An Evaluation of Six Solid-Phase Thyrotropin (TSH) Kits

By *W. G. Wood, D. Waller* and *U. Hantke*

Klinische Laboratorien und Poliklinik, Klinik für Innere Medizin (Direktor: Prof. Dr. med. P. C. Scriba) Medizinische Universität zu Lübeck, Lübeck

Summary: This article describes an objective evaluation of six thyrotropin (TSH) kits. One was a radioimmunoassay kit taken for comparison, three were immunoradiometric assays and one was an immunoenzymometric assay. The laboratory internal immunoluminometric assay for thyrotropin was used to measure the concentrations of thyrotropin in the kit standards using a standard curve of WHO 68/38 international reference preparation in serum from a thyrotoxic patient as matrix. The in-house assay was used to demonstrate the "sensitivity" to citrated plasma and the fact that kit standards could only measure "correctly" when used in its own kit.

The study was carried out in a "blind" way, the assayist and organiser not knowing from which of the four groups under test (blood donors — serum and citrated plasma, thyroliberin-test and thyroid outpatient clinic patients) the samples came until the study had been completed. The immunometric assay kits were able to differentiate statistically between euthroid and untreated hyperthyroid patients, although one IRMA kit (Kit F) had a large "grey zone" where both euthyroid and hyperthyroid patients overlapped.

Compound precision profiles covering the range 0 — 5 mU/l thyrotropin were good, a mean coefficient of variation under 5% within the range 0.5 — 5 mU/l being demonstrated by 3 immunometric assays. The immunometric assay kit with the most cumbersome methodology showed, as was to be expected, the worst precision.

The euthyroid ranges for thyrotropin were similar in 3 immunometric assay kits using the WHO 68/38 reference material as calibrator (Kits C and E, 0.25 — 3 mU/l) and correlated well with one kit using the 2nd IRP (NIBSC 80/558) as calibration material (Kit D 0.33 — 4 mU/l), although the results were around 30% higher in Kit D. The second kit (Kit A) using the 2nd IRP material as calibrator gave identical values with the kits using the WHO 68/38 reference thyrotropin-preparation for calibration purposes. In a further kit (Kit F) it was stated that both thyrotropin international reference preparations gave rise to identical serum values when used as calibrators.

The thyroliberin-test may have an additional role to play in monitoring returning pituitary function in thyrotoxic patients under treatment as the immunometric assay kits were easily able to measure a thyrotropin difference of 0.5 mU/l in the range 0 — 1 mU/l.

The conventional radioimmunoassay (Kit B) was unable to match the precision and sensitivity of the better immunometric assay kits in the range under 1 mU/l. It was also the only kit unable to differentiate statistically between euthyroid and hyperthyroid patients, and it is fair to say that this type of assay has had its day as far as the differentiation between the above two patient groups is concerned.

Beurteilung von sechs kommerziellen Festphasen-Testbestecks zur Bestimmung von Thyrotropin (TSH)

Zusammenfassung: Die vorliegende Arbeit beschreibt eine objektive Vergleichsstudie an sechs verschiedenen Thyrotropin (TSH)-Testbestecken, von denen drei auf immunoradiometrischer und ein Kit auf immunoenzymometrischer Basis beruhten.

Ein Radioimmunoassay-Kit wurde zu Vergleichszwecken mit den methodisch anderen Tests herangezogen. Der laborinterne immunoluminometrische Assay diente unter Verwendung einer mit dem internationalen Referenzmaterial WHO 68/38 im Serum eines Patienten mit Thyreotoxikose als Matrix hergestellten Standardreihe als „Bezugsassay" für die Bestimmung der Thyrotropin-Konzentrationen in den einzelnen Test-Standardreihen, um zu zeigen, daß die jeweiligen Standardreihen nur im eigenem Kit eingesetzt werden können, wenn „richtige" Ergebnisse erzielt werden sollen, und nicht von Kit zu Kit austauschbar sind.

Der als Blindstudie durchgeführte Vergleich beinhaltete Proben vier verschiedener Gruppen, und zwar Blutspender (Seren und Citratplasmaproben), Thyroliberin-Testseren (vor und nach 200 µg Thyroliberin (TRH) iv.) und Seren von Schilddrüsenpatienten der Poliklinik der Medizinischen Universität zu Lübeck. Bis zum Abschluß der Studie waren keiner der an der Durchführung oder Organisation beteiligten Personen bekannt, zu welcher der Gruppen die einzelnen Proben zuzuordnen waren.

Die immunometrischen Testbestecke waren in der Lage, statistisch zwischen euthyreoten und unbehandelten hyperthyreoten Patienten zu unterscheiden. Allerdings wies einer dieser Tests (Kit F) eine breite „Grauzone" auf, in der sich die Werte von eu- und hyperthyreoten Patienten überschnitten.

Die Präzisionsprofile im Bereich zwischen 0 und 5 mE/l Thyrotropin waren gut. Drei immunometrische Testbestecke konnten im Bereich 0,5 bis 5 mE/l Thyrotropin einen mittleren Variationskoeffizienten unter 5% aufweisen. Der in der Handhabung umständlichste Kit zeigte erwartungsgemäß auch die schlechteste Präzision.

Der Thyrotropin-Konzentrationsbereich für euthyreote Patienten war in zwei Kits auf Immunometriebasis, die das WHO 68/38 Referenzmaterial zur Kalibrierung verwenden, mit 0,25 – 3 mE/l gleich und zeigte in einem Kit eine gute Korrelation mit dem zwischen 0,33 und 4 mE/l liegenden Bereich eines mit dem 2. IRP (NIBSC 80/558) als Kalibriermaterial arbeitenden Test, sofern man den Faktor 1,3 für die Konzentrationseinheiten berücksichtigte. Der andere Kit, der den 2. IRP als Kalibrator benutzte, gab identische Werte mit den zwei Kits, die mit dem WHO 68/38 Referenzpräparat kalibriert wurden.

Der Thyroliberin-Test wird bei der Behandlung thyreotoxischer Patienten eine zusätzliche Hilfe zur Kontrolle der wiederkehrenden Hypophysenfunktion sein, da die immunometrischen Testbestecke leicht in der Lage sind, im Konzentrationsbereich von 0 bis 1 mE/l Unterschiede in der Größenordnung von 0,5 mE/l zu messen.

Der konventionelle Radioimmunoassay erreichte erwartungsgemäß im Konzentrationsbereich unter 1 mE/l nicht die Präzision und die Empfindlichkeit der besseren immunometrischen Tests und konnte als einziger Test nicht statistisch zwischen euthyreoten und hyperthyreoten Patienten unterscheiden. Es ist deshalb durchaus zulässig zu sagen, daß die neuen Thyrotropin-„Sandwichverfahren" in dieser Fragestellung den RIA-Test ersetzen werden.

Introduction

During recent months, an increasing number of sensitive thyrotropin assay kits has appeared on the market and several reports of the performance of such kits have already been published (1 – 4).

On the clinical side, the question has been raised as to whether such sensitive kits may alleviate the majority of thyroliberin (TRH) stimulation tests at present carried out to detect borderline hyperthyroidism (1 – 3).

The present study was carried out in order to provide an answer, if possible, to the latter question. The study was organised in a "blind" way, so that the analyst only received numbered sera and had no idea from what sort of patients or control groups these came. In addition, to increase the objectivity, each kit producer whose product was included in the study was invited to send a qualified technician to perform the test so that any claim of the producer that the persons carrying out the test were not acquainted with it, therefore the results were not optimal, could be excluded. This offer was taken up in the case of 4 from the 6 kits tested. The fact that the laboratory of the author was not concerned with the routine in-vitro thyroid-pituitary axis diagnosis removed the possibility of being influenced by the in-house routine method. In fact the laboratory-own method, an immunoluminometric assay for thyrotropin (5) was only used to compare the standards from all kits in an independent assay as well as to demonstrate the difference in results obtained from serum and citrated plasma samples.

It was not sufficient to have a sensitive kit, as far as the standard curve was concerned, and the nature of the study allowed an evaluation of each kit's ability to differentiate between euthyroid and untreated hyperthyroid patients, as well as to establish, when in a limited way, expected reference ranges for these two groups of patients.

Materials and Methods

The following thyrotropin kits were included in the study and allotted the codes used throughout the rest of the publication:

Kit A — Behring coated tube immunoradiometric assay — Riagnost TSH c. t. (Hoechst, Frankfurt am Main, D.)

Kit B — Henning radioimmunoassay TSH-RIASOL — overnight version (Henning-Berlin, Berlin).

Kit C — Henning immunoradiometric assay — TSH IRMAclon (Henning Berlin)

Kit D — Serono immunoradiometric assay — TSH-MAIACLONE (Serono, Freiburg im Breisgau, D).

Kit E — Boehringer-Mannheim immunoenzymometric assay overnight version — Enzymun Test TSH (Boehringer-Mannheim, Mannheim, D.)

Kit F — Boots-Celltech immunoradiometric assay — TSH IRMA (Röhm Pharma, Weiterstadt, D).

Kits A and D used the 2nd IRP-TSH preparation (NIBSC 80/558) for calibration, the rest, the WHO 68/38 reference material to calibrate the kit standards.

The house-internal immunoluminometric assay was performed as already published (5) and is included as method G, where used.

Table 1 shows the main points of the six kits used as far as components and total assay time are concerned. More detailed information on the assay schemes as performed in this study are to be seen in table 2.

Radioactivity was counted in a 16-well gamma spectrometer (counting efficiency 0.74—0.76) (Nuclear Enterprises, Sighthill, GB) data being reduced off-line with a CBM 8296 desk-top computer (Commodore Business Machines, Frankfurt am Main, D) using a spline function with automatic smoothing factor. All data were processed with this programme.

The colorimetric determination of kit E was performed on an ES-11 photometer coupled on- or off-line to an Epson HX-20 mini-computer (Boehringer-Mannheim).

The luminescence measurements were made on a LKB-1251 25-sample luminometer (LKB-Wallac, Turku, SF.), data reduction being off-line with the spline function described above.

All assays were performed exactly as stated in the kit instructions, except for kit D, where two wash steps were performed before radioactivity counting. In kit D, half the samples were decanted as stated in the instructions, the other half were carefully aspirated off using the set-up present in the laboratory and used for all routine radioimmunoassays performed in this laboratory. Kit E was performed with the normal reagents, but with a modified procedure to improve sensitivity. Kits A, C and E were not commercially available at the time of testing but should be on the market in the first half of 1985 according to manufacturers' claims in the form here tested.

Statistics were kept to a minimum, mainly using non-parametric methods, namely the Mann-Whitney U-Test for comparison of two unrelated samples and the Wilcoxon rank test for paired data. Other data are given in terms of the relevant percentiles and median. Correlations using the paired t-test were made on the euthyroid reference collectives, using kit A as the reference method, this because kit A had the lowest number of sera in which the thyrotropin-concentration was below the detection limit (one serum sample).

Tab. 1. Main details and components of the six kits tested.

Parameter	Kit A	Kit B	Kit C	Kit D	Kit E	Kit F
Method	IRMA	RIA	IRMA	IRMA	IEMA	IRMA
Calibrator-WHO No.	80/558	68/38	68/38	80/558	68/38	68/38
Kit instructions						
i. Assay details	n. a.	good	n. a.	good	n. a.	good
ii. Clinical data	n. a.	poor	n. a.	good	n. a.	good
No. of standards/	7	8	9	7	4 (7)	8
control sera	1	1	1	1	1	3
Range covered (mU/l)	0—50	0—50	0—50	0—50	0—6	0—256
No. of tubes/kit	100	100	100	100/250	100	120/240
Incubation (h)	2	22—24	2.5	2.15	20—22	3.7
No. of incubation/	1	2	2	2	3	2
wash steps	2	1	1	1 (2)	2	4
Radioactivity (kBq)	<300	<40	<160	<370	0	<220

Key:

In Kit E the number of standards covering the range 0.1—5.8 mU/l thyrotropin was 4. Three additional concentrations (see tab. 7) were incorporated, these being made up by dilution of kit standards with thyrotropin-zero serum.

Kits A, C and E were still in the final stages of development, hence the designation n. a. where data were not available.

In Kit D, one wash step is given; in this study, two wash steps were used to improve the dynamic range and precision.

Kits A and E used antibody coated tubes, Kits B, D and F used one of the antibodies immobilised onto a microparticular matrix.

Tab. 2. Assay details for each kit as performed in this study.

Kit A

200 µl serum/standard
100 µl tracer

Incubate 2 h on horizontal rotator at ambient temperature.
Aspirate and wash with assay buffer
Count for 60 s in gamma spectrometer

Kit B

100 µl serum/standard
100 µl anti-thyrotropin

Incubate overnight at ambient temperature

100 µl labelled thyrotropin

Incubate 4 h at ambient temperature

100 µl anti-rabbit γ-globulin

Incubate 30 min at ambient temperature

Centrifuge at 3000 g for 15 min at 4 °C

Aspirate supernate and wash with 1 ml assay buffer and repeat centrifugation and aspiration steps

Count precipitate for 1 min in gamma spectrometer

Kit C

200 µl serum/standard
100 µl labelled anti-thyrotropin

Incubate 30 min at ambient temperature

100 µl unlabelled anti-thyrotropin

Incubate 90 min at ambient temperature

1000 µl solid-phase anti-sheep γ-globulin

Incubate 30 min at ambient temperature

Centrifuge at 3000 g for 15 min at 4 °C
Aspirate supernate and count for 60 s in gamma spectrometer.

Kit D

200 µl serum/standard
100 µl labelled anti-thyrotropin

Incubate 2 h at ambient temperature

200 µl separation reagent

Allow to stand on magnet-rack for 5 min, decant supernate and allow tubes to stand inverted on absorbent paper

500 µl wash reagent

Repeat stand and wash steps, decant supernate and count precipitate for 60 s in gamma spectrometer

Kit E

200 µl serum/standard
1000 µl incubation buffer

Incubate overnight at ambient temperature

Aspirate and wash with 1 ml water

1000 µl peroxidase-labelled anti-thyrotropin

Incubate 4 h at ambient temperature

Aspirate and wash as above

1000 µl substrate buffer

Incubate 60 min at ambient temperature

100 µl stop-reagent

Measure tube contents in photometer at 405 nm

Kit F

100 µl serum/standard
100 µl labelled anti-thyrotropin

Incubate on orbital shaker for 2 h at ambient temperature
100 µl solid-phase anti-thyrotropin

Incubate 60 min as above

1000 µl assay buffer

Allow to stand 5 min

2000 µl separation reagent

Allow to stand 15 min

Aspirate of supernate and repeat the above wash/stand/aspiration steps and count precipitate for 120 s in gamma spectrometer

The experimental groups tested were as follows:

Group I — Sera from 51 non-screened blood donors.

Group II — Citrated plasma from 49 non-screened blood donors (not identical with those in group I).

Group III — thyroliberin-test sera before and 30 minutes after 200 µg Protirelin iv. (Antepan-200®, Henning Berlin), in 12 supposedly euthyroid laboratory personnel, 5 untreated hyperthyroid patients and 1 hypophysectomised patient.

Group IV — Sera from 41 patients attending the thyroid outpatient clinic in the Medizinische Hochschule Lübeck during November 1984. The patients were not selected, the only criterium being the consent of the patient to take part in the study.

All assays were performed in the first half of December 1984 using 2 batches of reagents for each method. Kits B, C, D and F were carried out by representatives of the respective firms to guarantee optimal results. Kit A was tested by the author and kit E by a technical assistent in the laboratory of the author.

Results

Figure 1 shows the distribution of results for groups I and II, the non-screened blood donors, serum and plasma values respectively. The 16th, 50th and 84th percentiles are shown together with the ranges. In group I 3 blood donors were found to be hyperthyroid and 1 blood donor to be hypothyroid, using serum thyroxine (T_4), thyroxine binding globulin (TBG) and the T_4: TBG ratio as criteria.

Kits A, C, D and E recognised all four cases correctly. Kits B and F recognised the hypothyroid case. Kit B was unable to recognise any of the hyperthyroid cases and Kit F failed to detect two of the hyperthyroid cases.
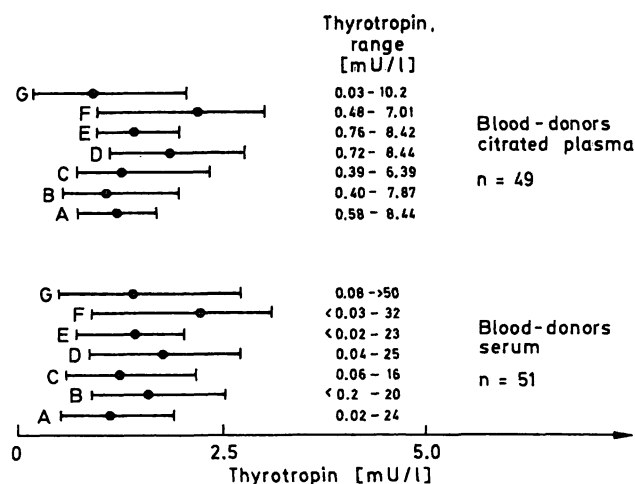
Fig. 1. This figure shows the ranges of thyrotropin measured in all 6 kits as well as in the in-house method (G). The 16th, 50th and 84th percentiles are shown diagrammatically. Values are given for both serum and citrated plasma. The concentrations refer to WHO 68/38, except for kits A and D, which refer to MRC 80/558 standard.

Figure 2 shows the results of group III, the thyroliberin tests on 12 laboratory personnel and on 5 non-treated hyperthyroid patients and 1 hypophysectomised patient. The hypophysectomised patient (Patient 3 of the second group in figure 2) showed paradoxical results. In Kits A and D, no response to thyroliberin was seen, whereas in the other kits, all showed a measureable thyrotropin increase after thyroliberin.

Figure 3 shows the distribution of the individual values from group IV, the thyroid outpatient clinic patients. Individual results are presented to allow a comparison between kits for each patient as measured in each of the 6 kits.

Table 3 shows the differentiation between euthyroid and untreated hyperthyroid patients together with the spread of results in both groups as measured in each of the 6 kits. The values for table 3 were taken
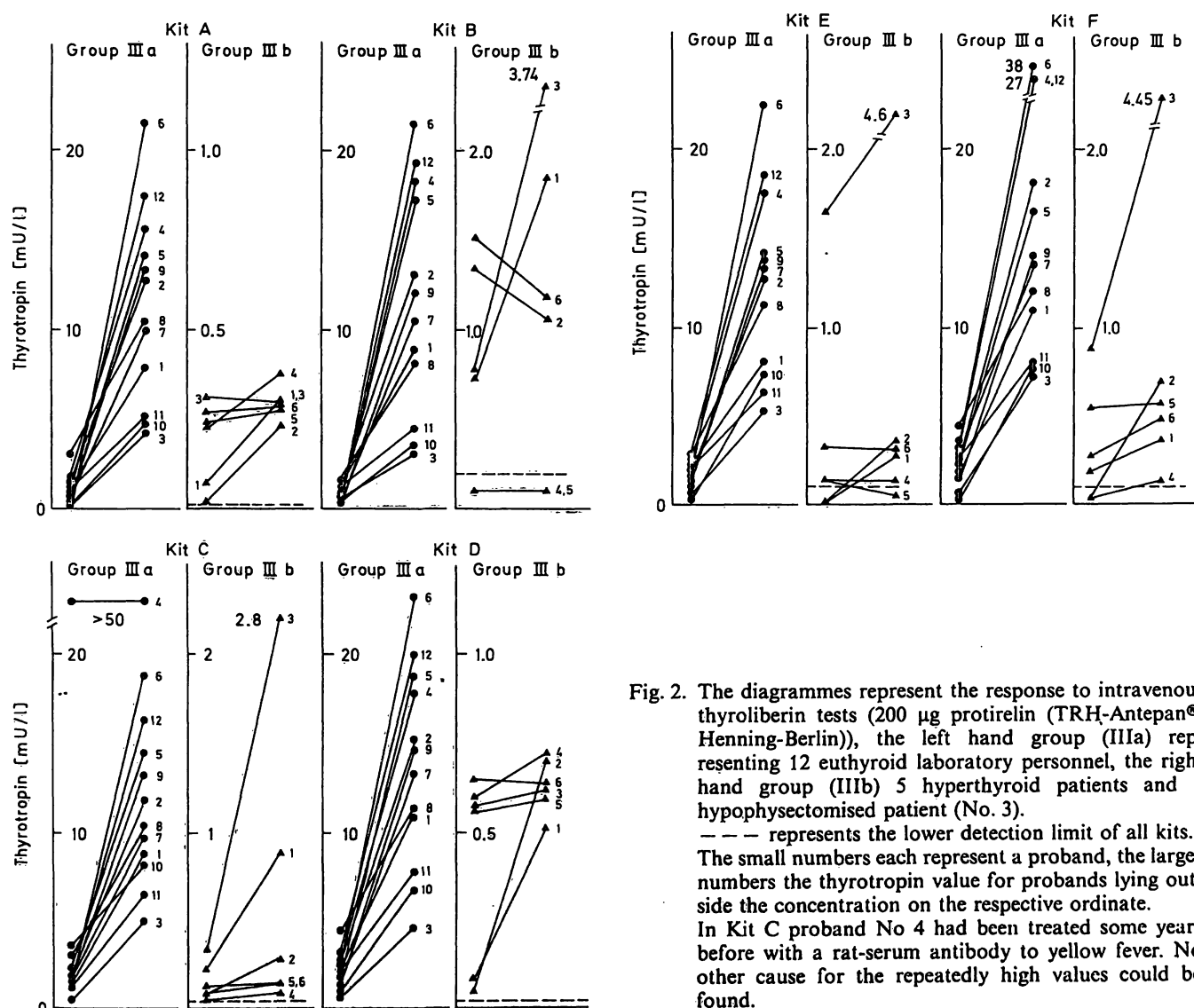

Fig. 2. The diagrammes represent the response to intravenous thyroliberin tests (200 μg protirelin (TRH-Antepan®, Henning-Berlin)), the left hand group (IIIa) representing 12 euthyroid laboratory personnel, the right hand group (IIIb) 5 hyperthyroid patients and 1 hypophysectomised patient (No. 3).

— — — represents the lower detection limit of all kits. The small numbers each represent a proband, the larger numbers the thyrotropin value for probands lying outside the concentration on the respective ordinate.

In Kit C proband No 4 had been treated some years before with a rat-serum antibody to yellow fever. No other cause for the repeatedly high values could be found.
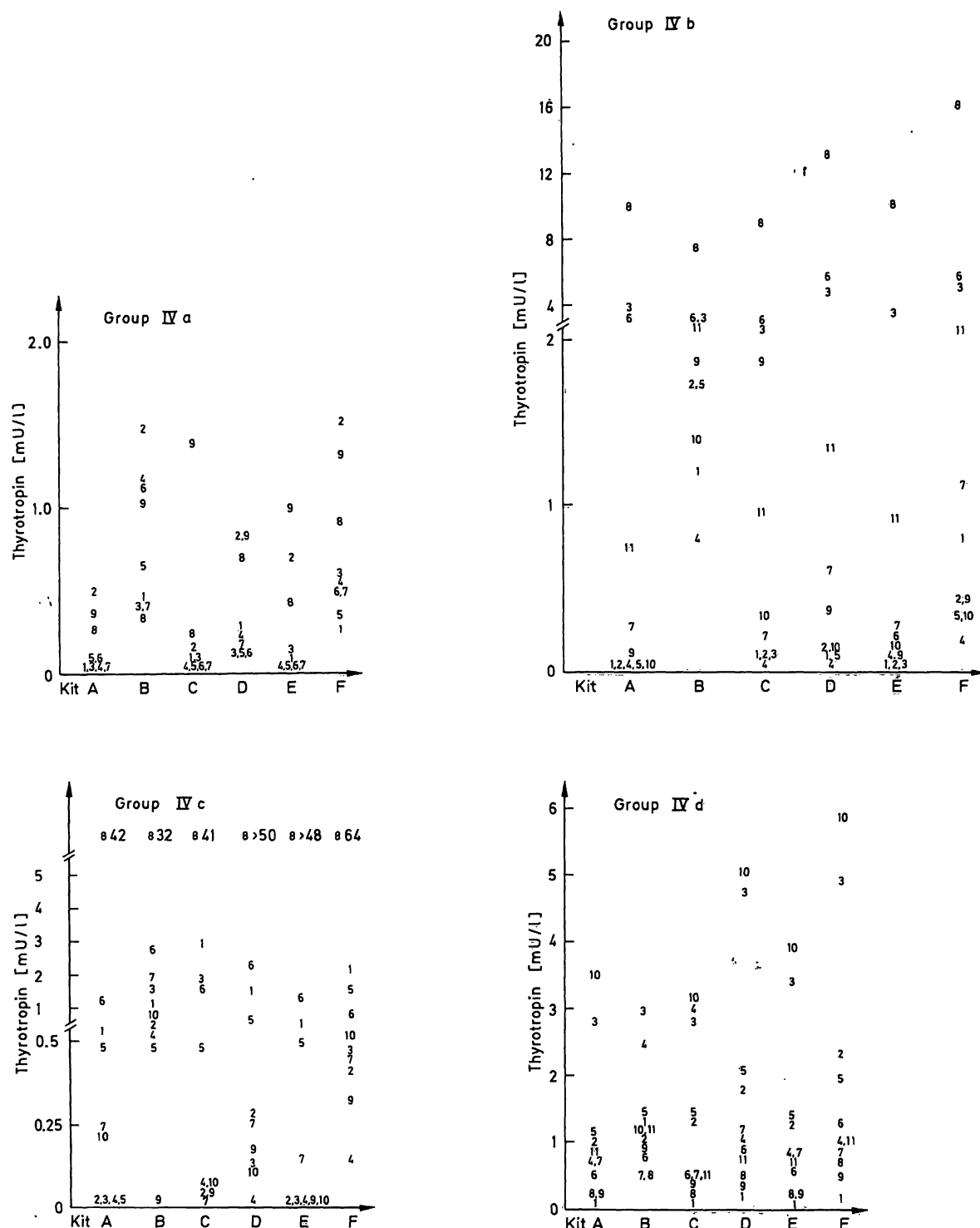
Fig. 3. The diagrammes show the individual results of each patient in Group IV, the thyroid outpatient clinic, divided into the following 4 groups:

IVa. Goitre.
   Patients 1 & 2 grade Ia, patients 3 & 4 grade III and patient 8 recurrence. No further details were given for patients 5, 6, 7 & 9.

IVb. Graves' disease and thyroiditis — histologically confirmed disease. Patients 4 and 6 had focal thyroiditis and Hashimoto thyroiditis, all others having Graves' disease.

IVc. Treated and non-treated hyperthyroid patients — no histological data. Patients 1,5 & 6 were under treatment, patient 8 had received an overdose of thiamazol (3 × 20 mg/d Favistan®), the remainder being untreated at the time of blood sampling.

IVd. These patients were seen as a miscellany, the composition of the group being depicted below:
   Patient 1 — thyroid adenoma, Patient 2 — Diarrhoea, Patient 3 — borderline hypothyroidism, Patient 4 — thyroid dysfunction, Patient 5 Lupus erythematosus, Patient 6 — medullary thyroid carcinoma (serum calcitonin above 1 mg/l!), Patients 7 & 10 hypophysectomy (patient 7 = patient 3 in fig. 2, Group IIIb), Patient 8 — Thyroidectomy post op., Patient 9 — euthyroid control, Patient 11 — pituitary insufficiency.

Tab. 3. Differentiation between eu- and untreated hyperthyroid states using serum samples.

Parameter	Kit A	Kit B	Kit C	Kit D	Kit E	Kit F
Euthyroids (n = 50)						
Range mU/l	0.26–2.83	n. d.–2.85	0.27–3.03	0.33–3.81	0.13–3.01	n. d.–3.97
16th percentile	0.63	0.62	0.64	0.97	0.73	0.91
50th percentile	1.16	1.07	1.40	1.83	1.44	2.02
84th percentile	1.87	1.71	2.09	2.65	2.02	3.07
Hyperthyroids (n = 24)						
Range mU/l	n. d.–0.22	n. d.–2.41	n. d.–0.28	n. d.–0.33	n. d.–0.34	n. d.–0.72
16th percentile	n. d.	n. d.	0.04	n. d.	n. d.	0.18
50th percentile	0.02	0.76	0.06	0.10	n. d.	0.37
84th percentile	0.05	1.56	0.16	0.17	0.08	0.49
Statistics						
z-Value	6.93	1.43	6.92	6.92	6.89	6.15
Overlapping results	0	24	1	0	1	23

Key:

n. d. — results lying below lower detection limit of the assay.

For the statistics the *Mann-Whitney* U-test with z-transformation. Z-values above 2.00 show a statistically significant difference between the two groups below the 5% level.

Overlapping results denote the number of hyperthyroid patients lying within the euthyroid range.

from groups I, III and IV, i. e. only serum values, after confirmation of a clinically euthyroid/hyperthyroid state (serum thyroxine, triiodothyronine, thyroxine binding globulin and thyroglobulin values as well as the thyroxine: thyroxine binding globulin ratio). All kits except Kit B were statistically able to differentiate between the two groups, although the degree of overlapping varied greatly from no overlap (Kits A and D) to an overlap of 23/24 cases in Kit F and 24/24 cases in Kit B.

Tab. 4. Control and test sera as measured in all kits.

Serum	Kit A	Kit B	Kit C	Kit D	Kit E	Kit F
HA	9.66	12.7	9.79	13.0	10.0	10.5
HB	2.74	3.42	2.71	4.40	3.00	2.84
HC	1.51	2.41	1.84	2.26	1.96	1.60
NS 14	0.05	n. d.	0.08	0.29	0.10	0.22
TSHFS	0.03	n. d.	0.16	0.24*	0.34	0.05
9040	0.02	0.69	0.06	0.12	n. d.	0.62

Key:

n. d.	— result below the lower limit of detection of the assay.
NS 14	— Affinity extracted human serum using solid-phase polyclonal anti-thyrotropin
TSHFS	— as NS 14 but using an alpha-specific monoclonal anti-thyrotropin
9040	— Serum from an untreated hyperthyroid patient with a thyroxine:thyroxine binding globulin ratio above 10.

Table 4 shows the performance of all kits in measuring 5 poolsera and a hyperthyroid patient serum (9040), the latter having a thyroxine: thyroxine binding globulin ratio above 10. Sera NS 14 and TSHFS had been rendered "TSH-free" with affinity chromatography using polyclonal (NS 14) and alpha-specific monoclonal antibodies (TSHFS).

Figure 4 shows the compound precision profiles (5, 6) for all samples measured within the concentration range lower detection limit to 5 mU/l thyrotropin, each decile representing 0.5 mU/l. The results are taken from 10 assays from each kit.

Table 5 shows the intra-assay coefficients of variation for different sera and kit control sera covering the range up to 12 mU/l thyrotropin.

Table 6 shows the dynamic range for the immunometric methods (all except Kit B) the dynamic range being expressed as the ratio between the counts in the standard in question and the counts in the zero standard. The mean and coefficient of variation are derived in each case from 10 assays. The larger the coefficient of variation, the more variable the shape of the standard curve. In the case of Kit D, the effect of washing the precipitate twice, instead of once as suggested in the kit protocol, almost doubled the dynamic range and improved the precision.
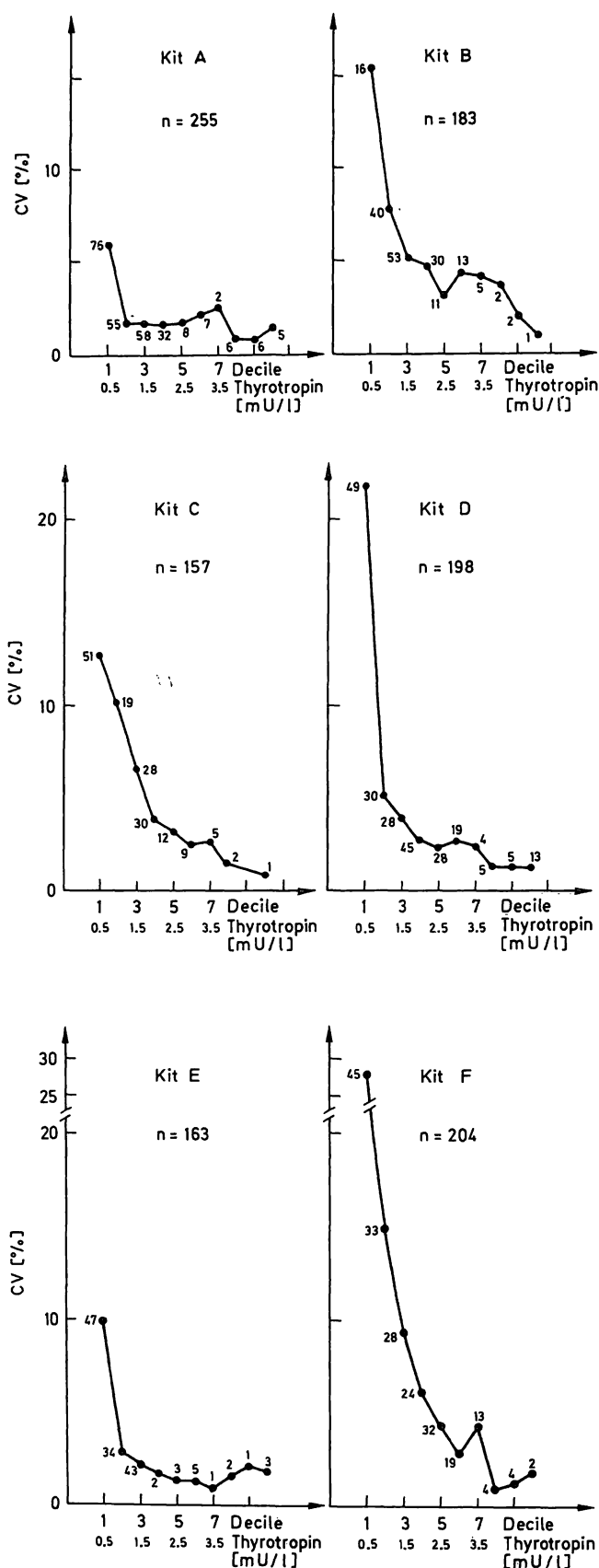
Fig. 4. The diagrammes show the compound precision profiles from 10 assays with each kit, using the values lying between the lower detection limit of the assay and 5 mU/l thyrotropin. The deciles represent concentration steps of 0.5 mU/l and the numbers by each point the number of data (from duplicate determinations) used in each decile. The total number of duplicates used is also shown for each kit.

Tab. 5. Interassay coefficients of variation in the six kits tested.

Kit/Serum	n	Mean (mU/l)	CV (%)
Kit A			
HA	10	9.66	4.19
HB	10	2.73	3.63
HC	10	1.51	2.73
Kit Cont (4.0)	10	4.23	2.73
NS 14	10	0.05	35.0
Kit B			
HA	10	12.7	6.86
HB	10	3.42	7.03
HC	10	2.41	7.37
Kit C			
9100	10	1.51	3.76
9107	10	12.4	1.56
9201	10	3.62	2.84
NS 14	10	0.08	52.3
Kit D			
HA	10	12.7	10.3
HB	10	4.54	10.2
HC	10	2.08	12.3
Kit Cont (5.0)	10	5.18	8.60
NS 14	10	0.32	16.9
Kit E			
9100	10	1.74	4.69
Kit Cont (5.2)	10	5.23	4.51
NS 14	10	0.34	9.84
Kit F			
Kit Cont A (0.5)	16	0.94	10.6
Kit Cont B (2.2)	16	2.44	15.3
Kit Cont C (11.5)	16	12.4	4.50
NS 14	28	0.22	49.8

Key:

NS 14 — The properties of this serum are described in table 4. Sera HA, HB and HC are identical with those in table 4.

The in-house luminescence immunoassay showed that when the kit standards were measured as unknowns, the concentrations measured differed widely from those given, especially at low concentrations of thyrotropin and where the matrix had a low protein content. The version of the in-house assay used here was chosen because of its sensitivity to matrix effects and show the inherent dangers of trying to use standards from kit x in kit y.

Table 7 shows the correlation of the euthyroid patients in table 4 using Kit A as x-value, this because Kit A had only a single serum in the whole study which lay below the lower detection limit, no other kit being able to match this performance.

Tab. 6. Kit dynamic ranges for immunometric methods — expressed as counts in the standard in question (Cs) divided by counts in the zero standard (Co).

	Standard (mU/l)	Mean Cs/Co	CV (%)	Total counts	Zero counts
Kit A	0.15	3.15	17.6	98 450	88
	0.50	7.02	15.4		
	1.50	16.7	14.0		
	5.00	48.7	15.6		
	15.0	140	16.5		
	50.0	401	8.30		
Kit C	0.39	1.84	5.25	134 000	780
	0.78	2.30	11.4		
	1.56	3.01	9.97		
	3.12	5.78	7.72		
	6.25	10.3	7.71		
	12.5	19.4	8.98		
	25.0	35.3	9.22		
	50.0	61.5	8.43		
Kit D	0.50	2.20	14.2	209 200	440
	1.00	3.63	8.62		
	5.00	13.7	12.1		
	10.0	26.4	10.3		
	30.0	77.4	10.2		
	50.0	124	8.54		
Kit E	0.10	1.90	15.5	—	0.042A
	0.25	3.45	10.9		
	0.50	5.07	5.42		
	1.00	9.20	6.70		
	1.80	16.6	7.80		
	5.80	40.3	5.78		
Kit F	0.09	1.31	19.2	55 200	97
	0.29	2.19	20.9		
	0.91	4.62	23.2		
	4.37	15.9	26.4		
	15.7	52.0	26.2		
	63.8	166	27.7		
	256	298	27.7		

Key:

Total counts represents the radioactivity given to each tube expressed in counts per minute. For Kit E is this irrelevant.

Zero counts represents the counts per minute bound in the zero standard not corrected for unspecific binding. For Kit E this value represents the absorbance at 405 nm.

Kit D — The precipitate was washed twice before counting.

Kit F — The precipitate was counted for 120 s in contrast to other kits (60 s).

Tab. 7. Correlation of values from 50 clinically euthyroid patient sera — comparison with Kit A (x-values).

	r	ayx	byx	axy	bxy
Kit B	0.347	0.677	0.423	0.909	0.284
Kit C	0.651	0.546	0.718	0.398	0.591
Kit D	0.938	0.342	1.25	−0.090	0.703
Kit E	0.857	0.207	0.992	0.177	0.741
Kit F	0.811	0.223	1.52	0.303	0.431

Note:

The values are for the regression lines $y = ax + b$ and $x = a + by$

All values in Kit B which were under the detection limit of the assay (n = 10) were entered as 0.01 mU/l.

Discussion

Although it was and is not the aim of this study to exalt or condemn, it is clear that the days of the conventional radioimmunoassay are numbered where a rapid discrimination between euthyroid and hyperthyroid patients is concerned, even when such assays claim to be supersensitive, but take 4 days to produce results (4). The specificity of radio-immunoassays for hormones such as thyrotropin is questionable as all fragments and aggregation products containing the epitopes recognised by the antibody are bound and can give rise to falsely elevated results.

The new immunometric assays based completely or partly upon monoclonal antibodies herald a new era in thyroid diagnosis and surveillance. At the same time, new problems have been presented to the kit producer, especially in the production of an acceptable thyrotropin-free matrix for the standards, and it may be, that the tenets upheld by this group for several years, (8, 9) i. e. thyrotropin standards dissolved in thyrotropin-free human serum, may now no longer be practical in assays capable of measuring under 0.05 mU/l thyrotropin. That the standard matrix is different in each kit can be seen in the results from electropherograms which have not been presented here, but in which the total protein and albumin: globulin ratios were often highly abnormal. One kit (Kit F) used a non-human based matrix for the standards.

The excellent correlation between, for example Kit A and Kit D for the euthyroid patients (tab. 7), does not detract from the fact that the slope of the regression curve shows results which differ by 30%. Such a difference in values was not expected as both kits used standards calibrated against the 2nd IRP — (NIBSC 80/558). This shows that kits must be optimally calibrated to give similar values, even if the matrix is not "pure" thyrotropin-free human serum and the secondary (calibration) standard not identical with the WHO reference preparation, which is the case in all kits. The degree of "thyrotropin-freeness" must be seen with respect to the epitopes on the antibodies used, and it may be that a matrix is thyrotropin-free in one system, but not in another (see tab. 5 and 6 — sera NS 14 and TSHFS).

One of the main points emerging from this study was the excellent precision and sensitivity of certain kits coupled with a short incubation time of as little as 2 h (Kit A). Even setting up an assay with 300 tubes using this kit in which the pipetting of serum took 110 minutes did not influence the kit performance in

terms of the standard curve and control serum values from the end of the assay read off the standard curve at the beginning of the assay and vice versa. Another interesting result was that the overnight version of the enzyme-labelled assay (Kit E) was as sensitive as the immunoradiometric tests. This allows the complete thyroid in-vitro diagnosis (with the possible exception of thyroglobulin) to be performed with commercially available non-radioactively labelled techniques. Even though the range of Kit E was relatively small (0−6 mU/l thyrotropin) in the overnight version, this is large enough for the initial differentiation between hyperthyroid, euthyroid and hypothyroid patients. The same-day version of Kit E allows the range 0.2−50 mU/l thyrotropin to be covered.

The hypophysectomised patient gave rise to interesting thyrotropin values before and after thyroliberin-test. Although all immunometric tests claimed to measure "intact-thyrotropin", the degree of "intactness" was different as Kits A and D showed no response to thyroliberin, whereas Kits D, E and F showed a marked response of "thyrotropin" to thyroliberin-stimulation. These results remain paradoxical, especially in terms of the therapy chosen, but serve as an example that even in these kits "thyrotropin" is not always thyrotropin! (10)

The kit with the most complicated and somewhat clumsy methodology (Kit F) gave rise to the worst results with respect to inter- and intra-assay precision, even in the hands of the technician from the firm offering the kit (fig. 4, Tab. 5−7). These results confirm earlier reports with this kit (11). Kit F exaggerated its claims to sensitivity as it was unable in many cases to differentiate between untreated hyperthyroid patients and clinically euthyroid patients and occupied the penultimate place before the conventional RIA (Kit B) with regard to this point. The sensitivity of the assay on paper cannot be treated separately from the ability to differentiate between the groups in question. The differing results from Kit F may reflect the non-human standard matrix to some extent.

The kit with the largest dynamic range (signal ratio highest standard: zero standard (signal to noise ratio)) was Kit A, followed by Kit E (in the range 0−6 mU/l thyrotropin). Both these kits had a low signal in the zero standard (see tab. 7). Table 7 also shows the main drawback of the immunoradiometric technology, namely the large amounts of radioactivity present in the kits, this being in certain cases 10-fold that in the conventional RIA (Kit D 209 200 counts/min · 100 μl, Kit B 22 400 counts/min · 100 μl), a

point which must be taken into account when operating with a licence allowing a limited amount of ^{125}I. The immunoradiometric assay kit with the lowest activity was Kit F which had only a quarter of the radioactivity in Kit D. The precision profiles showed that kit A had the best precision profile although Kits A, D and E had excellent precision in the range 0.5−5 mU/l thyrotropin (fig. 4). This confirms the data published by Hunter (6) who showed that immunometric assays should give a better precision profile than radioimmunoassays.

On the other hand, the most stable curve form (tab. 6) was exhibited in kits C, D, and E. The latter point may play a role in curve-fitting procedures where a computer software package is used in which the user has little or no chance to alter the routines to suit the assays used. Figure 1 shows that the kits were not affected by calcium-free plasma, in contrast to the in-house assay (Method G) which reacted allergically to plasma when using a certain labelled second antibody. Replacement of the label removed the anomalous effects without affecting the sensitivity of this assay. This test was carried out as it is often stated that certain antibodies are "plasma" or "calcium" sensitive (12).

All immunometric methods gave similar euthyroid reference ranges, when the calibration preparation is taken into account. This range appears to be smaller than that given for the RIA (13−15) and to that given in some of the kit protocols (Kits D and F). As Kits A, C and E were still not available commercially, the protocols being somewhat ad-hoc in nature and not containing relevant clinical data, it was impossible to compare the ranges here found with those in the kit instructions. The narrower (and lower) euthyroid ranges may reflect the improved specificity of the kits in measuring "intact" thyrotropin.

Figures 2 and 3 were included to show that each kit retains some "individuality" with respect to serum values from each patient, which means that the established practice of choosing a kit as the laboratory method still holds, even when the overall results tend towards more homogeneous and comparable reference ranges. In figure 2 patient 4 of the euthyroid group gave anomalous results inasmuch as the thyroid status was normal but the basal thyrotropin with kit C was above 50 mU/l! One possible explanation could be that the proband had been treated with rat-serum for yellow fever several years before, and that interfering antibodies were present. This was partially supported by the fact that when the incubation buffer was changed this anomaly disappeared,

basal thyrotropin being 0.6 mU/l, the stimulated value 12.9 mU/l, which agreed with the results from other kits. As a result of these findings, the assay buffer of Kit C was changed by the producer accordingly.

The era of monoclonal antibodies should make the long-term comparison of results better inasmuch as the reagents are standardised and available in large amounts suitable for production over several years. Moreover it is possible to choose epitopes on the thyrotropin molecule which will only measure the intact molecule. The change from radioactive to non-radioactive labelling techniques should allow the change to alternative methodology without having to worry about comparison of results.

The need for a "conventional" thyroliberin-test will be much reduced due to increased sensitivity and better differentiation of eu- and hyperthyroid patients with the new immunometric assays for thyrotropin. The degree of the reduction in thyroliberin-tests must be awaited in terms of clinical practice, where other in-vitro and clinical parameters are used to make up

the complete diagnostic picture. Finally, the thyroliberin-test receives a new and interesting role, namely the monitoring of returning pituitary-thyroid axis interplay under therapy of thyrotoxic patients, as all immunometric kits were able to measure a thyrotropin difference of 0.5 mU/l in the range 0 − 1 mU/l. With the advent of sensitive and precise thyrotropin kits a new era in in-vitro thyroid surveillance is dawning.

Acknowledgement

References

1. Seth, J., Kellett, H. A., Caldwell, G., Sweeting, V. M., Becket,, G. J., Gow, S. J. & Toft, A. D. (1984) Brit. Med. J. *289*, 1334−1336.
2. Allen, K. R. & Watson, D. (1984) Clin. Chem. *30*, 503.
3. Weeks, I. & Woodhead, J. S. (1984) Chemiluminescent acridinium esters as labels in immunoassay. In: Analytical applications of bioluminescence and chemiluminescence, (Kricka, L. J., Stanley, P. E., Thorpe, G. H. G. & Whitehead, T. P., eds.) Academic Press, London, pp. 185−188.
4. Mewissen, B. & Hehrmann, R. (1985) Akt. Endokr. Stoffw. *6*, 1−6.
5. Gadow, A., Wood, W. G. & Scriba, P. C. (1984) Akt. Endokr. Stoffw. *5*, 13−21.
6. Hunter, W. M. (1982) Recent advances in radioimmunoassay and related procedures. In: Radioimmunoassay and related procedures in medicine 1982, IAEA, Vienna, pp. 3−21.
7. Wood, W. G. (1983) Ärztl. Lab. *29*, 311−316.
8. Erhardt, F. W., Marschner, I., Pickhardt, C. R. & Scriba, P. C. (1973) J. Clin. Chem. Clin. Biochem. *11*, 381−387.

9. Erhardt, F. W. & Scriba, P. C. (1974) Ärztl. Lab. *20*, 191−198.
10. Beck Peccoz, P., Amr, S., Menezes-Ferreira, M. M., Faglia, G. & Weintraub, B. D. (1985) N. Engl. J. Med. *312*, 1085−1090.
11. Panitz, N. (1984) A German evaluation of a TSH-IRMA. In: A simple sensitive TSH immunoradiometric assay for the discrimination of thyroid function − Proceedings of a Boots-Celltech Symposium, *26. 03. 84* Edinburgh, GB. pp. 28−33.
12. Hunter, W. M. (1978) In: Handbook of experimental immunology (Weir, D. M., ed.) Vol. 1 Immunochemistry, Chapter 14 − Radioimmunoassay, Blackwell Scientific Publications, Oxford, London.
13. Meinhold, H. (1983) Nucl. Med. *22*, 196−203.
14. Pfreundschuh, M. & Hüfner, M. (1977) Acta. Biol. Med. Germ. *36*, 101−106.
15. Wood, W. G. (1980) NUC Compact *11*, 60−64.

Priv.-Doz. Dr. W. G. Wood
Klinische Laboratorien
Klinik für Innere Medizin
Medizinische Universität zu Lübeck
Ratzeburger Allee 160
D-2400 Lübeck 1