fundamental and Clinical Text*, edn 7, pp 377–396. Eds LE Braverman & RD Utiger. Philadelphia: JB Lippincott, 1991.
- Ladenson PW. Optimal laboratory testing for diagnosis and monitoring of thyroid nodules, goiter, and thyroid cancer. *Clinical Chemistry* 1996 **42** 183–187.

Received 23 July 1999

Accepted 16 February 2000