Comparison of Three Commercial Kits for Determination of Triiodothyronine (T_3) by Radioassay*†

Robert O. Pick, Douglas Daniels, Jerome A. Waliszewski, and Martin L. Nusynowitz

William Beaumont Army Medical Center, El Paso, Texas

Three T_3 radioassay kits were evaluated from the standpoint of accuracy, specificity, precision, reproducibility, sensitivity, technical performance factors, and cost. Whereas each has advantages and disadvantages in regard to each of these factors, all three kits provided acceptable results and were deemed satisfactory for use in the clinical nuclear medicine laboratory.

A number of clinically important thyroid disease states in which serum T_3 concentration plays a central role have been recognized in recent years. Among these are T_3 thyrotoxicosis arising de novo, T_3 thyrotoxicosis occurring after radioiodine therapy of the thyroid gland for the hyperthyroidism of Graves' disease, and chronic thyroiditis with euthyroidism, where a relative if not absolute increased production of T_3 compensates for a fall in thyroxine concentration (*1*). The determination of serum T_3 concentration has assumed, therefore, great practical importance in the evaluation of thyroid disorders. The advent of commercially available kits for the measurement of serum T_3 by radioassay necessitates their comparative evaluation to ascertain if requisite validity obtains for routine use in the nuclear medicine laboratory. We have evaluated three commercial T_3 radioassay kits from the point of view of accuracy, specificity, precision, reproducibility, sensitivity, and technical factors and cost, and report our findings herein (*2*).

Materials and Methods

Serum was obtained from 23 healthy young men and used in each of the three assays. The kits studied were manufactured by Pantex, Inc., Malibu, CA 90205 (kit P), Mallinckrodt/Nuclear, St. Louis, MO 63147 (kit M), and Nuclear Diagnostics, Inc., Troy, MI 48084 (kit N). Kit M uses magnesium-8-anilino-1-naphthalene sulfonate (ANS) to displace T_3 from serum and employs a resin strip for the separation of unbound from bound T_3; kit P utilizes a direct double-antibody technique and also employs ANS; kit N uses an alcoholic extraction procedure for the separation of T_3 from serum and employs a resin column for the separation of unbound from bound T_3.

The assays were performed according to instructions supplied by the manufacturer, with one exception: duplicate determinations were made on duplicate extractions using kit N instead of making "duplicate" determinations on the same extraction. All determinations were made in duplicate on each of two days by the same technologist using the same 23 serums for each of the three kits. For tests of accuracy, reagent grade triiodothyronine (Sigma Chemical Co., Catalogue T2377) was dissolved in a mixture of 100% ethanol and 12 N NaOH and added to pooled serum; recovery of added T_3 was determined in each assay. In addition, within-run precision was ascertained by performing 10 determinations on the same sample in one run. After the initial evaluation the results obtained with kit N were quite poor. The manufacturer was contacted and his representatives informed us that the instructions supplied were not explicit; the kits required performance in plastic tubes instead of the glass tubes that we employed for the evaluation. The parameters of performance for kit N were therefore repeated in plastic tubes and the results reported were obtained with the repeat determination of kit N. Standard statistical methods were employed (*3*).

Results

Accuracy, the extent to which the mean of an infinite number of measurements of a substance agrees with the exact amount of the substance present (*2*), was evaluated by testing for agreement with accepted values and by determining the percentage of T_3 added to a serum recovered by the assays. Table 1 shows the mean, standard deviation, and 95% confidence interval of normal for each kit. All three kits showed similar results, which are in good agreement with values reported in the literature. Also shown in Table 1 are the initial results obtained with kit N; note the lower values

For reprints contact: M. L. Nusynowitz, MD, PO Box 70014, William Beaumont Army Medical Center, El Paso, TX 79920.

*The mentioning of trade names or manufacturers does not constitute approval of or endorsement by the US Government.

†Presented at the 22nd Annual Meeting of the Society of Nuclear Medicine, Philadelphia, PA, June 1975.

TABLE 1. Results of T_3 Radioassays From Analysis of Serum in Normal Men

Kit (No. of serums)	Mean T_3 ng/100 ml	s.d.	95% confidence interval
P(23)	116	39	38-193
M(23)	112	31	51-173
N(23)	77	28	22-132
N(20)*	104	36	29-179

**Second determination on 20 different individuals.*

obtained when the determinations are performed in glass tubes. Accuracy, as judged in terms of recovery data, showed that recovery averaged 95% and 87% for kits P and M, respectively, whereas mean recovery of kit N was only 57% (Table 2).

Specificity, the extent of freedom of the test from interference by substances other than the one intended to be measured, was evaluated by dilution of hyperthyroid serums and comparison with standard curves. Figure 1 shows the standard curve obtained with kit P. Also plotted are count rates expressed as percentage of the zero concentration standard and the T_3 concentrations of dilutions of serums from two hyperthyroid patients. These concentrations are computed by multiplying the dilution factor by the concentration obtained from the standard curve for the undiluted serum. The three regression equations obtained from the points in the linear region for the standards and the patients' values are identical; there is no significant difference in slopes or intercepts among them, thus fulfilling a necessary (albeit not sufficient) condition for specificity of kit P. Similar findings were obtained with kits M and N (Figs. 2 and 3). Kit N measures percentage free instead of percentage bound, so that the curve slopes upward instead of downward as with kits M and P.

Precision, the extent to which a given set of measurements of the same sample agrees with the mean, was determined by performing 10 replicate determinations on the same serum in the same run. The coefficients of variation for the 10 determinations were 19% for kit P, 13% for kit M, and 16% for kit N (Table 3).

Reproducibility, the extent to which an estimate is duplicated upon repeated measurement, was evaluated by an analysis of variance of a randomized complete block hierarchial design, for within-run and between-run variation (*3*) (Table 4). Variance between duplicate de-

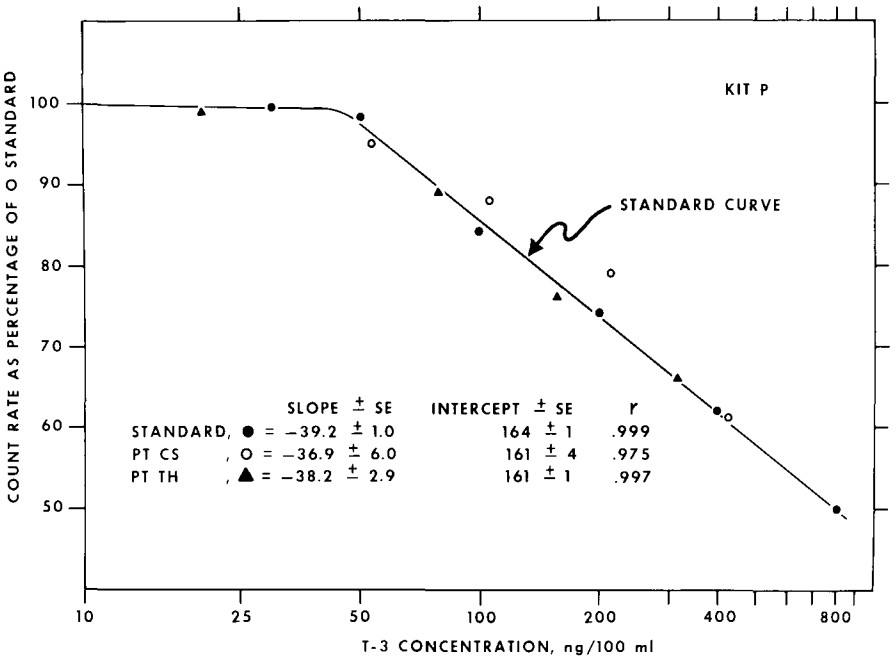

FIG. 1. Standard curve and values from dilutions of serums from two hyperthyroid patients by using kit P. Slopes and intercepts are determined from regression lines by using values obtained from standard and serum dilutions separately.

TABLE 2. Comparison of T_3 Radioassay Kits—Recovery of T_3 Added to Pooled Serums

T_3 added (ng/100 ml)	Kit P		Kit M		Kit N	
	Recovered (ng/100 ml)	Recovered (%)	Recovered (ng/100 ml)	Recovered (%)	Recovered (ng/100 ml)	Recovered (%)
53	40	75	53	100	30	57
106	103	97	93	88	66	62
212	233	110	152	72	113	53
Mean		94		87		57
s.e.		10		8		3
Significance of difference from 100% recovery		N.S.		N.S.		$p < 0.005$

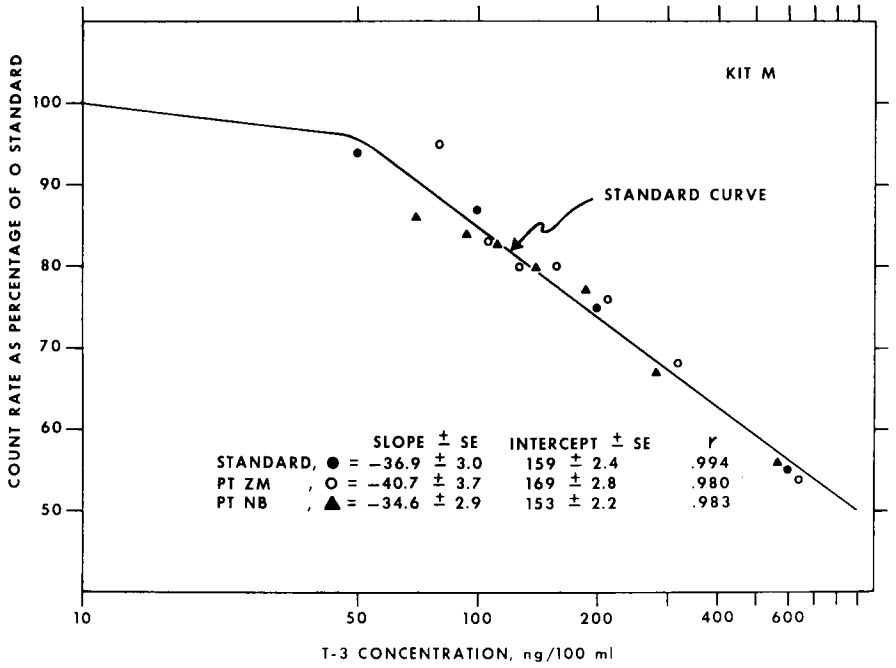

FIG. 2. Standard curve and values from dilutions of serums from two hyperthyroid patients by using kit M.

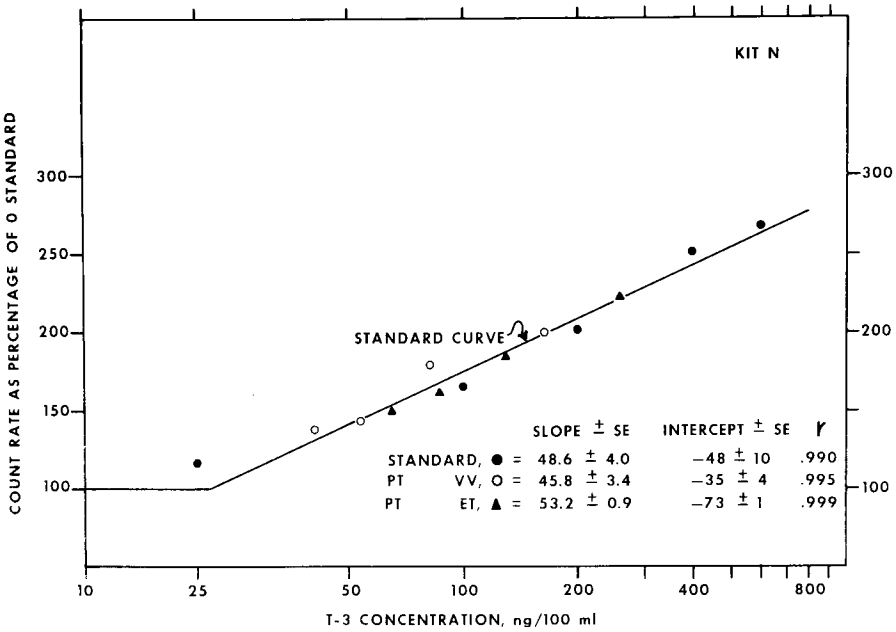

FIG. 3. Standard curve and values from dilutions of serums from two hyperthyroid patients by using kit N. Curve slopes upward since percentage free rather than percentage bound is determined with this kit.

terminations was lowest for kit M, next best for kit N, and highest for kit P. The between-days variance was very highly significant for kits P and M and of no significance for the kit N. These results also show that differences among serums are much greater than differences among duplicates of the same serum for each kit, a desirable feature for any assay. Although the between-days variance for kit P was significantly different, comparison of the days-to-serum variance ratio for this kit was not significant, enabling distinction of serums from day to day; kit M suffered in this respect. Between-duplicates variance data are summarized in Table 5, which shows the reproducibility of the three kits by expressing the standard deviation of the differences between duplicate determinations relative to the mean. The results show that, in terms of reproducibility between duplicate determinations, kit M has the smallest coefficient of variation, and kit P the largest.

Sensitivity, the smallest amount of unlabeled hormone which can be distinguished by the kit from no hormone, was estimated from the graphs of the standard curves (Figs. 1–3). The results show essentially no difference in sensitivity among the three radioassay kits (Table 6). Whereas kit N has the lowest value, it is not appreciably different from the other two values when expressed relative to the mean T_3 as determined by each kit. Furthermore, this may be a manifestation of the lower recoveries with kit N, accounting for a somewhat lower mean and range as well (Table 1).

Table 7 compares the technical and cost factors of

the three kits. This chart is based upon the performance of an equal number of serum samples in duplicate, and

TABLE 3. Precision of T_3 Radioassay Kits—Ten Replicate Determinations (Same Serum, Same Day)

	Kit P	Kit M	Kit N
Coefficient of variation	19%	13%	16%
Mean T_3 (ng/100 ml)	124	106	82

TABLE 4. Reproducibility of T_3 Radioassay Kits—Summary of Analysis of Variance of Results of Serum Triiodothyronine Concentrations (ng/100 ml)

	Variance		
	Kit P	Kit M	Kit N
Variation source			
Between days	7,927	26,148	233
Among serums	4,379	2,226	5,040
Interaction: days × serums	442	412	150
Between duplicates	566	125	213
Variance Ratio (F)			
Days/interaction	17.9*	63.4*	1.6†
Serums/duplicates	7.7*	17.8*	23.7*
Days/serums	1.8†	11.7‡	0.1†

*$p < 0.001$.
†*Not significant.*
‡$p < 0.01$.

TABLE 5. Reproducibility of T_3 Radioassay Kits—Variation Between Duplicate Determinations

Kit (No. of serums)	Mean T_3 (ng/100 ml)	Coefficient of variation (paired differences)
P(23)	116	21%
M(23)	112	10%
N(20)	104	14%

TABLE 6. Sensitivity of T_3 Radioassay Kits

Kit	T_3 (ng/100 ml)
P	42
M	39
N	32

TABLE 7. Comparison of T_3 Test Kits Performance and Cost Factors

	Kit P	Kit M	Kit N
Incubation time	4 hr, 37°C, 4°C	2 hr, 37°C, 4°C	2 hr, 20°C
Centrifugation	1	0	1
Pipetting steps	162, 1 volume 142 automated	102, 2 volumes 68 automated	120, 4 volumes 60 automated
Extraction step	None	None	Alcohol
Reaction vials	Test tubes, not supplied	Vials, supplied	Test tubes, not supplied
	Convenient	Awkward	Convenient
Separation B/F	Second antibody	Resin strip	Resin column
	Convenient	Awkward	Awkward
Price per 100 tests	$85	$165	$220

TABLE 8. Summary of Results

Parameter	Kit P	Kit M	Kit N
Accuracy	1	1	2
Specificity	1	1	1
Precision	3	1	2
Reproducibility	2	3	1
Sensitivity	1	1	1
Performance and cost factors	1	2	3
Total	9	9	10

the standard curve specified by the manufacturer for each of the kits. In terms of convenience, kit P rates highest since no extraction is necessary: almost all the steps can be automated, and manipulations are convenient. Kit N was judged to be the most awkward because of the need to employ an extraction step and a resin column.

Table 8 summarizes the results of the factors evaluated. One point was assigned for a first place, two for the second place, and three for a third place. As can be seen, there are no significant differences among the three kits when all factors are considered.

Discussion

The results indicate that, in terms of accuracy, kits M and P are somewhat superior to kit N, whose relatively low recovery is probably due to the inadequacy of the extraction of T_3 from serum by the alcohol method. The differences in extraction could account for the differences in the mean values observed in Table 1. However, all three tests show mean and ranges similar to those reported by others (*1, 4*).

The results of dilution of hyperthyroid serums show regression slopes not significantly different from the standard curves for all three kits, attesting to the accuracy and specificity of all.

With respect to precision, kit M was clearly superior, demonstrating the lowest variability between replications of the same serum within a run; kits N and P followed in that order. On the other hand, day-to-day variation was lowest for kit N and worst for kit M (Tables 3 and 4). As stated above, the differences among serums for each kit were much greater than differences among duplicates of the same serum, a desirable feature since one may distinguish among high, low, and normal serums in any run.

References

1. Utiger RD: Serum triiodothyronine in man. *Ann Rev Med* 25:289–302, 1974

2. Midgley AR Jr, Niswender GD, Rebar RW: Principles for the assessment of the reliability of radioimmunoassay methods (precision, accuracy, sensitivity, specificity). *Acta Endocrinol* 63(Suppl 142):163–184, 1969

3. Batson HC: *An Introduction to Statistics in the Medical Sciences*. Minneapolis, Burgess Publishing Co., 1963

4. Larsen PR: Triiodothyronine: Review of recent studies of its physiology and pathophysiology in man. *Metabolism* 21:1073–1092, 1972