Clinical Chemistry 45:12
2207–2215 (1999)

Endocrinology and Metabolism

Comparison of Pituitary and Recombinant Human Thyroid-stimulating Hormone (rhTSH) in a Multicenter Collaborative Study: Establishment of the First World Health Organization Reference Reagent for rhTSH

Brian Rafferty[1*] and Rose Gaines Das[2]

Background: **The increasing use of recombinant-DNA-derived materials in therapy and diagnosis poses a new challenge for biological standardization, that of developing reference preparations appropriate for both the native and recombinant products. Here we report the results of an international collaborative study that was carried out under the auspices of WHO to assess the suitability of a preparation of recombinant thyroid-stimulating hormone (rTSH; 94/674) to serve as a potential standard for the calibration of diagnostic immunoassays compared with the International Reference Preparation (IRP) for human TSH (80/558).**

Methods: **Coded samples were provided to the 33 laboratories in the study, and participants were asked to perform TSH assays currently in use in their laboratories. Twenty-eight laboratories contributed 93 immunoassays in 41 different method-laboratory combinations, and an additional 5 laboratories contributed bioassay data. All data were analyzed centrally at the National Institute for Biological Standards and Control.**

Results: **The results obtained in different laboratories and with different assay systems revealed significant variability between estimates of rTSH relative to the IRP. These ranged from 5.51 mIU (95% limits, 3.95–7.67 mIU) per ampoule by RIA to 7.15 mIU (95% limits, 6.7–7.63 mIU) per ampoule by immunofluorometric assay. However, the results showed that the assignment of a value of 6.70 mIU per ampoule of 94/674 would give reasonable continuity with the IRP in many assay systems.**

Conclusions: **The preparation was established as the First WHO Reference Reagent for TSH, human, recombinant, to provide a means of validating assay performance and to maintain continuity with the IRP without compromising clinical data.**

Immunoassays for thyroid-stimulating hormone (TSH)[3] are widely applied in the diagnosis and management of thyroid dysfunction. The calibrators in these assays are secondary standards whose activities have been assigned relative to a primary standard, typically the Second International Reference Preparation (IRP) for human TSH *(1)*. Because of the safety issues involved in the use of TSH of pituitary origin (from which the primary and most secondary standards are derived), increasing use is being made of preparations of recombinant human TSH (rhTSH) as calibrators in hTSH immunoassay systems. The expression and large-scale production of rhTSH in Chinese hamster ovary (CHO) cells have been reported *(2, 3)*, and clinical trials for its proposed diagnostic use to stimulate radioiodine uptake into residual or metastatic thyroid tissue in patients with thyroid carcinoma are underway *(4)*. Several studies have been carried out to assess the biological and immunological characteristics of the recombinant molecule *(5, 6)* and its use for the preparation of immunoassay reagents *(7)*. With the ready availability of rhTSH, it is anticipated that increasing use

Divisions of [1] Endocrinology and [2] Informatics, National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Herts EN6 3QG, UK.

*Author for correspondence. Fax 44-1707-646730; e-mail brafferty@nibsc.ac.uk.

Received in revised form July 19, 1999; accepted September 17, 1999.

[3] Nonstandard abbreviations: TSH, thyroid-stimulating hormone; IRP, International Reference Preparation; rhTSH, recombinant human TSH; CHO, Chinese hamster ovary; NIBSC, National Institute for Biological Standards and Control; IFMA, immunofluorometric assay; ICLMA, immunochemiluminometric assay; PAb, polyclonal antibody; and MAb, monoclonal antibody.

will be made of this material as a calibrator in hTSH immunoassay systems.

To address this point, a candidate preparation of rhTSH was obtained and a collaborative study organized to establish whether it could be calibrated adequately by immunoassay in terms of the existing Second IRP of pituitary TSH *(8)*. The aims of the study were: (*a*) to compare rhTSH and pituitary TSH in a variety of immunoassay systems; (*b*) to assess the suitability of the candidate preparation of rhTSH to serve as a standard for the calibration of diagnostic immunoassays; (*c*) to assess the stability of the candidate preparation of rhTSH after accelerated thermal degradation; and (*d*) to confirm the bioactivity of rhTSH after ampouling and lyophilization procedures.

Materials and Methods

PREPARATIONS OF TSH

The ampouled materials supplied to participants for the study are listed in Table 1. In addition, participants were asked to include their in-house standards in the assays. For immunoassay, each participant received a basic set of ampoules containing the Second IRP, rhTSH (94/674), and pituitary TSH preparation 81/502, and in addition they received accelerated degradation samples selected on the basis of assay capacity, sample availability, and a design giving information across as wide a range of assay systems as feasible. For bioassay, the Second IRP was replaced with the pituitary-derived bioassay standard, 84/703 [National Institute for Biological Standards and Control (NIBSC) Std B]. The Second IRP and pituitary TSH preparations 81/502 and 84/703 have been described *(1)*. The rhTSH 94/674 was ampouled at NIBSC, following procedures described by WHO *(9)*. Highly purified (>95% by sodium dodecyl sulfate-polyacrylamide gel electrophoresis) rhTSH, expressed in CHO cells *(2)*, was kindly donated to WHO by the Genzyme Corporation, Framingham, MA. The preparation was received as a lyophilized white powder (nominal 2 mg) that, after dilution in 2 L of a solution containing 1 g/L human serum albumin and 5 g/L lactose, was distributed into ampoules as 1.0-mL aliquots. Each ampoule (coded 94/674) has a nominal content of 1 μg of TSH, 1 mg of human plasma albumin, and 5 mg of lactose.

PARTICIPANTS IN THE STUDY

Thirty-three laboratories in 11 countries took part in the study and are listed alphabetically, by country, below. Throughout this report, each participating laboratory is referred to by a code number, which was assigned randomly and which does not reflect the order of listing:

Drs. T. Byrne and C. Morris, Therapeutic Goods Administration Laboratories, P.O. Box 100, Woden, ACT 2606, Australia.
Dr. E. Poncelet, Medgenix Diagnostics, Zoning Industrial B-6220 Fleurus, Belgium.
Ms. C. Peroni, Dr. M.T. Ribela, and Dr. P. Bartolini, Instituto de Pesquisas, Energeticas e Nucleares, University of Sao Paulo, Caixa Postal 11049 (Pinheiros), Sao Paulo, Brazil.
Dr. J.G.H. Vieira, Disciplina de Endocrinologia/Escola Paulista de Medicina, Caixa Postal 20266, 04034-970 Sao Paulo, Brazil.
Dr. Leng Wei, NICPBP, Temple of Heaven, Beijing 100050, Peoples Republic of China.
Dr. K. Lahtinen, Wallac Oy, P.O. Box 10, SF-20101 Turku, Finland.
Dr. B. Incaurgarat, Biomerieux SA, 69280 Marcy L'Etoile, France.
Prof. H. Rochat, Lab Biochimie, Hopital Nord, 13915 Marseilles, France.
Prof. C. Ronin, URA 1296 CNRS-IBSM, GLM-CNRS, 13402 Marseilles, France.
Drs. C. Ullrich and A. Bergmann, BRAHMS Diagnostica GmbH, Postfach 42 04 41, 12064 Berlin, Germany.
Dr. M. Fiorani, Biodata Spa, 00012 Guidonia Montecelio, Rome, Italy.

Table 1. Preparations supplied to participants in collaborative study.

TSH preparation	Defined activity	Nominal ampoule contents
2nd IRP for TSH for Immunoassay (80/558)	37 mIU/ampoule	7.5 μg of TSH extract, 5 mg of lactose, 1 mg of human albumin
rhTSH (94/674)		1 μg of rhTSH, 5 mg of lactose, 1 mg of human albumin
Accelerated degradation samples of 94/674 stored		Contents assumed identical to 94/674
215 days at 20 °C		
215 days at 37 °C		
215 days at 45 °C		
Pituitary TSH (81/502)	7.75 mIU/ampoule	2 μg of TSH extract, 5 mg of lactose, 1 mg of human albumin
Accelerated degradation samples of 81/502 stored		Contents assumed identical to 81/502
15 years at 20 °C		
15 years at 37 °C		
Pituitary TSH (84/703; NIBSC Res Std B for bioassay)	37 mIU/ampoule	7.5 μg of TSH extract, 5 mg of lactose, 1 mg of human albumin

Dr. P.B. Romelli, Technogenetics SRL, 20099 Sesto San Giovanni, Italy.
Mr. D. Allan and Dr. S. Edwards, Johnson & Johnson Clinical Diagnostics, Chalfont St Giles, Bucks HP8 4SP, UK.
Dr. S.P. Bidey, Endocrine Sciences Research Group, University of Manchester, Oxford Road, Manchester M13 9PT, UK.
Ms. L. Flack and Mr. J.B.P. Titman, Anagen (UK) Ltd, Aldershot, Hants GU12 4UH, UK.
Drs. J.A. Little and R. Edwards, NETRIA, St. Bartholomew Close, London EC1A 7BE, UK.
Prof. D.S. Munro, 26, Endcliffe Grove Ave., Sheffield S10 3EJ, UK.
Mr. P. Puntin, IDS Ltd, Boldon Business Park, Boldon, Tyne & Wear NE35 9PD, UK.
Mr. B. Rafferty, Division of Endocrinology, National Institute for Biological Standards & Control, Potters Bar, Herts EN6 3QG, UK.
Dr. B. Rees Smith, RSR Limited, Pentwyn, Cardiff CF2 7HE, UK.
Drs. M. Watts and M. Walker, SCL Bioscience Services Ltd, Cambridge CB4 4ZA, UK.
Dr. L.B. Garcia Aguirre, National Direction of Nuclear Technology, Mercedes 1041, Montevideo, Uruguay.
Ms. D. Bethell and Ms. S. Sasaki, Medix Biotech, Inc, San Carlos, CA 94070.
Dr. P. Bodlaender, Diagnostic Products Corporation, Los Angeles, CA 90045-5597.
Mr. D. Carter and Mr. J. Kim, Corning Nichols Institute, San Juan Capistrano, CA 92690.
Mr. D. Clark, Dade International Inc., P O Box 520672, Miami, FL 33152.
Dr. Z. Latif, Abbott Laboratories, Abbott Park, IL 60064.
Mr. W. Link, Diagnostics Group, Bio-Rad Laboratories, Hercules, CA 94547.
Mr. D. Lino, CIBA-Corning Diagnostics, E. Walpole, MA 02062.
Drs. P. Liu and P.D.K. Lee, Diagnostic Systems Laboratories Inc, Webster, TX 77598.
Dr. R. Mattaliano, Genzyme Biotherapeutics, Framingham, MA 01701-9322.
Dr. W.D. Odell, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, Utah 84132.
Ms. S. Wilson and Dr. R. Sarber, Genzyme Corporation, Cambridge, MA 02139.

DESIGN OF THE STUDY AND ASSAY METHODS CONTRIBUTED

Participants were asked to perform the TSH immunoassay(s) currently in use in their laboratory and to carry out at least two independent assays (i.e., using freshly reconstituted ampoules) of each assay type, each assay to include a set of ampouled preparations selected with regard to assay system capacity and to provide maximum information for the study. Each set of preparations included a coded duplicate preparation to provide an independent assessment of assay accuracy and precision. All preparations and any local standards were to be included at several doses to provide information on linearity and parallelism of the dose–response relationship. Selected participants were invited to provide bioassay data, following as far as practical, the broad design suggested for immunoassays. All raw data were reported to NIBSC for analysis.

STATISTICAL ANALYSIS

As far as possible for this study, an "assay" was defined as an independent test for each preparation beginning with freshly reconstituted ampoules.

In all immunoassays, the dose–response relationships could be reasonably described using a four-parameter logistic function. The fitted asymptotic values of the logistic function were used to transform the reported responses to logits. Logit responses were analyzed using weighted linear regression and an in-house program *(10)*, and the linearity and parallelism of the log dose-logit response lines were assessed. In assays showing apparently significant ($P < 0.05$) deviations from parallelism, the differences in slopes between different preparations were compared to the differences in slopes between the coded duplicate preparations; differences between different preparations that were less than the observed differences between the coded duplicate preparations were not considered to reflect consistent nonparallelism. Relative potencies were determined as the displacement of fitted parallel log dose-logit response lines. For some bioassays, there were insufficient data for satisfactory fitting of a four-parameter logistic relationship; for these assays, log dose-response data were analyzed using the methods for parallel line assays. Except as otherwise noted, there was no exclusion of data.

Groups of potency estimates were compared using unweighted analysis of variance of log potency estimates. Potency estimates were combined as unweighted geometric means, and 95% fiducial limits for the combined estimate were calculated using the variance of the logs of the estimates combined. In some cases, the variation between estimates was expressed as the geometric CV {gCV = [exp (standard deviation of log potency estimates) − 1] multiplied by 100}.

Estimates of the relative activity (to ampoules of the same preparation stored continuously at −20 °C) remaining in the ampoules of the TSH preparations after storage at increased temperatures were used to fit an Arrhenius equation relating degradation rate to absolute temperature assuming first-order decay *(11)* and, hence were used to predict the degradation rate of the preparations when stored at −20 °C.

Results

ASSAYS CONTRIBUTED AND PRELIMINARY ANALYSIS

A total of 93 immunoassays by 41 different method-laboratory combinations, and 10 bioassays by 5 methods were contributed to the study. These are summarized in Tables 2 and 3, respectively. Immunoassay methods are categorized as either competitive (RIA) or noncompetitive, with the latter being further subdivided into radioisotopically (IRMA) or nonradioisotopically [ELISA, immunofluorometric assay (IFMA), and immunochemiluminometric assay (ICLMA)] labeled systems. The antibodies used in the RIAs were polyclonal (PAb) in origin. Where known, the antibody formats used in the noncompetitive systems were solid-phase monoclonal antibody/detector monoclonal antibody (MAb/MAb; 50%), MAb/PAb (36%), or PAb/MAb (14%), none of which was specific for a particular system. No classification was made as to whether the systems were manual or automated. As far as could be determined, methods are unique to the laboratories except that two of the IFMAs were each reported to be used in two laboratories (one method in laboratories 10 and 13, and one in laboratories 12 and 19) and the same IRMA was used in laboratories 9 and 16. No in vivo bioassays were available. Because of the limited data available from the in vitro bioassays, no formal comparisons among them or of bioassay and immunoassay results were attempted. Estimates from some laboratories were excluded from some of the subsequent calculations because of noninclusion of the IRP in their assays (laboratory 2) or inappropriate reconstitution of ampoule contents in water (laboratories 4 and 12). Log dose-logit response lines were generally linear. With only a few exceptions, the individually fitted log dose-logit response lines did not show any consistent significant deviations from parallelism. A full listing of estimates of activity for all preparations (shown in Figs. 1, 2, and 3) may be obtained from the authors upon request.

Table 2. Immunoassay methods used in the study.

Assay type and laboratory code	Assay sources[a]
RIA	
07, 17, 28	University of Sao Paulo (in-house); Tianjin Corporation; University of Utah (in-house)
IRMA	
02, 03, 04, 05, 07, 08, 09, 15, 16, 17, 18, 21, 22, 26, 28	Bio-Rad RADIAS; bioMerieux CoatRia; LifeScreen coated tube; Nichols Allegro; DPC Coat-a-Count (2 labs); University of Utah (in-house); University of Sao Paulo (in-house); TSH MYRIA; Beijing Northern Institute; DSL ACTIVE coated tube; IDS TSH Assay; Medgenix TSH-IRMA; NETRIA (UK); Amerlite
ELISA	
06, 08, 15, 17, 18, 20, 24, 25	Serono SR1 IEMA; BioClone; Bio-Rad RADIAS; Medgenix TSH-EASIA; Tianjin Bioresearch Corporation; NETRIA (UK); Medix TSH-ELISA; NIBSC (in-house)
IFMA	
01, 10 (× 2), 11, 12, 13 (× 2), 19 (× 2), 26	Wallac AutoDELFIA (2 labs); Abbott IMx (2 labs); Abbott AxSYM; Anagen AuraFlex; bioMerieux Vidas; Baxter Stratus
ICLMA	
04, 14, 16, 21, 23, 27	Ciba-Corning ACS:180; DPC Immulite (2 labs); Nichols Chemiluminescent; BRAHMS LUMItest; Amerlite

[a] Not in the same order as laboratory codes.

Table 3. Bioassay methods used in the study.

Laboratory code	Bioassay method	mIU of 84/703 per ampoule of 94/764
29	[^{3}H]Thymidine incorporation into FRTL-5 thyroid cells	3.12 (2 assays)
30	[^{3}H]Adenine incorporation into JP09 cell line transfected with human TSH receptor	1.85 (1 assay)
31[a]	Binding to porcine receptor preparation	6.4
32	cAMP generation after binding of TSH to thyroid microsomal preparation	5.82 (2 assays)
33	Iodide uptake into FRTL-5 cells	2.95[b]

[a] Raw data unavailable.
[b] mIU of 81/502.

VARIABILITY OF IMMUNOASSAY ESTIMATES

Estimates of the activity of the coded duplicates of 94/674 relative to one another provided an independent measure of assay variability. Individual estimates were between 0.64 and 1.34, and the geometric mean (95% limits) of all individual estimates was 0.98 (0.96–1.01), in good agreement with the value of 1 expected for identical preparations.

Deviations of the observed individual values from the expected relative activity of 1.0 gave a pooled (over all immunoassay methods and laboratories) within-assay mean square for log potency estimates of ~0.01, corresponding to an ~95% interval of 82–122% for comparison of two identical materials within an assay in these systems. The between-assay mean square (also pooled over all methods and laboratories) was smaller than this, suggesting that there may be a tendency for greater similarity of estimates for the same preparations between assays than for independent tests of the same preparations within the same assay. It was noted that the within- and between-assay variability, pooled over the broad groupings of methods, was 2- to 10-fold larger for ELISAs than for the other methods.

COMPARISON OF rTSH 94/674 AND THE PITUITARY hTSH IRP

The laboratory geometric mean estimates for each of the duplicate ampoules of 94/674, expressed as mIU of the IRP per ampoule, are shown in Fig. 1. The between-laboratory mean square for these estimates was approximately three- to fourfold larger than that obtained for the

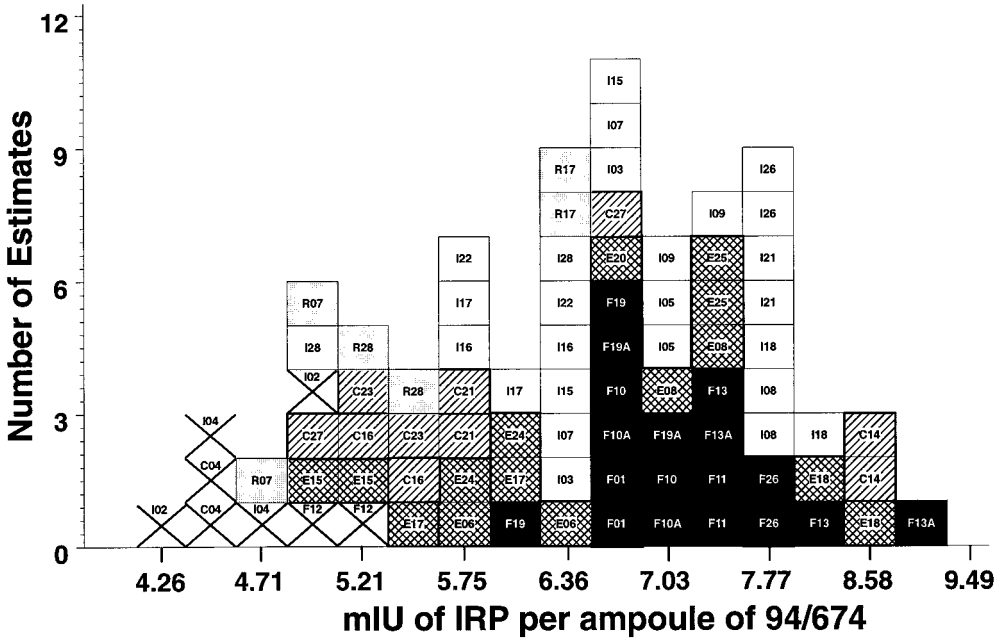

Fig. 1. Laboratory geometric mean estimates for duplicate ampoules of 94/674 expressed as mIU of the IRP per ampoule.

Each *square* denotes a single combined estimate for each assay method in each laboratory. For 94/674, a single estimate was obtained for each of the coded duplicate ampoules (i.e., two estimates for each laboratory and method in most cases). The *label* in each *square* denotes the assay method (*first character*) and the laboratory code (*second two characters*). Methods codes: *C*, ICLMA; *F*, IFMA; *E*, ELISA; *I*, IRMA, *R*, RIA; methods are also identified by *shading*. Estimates shown as X were obtained in assays that did not include a fresh ampoule of the IRP, or in laboratories using water as the primary diluent.

direct comparison of the identical preparations, reflecting the markedly different estimates obtained in the different laboratories. Because of this between-laboratory variability and because not all laboratories contributed the same number of assays, a single geometric mean of all assay estimates for the two ampoules of 94/674 was obtained for each method in each laboratory, and these laboratory and method means were used to calculate the combined estimates. Estimates by RIA (mean, 5.51 mIU per ampoule; 95% limits, 3.95–7.67 mIU per ampoule; n = 3) were notably smaller than estimates by the other broadly defined methods, and this difference would be more marked if the estimates from laboratory 17 were excluded. Estimates by ICLMA were also generally smaller than estimates by other methods if the ICLMA of laboratory 14 was excluded (mean, 5.55 mIU per ampoule; 95% limits, 5.20–5.92 mIU per ampoule; n = 4). Estimates by IFMA (mean, 7.15 mIU per ampoule; 95% limits, 6.7–7.63 mIU per ampoule; n = 9) were generally larger than estimates by the other methods. Estimates by ELISA (mean, 6.50 mIU per ampoule; 95% limits, 5.69–7.43 mIU per ampoule; n = 8) and IRMA (mean, 6.77 mIU per ampoule; 95% limits, 6.31–7.26 mIU per ampoule; n = 13) spanned the ranges of estimates by the various other methods, with marked differences between estimates from different laboratories within these methods. The mean estimate over all methods, excluding data as described previously, was 6.59 mIU per ampoule (95% limits, 6.28–6.92 mIU per ampoule; n = 38).

A limited number of bioassays were carried out as summarized in Table 3. Laboratory 31 also reported an estimate of 2.4 mIU for an ampoule of 94/674 stored at 45 °C for 400 days.

COMPARISON OF THE VARIOUS LOCAL STANDARDS WITH THE IRP

Geometric mean estimates of the relative activity of the various local standards expressed as IU of the IRP equivalent in activity to a nominal IU of the local standard are shown in Table 4. These estimates were all ~1.0, as might be expected given the length of time for which the IRP has been in use, although there were significant ($P < 0.05$) differences between laboratories when the local standards were compared directly with the IRP. Individual estimates ranged from 0.58 to 1.31 IU per nominal IU, with a mean of individual estimates of 0.97 IU (95% limits, 0.89–1.05 IU) per nominal IU.

COMPARISON OF PITUITARY hTSH 81/502 AND THE IRP

Laboratory geometric mean estimates of 81/502 expressed as mIU of the IRP per ampoule are shown in Fig. 2. The between-laboratory mean square, pooled over all laboratories and methods, for these estimates was approximately twofold greater than that for the coded duplicate preparations, suggesting that there were significant differences between laboratories. For example, the RIAs in use in laboratories 07 and 28 gave estimates that were notably smaller than the estimate given by the RIA in use in laboratory 17, a relationship similar to that seen for comparison of 94/674 with the IRP. There was no clear difference between the broad categories of assay, although estimates by IFMA tended to be slightly larger (mean, 7.60 mIU per ampoule; 95% limits, 7.02–8.22 mIU per ampoule) than estimates by the other methods. The geometric mean of all laboratory and method geometric means was 7.30 mIU per ampoule (7.00–7.60 mIU per ampoule; n = 37), omitting data as described previously.

Table 4. Geometric mean estimates of the activity of the various in-house standards expressed as IU of the IRP equivalent in activity to 1 nominal IU of in-house standard.

Laboratory code	Assay type	In-house standard relative to IRP
01[a]	IFMA	1.04
03	IRMA	1.12
05	IRMA	0.90
06	ELISA	0.63
08	ELISA	1.08
08	IRMA	1.13
09	IRMA	1.26
10	IFMA	0.99
10A	IFMA	1.13
12[a]	IFMA	0.59
15	IRMA	0.98
16	IRMA	0.96
17	ELISA	1.08
17	IRMA	0.86
17	RIA	1.01
20	ELISA	1.04
23[a]	ICLMA	0.94
26	IRMA	0.97
26	IFMA	0.97
28	IRMA	1.02
28	RIA	0.78

[a] Raw data for in-house standard not available; estimate based on values supplied by participant.

comparison of rTSH 94/674 and pituitary hTSH 81/502

Laboratory geometric mean estimates for each of the duplicate ampoules of 94/674, expressed as milliunits of 81/502 using the value of 7.75 mIU per ampoule previously assigned to 81/502, are shown in Fig. 3. The between-laboratory mean square, pooled over all laboratories and methods, for these estimates was approximately fourfold larger than that obtained for the comparison of the coded duplicate preparations. However, the between-laboratory mean squares for IFMA or RIA assays considered separately were similar to those obtained for the coded duplicate preparations, although the geometric mean estimates by these two methods differed significantly, being 7.29 mIU (6.93–7.67 mIU) per ampoule by IFMA or 6.38 mIU (6.17–6.60 mIU) per ampoule by RIA. In contrast, the between-laboratory mean square for estimates by ICLMA was approximately 10-fold larger than that of the coded duplicate preparations, and these methods gave both the smallest and largest estimates obtained for this comparison, from 5 to 10 mIU per ampoule. Both for this comparison and for the comparison of 94/674 with the IRP, the ICLMA in use in laboratory 14 gave substantially larger estimates than those given by the other ICLMAs.

stability of rTSH 94/674 and pituitary TSH 81/502

Estimates of the activity of the sample of 94/674 that had been stored for 215 days at 20 °C relative to the samples stored continuously at −20 °C showed no detectable loss of activity. The geometric mean of laboratory geometric mean estimates for the sample stored at 20 °C was 0.99 ampoules (95% limits, 0.95–1.02 ampoules; n = 20) of 20 °C sample having equivalent activity to one ampoule of 94/674. Samples stored for 215 days at 37 or 45 °C showed substantial losses of activity in nearly all assay systems. Predicted losses of activity for samples stored continuously at −20 °C vary depending on results used and in some cases, show poor agreement with the theoretical model. However, the apparent immunological stability of the samples stored at 20 °C, and the predicted stability when this information is combined with the apparently moderate losses of activity of samples stored at 37 or 45 °C indicate that 94/674 under the usual conditions of storage at −20 °C is likely to be sufficiently stable to serve as a standard.

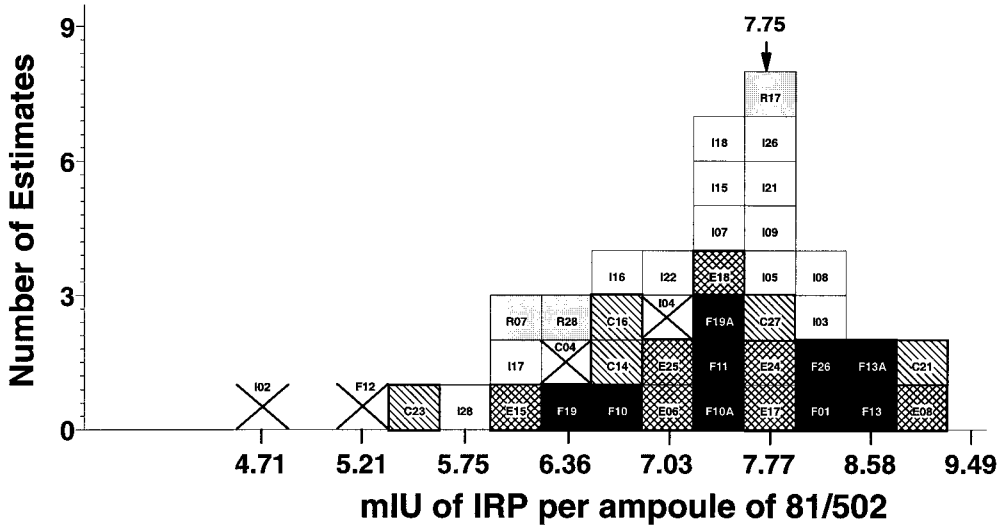

Fig. 2. Laboratory geometric mean estimates for ampoules of 81/502 expressed as mIU of the IRP per ampoule.

The *arrow* shows the value assigned previously by a collaborative study. For key, see legend for Fig. 1.

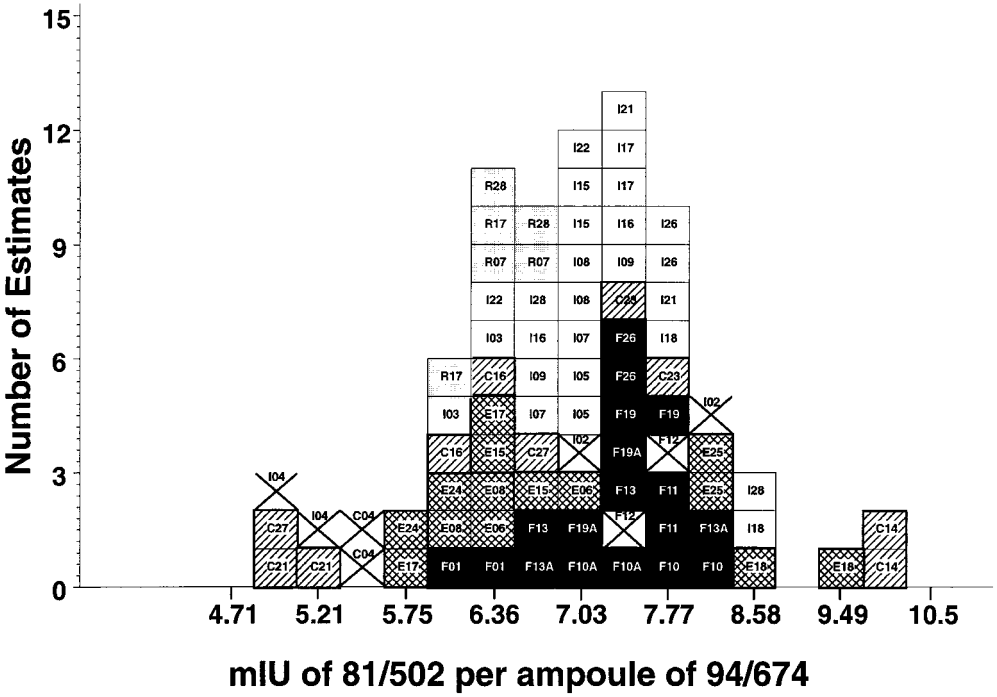

Fig. 3. Laboratory geometric mean estimates for duplicate ampoules of 94/674 expressed as mIU of the pituitary preparation 81/502, using its assigned value of 7.75 mIU per ampoule.

For key, see legend for Fig. 1.

The samples of 81/502 had been stored at relatively high temperatures for a considerable time, ~15 years. The immunologically detectable loss of activity ranged from ~15% to undetectable for the samples stored at 20 °C (mean remaining activity, 95%; 95% limits, 90–99%) and from 35% to 10% for the samples stored at 37 °C (mean remaining activity, 76%; 95% limits, 71–83%), indicating that, as for 94/674, the measured loss of activity is dependent on the assay system used for detection. (This suggests that the degradative changes induced in the molecule variously affect antibody recognition). Using the remaining activity at 20 and 37 °C of either 95% and 65%, as obtained by the system seeing the largest loss of activity, or 96% and 85%, as obtained by the system seeing the least loss of activity, gives predicted immunologically detectable losses of activity of <0.01% per year for the samples stored continuously at −20 °C.

Discussion

Many internationally recognized reference preparations have been produced for biotechnology products, including growth hormone *(12)*, erythropoietin *(13)*, and follicle-stimulating hormone *(14)* along with numerous cytokines and growth factors. However, replacement of the native material by a recombinant preparation may lead to problems if the systems for which the standards are intended are not fully characterized. Glycosylation and other posttranslational modifications are highly cell- and process-dependent. The nature of these modifications may lead to differential recognition of the preparations, depending on the characteristics of the assay systems used, e.g., antibody specificity, leading to inappropriate relative estimates of activity. Pituitary TSH differs from rTSH in that the former contains predominantly sulfated oligosaccharides *(15)*, whereas the oligosaccharides of the CHO cell-derived material are exclusively sialylated *(5)*. The degree of glycosylation, particularly terminal sialylation, has been shown to affect the in vivo and in vitro bioactivities of TSH *(6, 16)*, but variations in glycosylation can also influence immunoreactivity, as discussed for the glycoprotein hormones in general *(17)* and TSH in particular *(18)*.

Comparison of the ampouled preparation of rhTSH (94/674) with the IRP (80/558) showed only limited evidence of nonparallelism in the 41 method/laboratory combinations included in the study. However, there was substantial variability between laboratory estimates of 94/674 relative to the IRP. The extent to which this was system-based is not clear; estimates by RIA were smaller than estimates from the other categories of assays, whereas estimates by IFMA were generally larger. The lower immunopotency of rhTSH by RIA compared with two-site noncompetitive systems has been described previously *(7)* and may be attributable to differential recognition by the antibody of labeled and unlabeled ligand from different sources. Other factors, such as surface adsorption or matrix effects, may influence assay performance by affecting recovery of material from the ampoules. However, most participants performed the initial dilutions in a bovine serum albumin- or serum-based diluent, and in addition, protein protection was offered by human serum albumin as part of the ampoule content formulation. Examination of the data from the largest category of assays (IRMA; n = 15) showed a group of four laboratories with a smaller mean estimate than the remaining members. However, there was no obvious cor-

relation of initial dilution volume and primary and secondary diluent composition with assay estimate. During the course of the study, reconstitution of ampoule contents in water was shown to cause problems in several laboratories; one laboratory reported nonlinearity of responses (data not analyzed), and estimates from two others were substantially smaller than those obtained in the majority of other assay systems. The incorporation of two different preparations of pituitary TSH (IRP and 81/502) provided evidence that at least some assay systems may distinguish between rhTSH and pituitary TSH because the interlaboratory variability for comparison of either pituitary preparation and the rhTSH was twofold greater than that for comparison of the two pituitary preparations. It may be, for example, that the large estimates obtained by laboratory 14, which used ICLMA, or laboratory 18, which used ELISA, for 94/674 relative to either the IRP or 81/502 reflect a relatively greater selectivity of these systems for the rhTSH than for the pituitary TSH, whereas the reverse is true for many of the ICLMAs and ELISAs, which gave generally lower estimates for these comparisons (Figs. 1 and 3).

The single geometric mean of all assay estimates for the coded duplicates of 94/674 relative to the IRP, excluding estimates in terms of an in-house standard or with water as primary diluent, was 6.59 mIU (6.28–6.92 mIU; n = 38) per ampoule. The mean estimate by RIA was smaller than estimates from the other categories of assays. RIA has mostly been superseded by two-site noncompetitive assays (only 3 laboratories of 28 offered RIA) and is likely to be replaced completely in the future. Removal of the RIA data gave a geometric mean of 6.70 mIU (6.38–7.02 mIU; n = 35) per ampoule.

The single geometric mean of all estimates of 94/674 relative to the pituitary preparation 81/502 was 6.91 mIU (6.60–7.23 mIU; n = 41) per ampoule. The potency (7.75 mIU per ampoule) of 81/502 had been assigned previously in the international study that established the Second IRP 80/558 *(1)*; therefore, inclusion of 81/502, together with samples of it stored continuously at increased temperatures for 15 years, also allowed the possible loss of immunoreactivity of both preparations to be monitored over this extended period. Although there was little evidence of lack of stability, the mean estimate of potency relative to the IRP in the current study was 7.30 mIU per ampoule (7.00–7.60 mIU per ampoule; n = 37). It is possible that this apparent decline in potency is a reflection of the optimization of assays in the intervening 15 years toward the pituitary preparation contained in the second IRP, or although less likely, because the potency was initially assigned in terms of both immunoassay and bioassay. Because there are significant ($P < 0.05$) interlaboratory differences in estimates, the possibility cannot be excluded that inclusion or exclusion of particular laboratories may have affected the overall mean estimate, both in this and the previous study. For example, IRMAs, excluding laboratories 16, 17, 22, and 28, gave a mean of 7.76 mIU per ampoule, and estimates from several laboratories were consistent with the value of 7.75 mIU per ampoule (Fig. 2).

Although it appeared that some assay systems distinguished between pituitary and recombinant TSH, the results of the study showed that assignment of a value of 0.0067 IU (6.70 mIU) per ampoule of rhTSH (94/674) would give reasonable continuity with the Second IRP for TSH in many assay systems. It appeared to be sufficiently stable to serve as a reference reagent because the predicted loss of immunoreactivity when stored continuously at −20 °C was <0.01% per year. The preparation was also shown to have appropriate biological activity. At its 47th meeting, therefore, the Expert Committee on Biological Standardization of WHO established the preparation in ampoules coded 94/674 as the First WHO Reference Reagent for TSH, human recombinant, with an assigned value of 0.0067 IU per ampoule *(19)*.

It is not intended that the recombinant preparation should replace the current IRP because there are sufficient stocks of the IRP to last for many years. Rather, it is envisaged that, as companies gradually change over to the use of recombinant DNA-derived materials as kit calibrators, a reference reagent that has been assessed in a variety of assay systems relative to the IRP [and which may have similarities to forms of circulating TSH seen in some clinical conditions *(20)*] will provide a means of validating assay performance of those systems that may distinguish between the native and recombinant materials. By characterizing the system in this manner, the assigned value would maintain continuity with the IRP and should not compromise the clinical data.

We gratefully acknowledge the important contributions of all the participants, Genzyme Corporation (Framingham, MA) for kindly donating the recombinant TSH, Julie Hopkins for preliminary analysis of the data, Margaret Lewis for assistance with the manuscript, Gill Creeber for data entry, and Dr. P. Dawson and the Standards Division for preparation of the ampouled materials.

References

1. Gaines Das RE, Bristow AF. The Second International Reference Preparation of Thyroid-stimulating Hormone, Human, for Immunoassay: calibration by bioassay and immunoassay in an international collaborative study. J Endocrinol 1985;104:367–79.
2. Cole ES, Lee K, Lauziere K, Kelton C, Chappel S, Weintraub B, et al. Recombinant human thyroid stimulating hormone: development of a biotechnology product for detection of metastatic lesions of thyroid carcinoma. Biotechnology 1993;11:1014–24.
3. Hussain A, Zimmerman CA, Boose JA, Froehlich J, Richardson A, Horowitz RS, et al. Large scale synthesis of recombinant human thyrotropin using methotrexate amplification: chromatographic, immunological and biological characterization. J Clin Endocrinol Metab 1996;81:1184–8.
4. Ladenson PW, Braverman LE, Mazzaferri EL, Robbins J, Cooper DS, Garber JR, et al. Comparison of recombinant human thyrotropin administration to thyroid hormone withdrawal for radioactive

iodine scanning in patients with thyroid carcinoma. N Engl J Med 1997;337:888–96.
5. Szkudlinski MW, Thotakura NR, Bucci I, Joshi LR, Tsai A, East-Palmer J, et al. Purification and characterization of recombinant human thyrotropin (TSH) isoforms produced by Chinese hamster ovary cells: the role of sialylation and sulfation in TSH bioactivity. Endocrinology 1993;133:1490–503.
6. Thotakura NR, Szkudlinski MW, Weintraub BD. Structure-function studies of oligosaccharides of recombinant human thyrotropin by sequential deglycosylation and resialylation. Glycobiology 1994; 4:525–33.
7. Ribela MT, Bianco AC, Bartolini P. The use of recombinant human thyrotropin produced by Chinese hamster ovary cells for the preparation of immunoassay reagents. J Clin Endocrinol Metab 1996;81:249–56.
8. WHO Technical Report Series No. 800 [Report]. Geneva, Switzerland: WHO, 1990:181–214.
9. WHO Technical Report Series No. 802 [Report]. Geneva, Switzerland: WHO, 1992:15–6.
10. Gaines Das RE, Tydeman MS. Iterative weighted regression analysis of logit responses. A computer program for the analysis of bioassays and immunoassays. Comput Programs Biomed 1980;15:13–22.
11. Kirkwood TB, Tydeman MS. Design and analysis of accelerated degradation tests for the stability of biological standards II. A flexible computer program for data analysis. J Biol Stand 1984; 12:207–14.
12. Bristow AF, Gaines Das R, Jeffcoate SL, Schulster D. The First International Standard for Somatotropin: report of an international collaborative study. Growth Regul 1995;5:133–41.
13. Storring PL, Gaines Das R. The International Standard for Recombinant DNA-derived Erythropoietin: collaborative study of four recombinant DNA-derived erythropoietins and two highly purified human urinary erythropoietins. J Endocrinol 1992;134:459–84.
14. Rose MP, Gaines Das RE. Characterisation, calibration and comparison by international collaborative study of international standards for the calibration of therapeutic preparations of FSH. J Endocrinol 1998;158:97–114.
15. Green ED, Baenziger JU. Asparagine-linked oligosaccharides on lutropin, follitropin and thyrotropin. J Biol Chem 1988;263:36–44.
16. Canonne C, Papandreou M-J, Medri G, Verrier B, Ronin C. Biological and immunological characterization of recombinant human thyrotrophin. Glycobiology 1995;5:473–81.
17. Storring PL. Assaying glycoprotein hormones–the influence of glycosylation on immunoreactivity [Review]. Trends Biotechnol 1992;10:427–32.
18. Zerfaoui M, Ronin C. Glycosylation is the structural basis for changes in polymorphism and immunoreactivity of pituitary glycoprotein hormones. Eur J Chem Clin Biochem 1996;34:749–53.
19. WHO Technical Report Series No. 878 [Report]. Geneva, Switzerland: WHO, 1998:16–7.
20. Papandreou M-J, Persani L, Asteria C, Ronin C, Beck-Peccoz P. Variable carbohydrate structures of circulating thyrotropin as studied by lectin affinity chromatography in different clinical conditions. J Clin Endocrinol Metab 1993;77:393–8.