INTER-LABORATORY ASSAY PERFORMANCE OF THREE THYROID RELATED HORMONE RADIOASSAYS*

I. SCHWARZ, M.T.C.P. RIBELA, R.R. ARKATEN,
J.E. OLIVEIRA, P. BARTOLINI
Division of Medicine,
Department of Application of Nuclear Techniques
in Biological Sciences,
National Nuclear Energy Commission (IPEN–CNEN),
São Paulo, Brazil

Abstract

INTER-LABORATORY ASSAY PERFORMANCE OF THREE THYROID RELATED HORMONE RADIOASSAYS.

The paper presents the results of a survey comparing the assay performances of seven Brazilian laboratories in the radioimmunoassay (RIA) of triiodothyronine (T3) and thyroxine (T4) and in the immunoradiometric assay (IRMA) of thyrotropin (TSH). All laboratories worked with matched bulk reagents from the North East Thames Region Immunoassay Unit, London, United Kingdom (NETRIA) and during the two year period of this study, fifty-eight T3, fifty-two T4 and fifty TSH assays were evaluated. For the inter-laboratory comparison, the results of three quality control sera (QCS) determinations in each of the assay batches were analysed. The analysis was based on the deviation of individual results from the trimmed mean calculated for the pooled results from all laboratories. Three lots of reagents of each analyte type were involved in the study, and two of them (the first lot of T3 and of T4) gave inconsistent inter-laboratory results, presenting high between-batch coefficients of variation (BBCV), of about 30–40%. The TSH-IRMA results in the medium and high dose ranges (but not in the low dose range) presented a considerably lower BBCV (about 8%) than all the T3 and T4 results (12–20%). The quality of the assays seemed to improve with time, and the BBCV of about 20% in the beginning fell to about 12% by the end of the two year study period. No statistical evidence for differences in performance among the participating laboratories could be detected.

1. INTRODUCTION

Radioassays in general may present problems of precision and reproducibility and their results can be influenced by factors such as staff training, work overload, reliability of assay techniques, standardization and reagent stability, especially the radioactive tracers [1].

* Work supported by IAEA Research Contract No. 4872.

The importance of quality assurance of general clinical chemistry laboratory analyses through external surveillance schemes is widely recognized [2], and is being more and more applied to radioassays, after their growing use in clinical laboratories since about 20 years [3]. In most developed countries laboratories are enrolled in external quality assessment programmes, but in Brazil, despite the relatively wide use of radioassays, this practice is very limited: clinical laboratories interested in a quality control surveillance have recourse to international programmes or apply only internal quality control practices. Until now, only one local approach towards external quality control was undertaken in Brazil, by the group of A.M. Fiori [4], which includes about ten laboratories from a small southern region of Brazil and two neighbouring countries.

In 1988 the IAEA began to support a Latin American Co-ordinated Research Programme, the ARCAL VIII Project, aiming towards the regional production of the reagents for the radioassays of the thyroid related hormones: triiodothyronine (T3), thyroxine (T4) and thyrotropin (TSH). Besides this, one of the main objectives of this project is the stimulation of quality control practices in the region, both at intra- and inter-laboratory levels.

The purpose of this paper is to provide some information on the radioassays performed on the above mentioned hormones, in seven Brazilian laboratories participating in the ARCAL VIII Project.

2. METHODS

2.1. Co-operating laboratories

Apart from our laboratory — the Brazilian Reference Laboratory — the six other participants are clinical laboratories of public or university hospitals. All have radioassay experience, but mainly employ commercial kits. These laboratories are located in the north-east, centre-south and south of Brazil, covering a large area of the country. To each laboratory a number was assigned, by which each will be cited in the text that follows. Laboratory 5 participated only in the first year of the project, and laboratory 7 joined it in the second half of the second year.

2.2. Assay reagents

The distribution of matched reagents for the development of the project was chosen according to the good results reported by a similar scheme in the United Kingdom [5]. All participants received the necessary assay reagents in bulk from the North East Thames Region Immunoassay Unit, London, United Kingdom (NETRIA), so all laboratories worked with the same set of reagents, excluding inter-laboratory variabilities associated with the quality of the reagents. Total serum T4

and T3 radioimmunoassays RIAs were incubated for 2 h and overnight, respectively, while the bound-free separation was performed by a polyethylene glycol (PEG) assisted double antiserum system. Serum THS was determined by a solid phase immunoradiometric assay (IRMA) method, using ^{125}I monoclonal anti-TSH as tracer and a polyclonal anti-TSH antiserum coupled to cellulose as solid phase; incubation time was overnight.

Together with the necessary assay reagents, three internal quality control sera (QCS), having low, medium and high concentrations of each hormone, were also supplied by NETRIA.

All reagents, except TSH IRMA solid phase, were supplied lyophilized, to be reconstituted by the co-operating laboratories, which also had to prepare incubation and wash buffers. Only three participants (laboratories 1, 2 and 4) had previous experience in the work with bulk reagents, but special instructions about buffer preparation, reagent aliquoting and storage, tracer dilution, as well as about calculations and internal quality control practices, were necessary at the beginning.

During the two year period of this study, three different lots of reagents were distributed, while the radioiodinated tracers were furnished every two months. In some assays performed with Lots II and III, tracers (^{125}I-T3, ^{125}I-T4 and ^{125}I monoclonal anti-TSH) prepared in our laboratory were also tested on a small scale, in comparison with the NETRIA tracers, and their results were included because no significant differences were observed.

2.3. Data analysis and calculations

The data presented in this paper refer to the determinations performed on the internal QCS from NETRIA. During the two year period of this study a total of fifty-eight T3, fifty-two T4 and fifty TSH assays were evaluated. For an inter-laboratory comparison we had to consider the following limitations: (a) the small number of participants; (b) the variable number of responses sent by each laboratory (1 to 25); (c) the lack of a known 'target' value. Besides this, three lots of QCS had to be analysed, each one with differing hormone levels. Because of these aspects, this survey focused more on the between-laboratory comparison of performances and the study of the variability of this performance with time, employing simple statistical calculations.

Assays discarded by reason of very poor precision or other gross errors were excluded from the analysis.

The first approach was a plot of all valid results in charts, as shown in Fig. 1. The visual inspection of result dispersion helped us in deciding whether the calculation of a mean target value for a particular set (A, B, C) and lot (I, II or III) of QCS seemed justified or not, for an inter-laboratory statistical treatment of assay performance, trying not to introduce a bias due to a particular laboratory, which possibly had more determinations than others. When the dispersion of the results of most

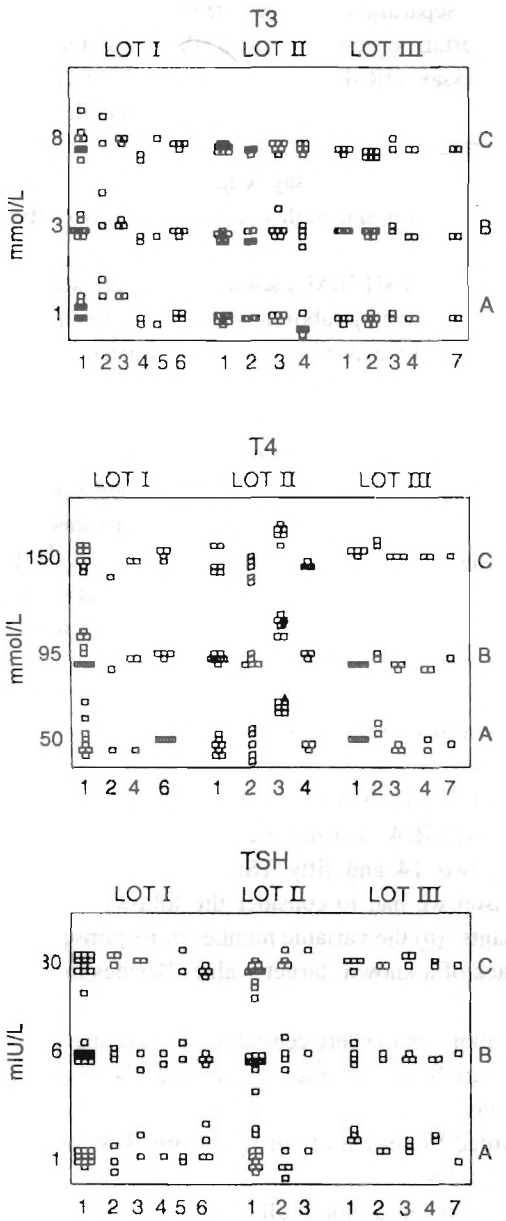

FIG. 1. Quality control charts of the results determined by seven laborqtories (1–7) for three quality control sera (A, B and C) in the T3, T4 and TSH radioassays.

TABLE I. MEAN AND BBCV OBTAINED IN THREE QUALITY CONTROL SERA (A, B AND C); POOLED RESULTS FROM n ASSAYS, PERFORMED IN SEVEN BRAZILIAN LABORATORIES

Assay	QC	Lot No.	n	Mean	BBCV
		I	18[a]	1.50 nmol/L	41
	A	II	20	1.19 nmol/L	15
		III	12	0.93 nmol/L	17
		I	17[a]	3.78 nmol/L	27
T3	B	II	20	2.79 nmol/L	24
		III	12	3.38 nmol/L	12
		I	15[a]	8.9 nmol/L	25
	C	II	21	6.1 nmol/L	17
		III	12	6.8 nmol/L	11
		I	16[a]	58 nmol/L	29
	A	II	13	399 nmol/L	32
		III	9	57 nmol/L	7
		I	16[a]	123 nmol/L	31
T4	B	II	13	98 nmol/L	15
		III	9	94 nmol/L	12
		I	14[a]	180 nmol/L	25
	C	II	13	151 nmol/L	19
		III	9	162 nmol/L	11
		I	19	0.82 IU/L	31
	A	II	11	0.64 IU/L	49
		III	10	1.28 IU/L	19
		I	20	6.7 IU/L	8
TSH	B	II	11	6.4 IU/L	8
		III	10	6.7 IU/L	7
		I	20	28.9 IU/L	12
	C	II	11	27.6 IU/L	7
		III	10	29.0 IU/L	4

[a] Excluded from the inter-laboratory comparison.

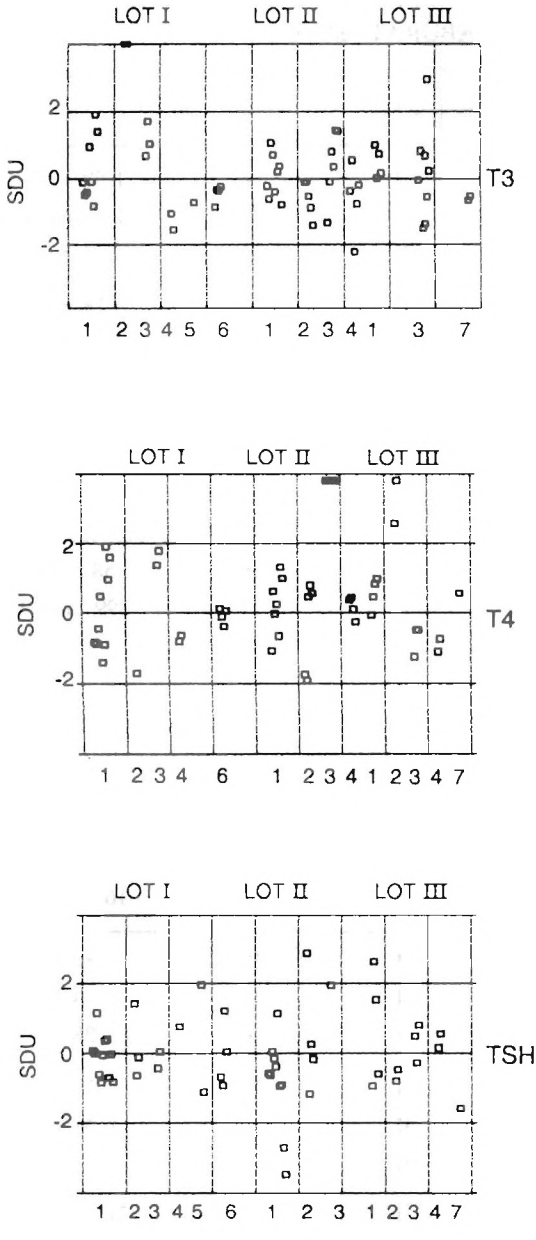

FIG. 2. Mean SDU variability of individual assay batches from seven laboratories, for T3, T4 and TSH radioassays.

laboratories proved to be around a consensus value, we proceeded with the statistical data treatment described below.

2.3.1. *Calculations*

Owing to the limitations outlined above, simple calculations were performed — the arithmetic mean, standard deviation (SD) and the between-batch coefficient of variation (BBCV) were calculated for each series of results. Deviation of each value from the calculated mean was estimated in terms of standard deviation units (SDU), as follows:

$$SDU = (x_i - x)/SD$$

where x_i = i^{th} value from the series; x = arithmetic mean of the series; SD = standard deviation.

2.3.2. *Trimming*

In a first approach, statistics were calculated for each laboratory having done a minimum of three assays for a particular lor of QCS. For a specific QCS and lot, the mean and the SDUs were calculated. Arranging the absolute SDU values of the three QCS in increasing order, 10% at the high extreme were discarded. By this procedure, possible outliers from each laboratory were excluded, even if they were not recognized as outliers in the final calculations. After this trimming procedure, the results from all laboratories were pooled and the overall mean and BBCV were calculated per QCS and lot. These results are shown in Table I.

Finally, for the purposer of inter-laboratory assay performance comparison, for each assay batch a mean SDU was calculated, considering the three SDU values corresponding to each QCS (A+B+C) of an assay. This gave us good information about the overall performance in each assay batch; the results are shown in Fig. 2.

2.3.3. *Criteria for discarding an assay*

In discarding an assay, we followed the general recommendations made as to when the results on the QCS should be considered as outliers, and repeating of analysis would be necessary [6], i.e. when:

(a) One QCS deviates from the mean by more than 3 SDU.
(b) Two QCS deviate from the mean (in the same direction) by more than 2 SDU.
(c) Three QCS deviate from the mean (in the same direction) by more than 1 SDU.

Results for the three analytes are summarized in Table II and the mean SDU values are shown batch by batch in Fig. 2.

TABLE II. OUTLIERS IN T3[a], T4[a] AND TSH ASSAYS PERFORMED IN THE SEVEN LABORATORIES. n = TOTAL NUMBER OF ASSAYS; d = DISCARDED ASSAYS

Laboratory	1	2	3[b]	4	5	6	7
n	44	21	21	16	2	4	4
d	6	7	5	3	1	1	0
n/d (%)	13.6	33.3	23.8	18.8	50.0	25.0	0

[a] Lot I assays not included.
[b] T4 Lot II not included.

3. RESULTS AND COMMENTS

Many assays obtained with the first lot of T3 and T4 RIAs had to be discarded because of unacceptable precision or invalid standard curves. Although these lots were not considered in the inter-laboratory evaluation, they may be responsible for the bad results relative to Lot I of the QCS shown in Fig. 1. A high between-laboratory dispersion is observed, each one of the six laboratories tending to have its own mean value; this is the reason why these assays were excluded from the inter-laboratory comparison. In Table I we can see that the BBCVs of these assays were very high. These poor quality results may be associated with less stable tracers, often used more than 30–40 days after labelling. From Lot II on, and for all TSH IRMA lots, the dispersion of the results was considerably lower, and mean values were calculated, although a relatively high dispersion was still apparent in Lot II assays. Particularly the T4 Lot II assays (n = 6) from laboratory 3 showed an extreme deviation, associated with the tracer preparation, and for which no other explanation could be found: other laboratories using the same tracer did not have problems.

The analysis of the results from Table I shows a quite acceptable inter-laboratory precision, with BBCV greater than 20% for only 4 of the 21 valid lots of QCS. Two of these poorly performing lots were low dose TSH, which seems to present in general a higher variability. On the other hand, in the medium and high dose ranges, the TSH IRMA results presented a considerably lower BBCV — about 8% — than the T3 and T4 results. The general performance showed a tendency to improve with time.

Figure 2 displays the overall assay performance of each laboratory (mean SDU/assay batch). Excluding the aberrant values from laboratory 3 on T4 Lot II, the percentage of discarded assays varied from 0 to 50%, as shown in Table II. But,

considering the small number of assays of some laboratories, there is no statistical evidence that one particular laboratory has a different assay performance or bias, compared to the others.

4. CONCLUSIONS

Within the limitations of the available data, we can conclude that all participating laboratories showed a similar performance and that the quality of assays seemed to improve with the progress of the project. The co-operating laboratories ended up without major problems in handling bulk reagents. This is also evidence of the robustness of NETRIA's assay systems.

REFERENCES

[1] SETH, J., HANNING, I., Factors associated with the quality of laboratory performance in the United Kingdom external quality assessment scheme for serum growth hormone, Clin. Chim. Acta **174** (1988) 185–196.

[2] ROSS, J.W., Evolution of evaluation criteria in the College of American Pathologists surveys, Arch. Pathol. Lab. Med. **112** (1988) 334–339.

[3] SETH, J., The external quality assessment of hormone assays, J. Clin. Biochem. Nutr. **2** (1987) 111–139.

[4] FIORI, A.M., FIORI, A.M.C., GARCIA, E.J., "Regional quality-control programme for the measurement of thyroid hormones by radioimmunoassay", Radioimmunoassay and Related Procedures in Medicine 1982 (Proc. Symp. Vienna, 1982), IAEA, Vienna (1982) 601–606.

[5] EDWARDS, R., PULL, P.I., "Development and operation of a regional immunoassay scheme", Immunoassays for Clinical Chemistry (HUNTER, W.M., CORRIE, J.E.T., Eds), Churchill Livingstone, Edinburgh (1983) 16–24.

[6] INTERNATIONAL ATOMIC ENERGY AGENCY, Radioimmunoassay Data Processing Program for IBM PC Computers, IAEA-TECDOC-509, Vienna (1988).