WHO/BS/2024.2480
ENGLISH ONLY

EXPERT COMMITTEE ON BIOLOGICAL STANDARDIZATION
Geneva, 7 to 11 October 2024

Proposed First WHO International Standard for Thyroglobulin Antibodies

Melanie Moore*[1], Jason Hockley[1], Peter Rigsby[1] and Ben Cowper[1]

[1]*Medicines and Healthcare products Regulatory Agency (MHRA), Blanche Lane, South Mimm, Potters Bar, EN6 3QG*
** Study coordinator (melanie.moore@mhra.gov.uk)*

NOTE:
This document has been prepared for the purpose of inviting comments and suggestions on the proposals contained therein, which will then be considered by the Expert Committee on Biological Standardization (ECBS). Comments MUST be received by **6 September 2024** and should be addressed to the World Health Organization, 1211 Geneva 27, Switzerland, attention: Technical Standards and Specifications (TSS). Comments may also be submitted electronically to the Responsible Officer: **Dr Ivana Knezevic** at email: knezevici@who.int.

Summary

A World Health Organisation (WHO) International Reference Reagent (IRR) for Anti-Thyroglobulin Antibodies, coded 65/093, has been available since 1970, and has since been widely used for the calibration of immunoassays for anti-thyroglobulin antibodies (herein referred to as TgAb), used in the diagnosis of autoimmune thyroid diseases. Stocks of 65/093 are now exhausted and therefore a replacement WHO International Standard is required. This study reports the preparation and evaluation of a candidate replacement WHO International Standard for Thyroglobulin Antibodies, coded 23/180, in a collaborative study organised by the UK Medicines and Healthcare products Regulatory Agency (MHRA) acting on behalf of WHO.

Seven laboratories in seven countries returned data, encompassing eighteen different immunoassay methods. The returned data has undergone statistical analysis, giving a range of estimated content of anti-thyroglobulin in 23/180, in terms of 65/093, of 326 to 1251 IU/ampoule, with an inter-laboratory GCV of 50%, a geometric mean of 712 IU/ampoule and a robust geometric mean of 735 IU/ampoule. Whilst the variability is quite high in the estimates of the candidate standard preparation, this is a reduction to that observed for the IRR 65/093, in which estimates in terms of kits standards ranged from 323 to 4001 IU/ampoule, with an inter-laboratory GCV of 77%, indicating that introduction of the candidate standard will not adversely affect the standardisation of TgAb assays. It was noted that continuity between the current IRR and candidate IS may be affected in some assay methods, however, improvements in the harmonisation of serum sample estimates in terms of the candidate standard were also observed, providing further assurance that the candidate standard will enable continued calibration of TgAb immunoassays.

The candidate IS also exhibited commutability with a set of serum samples, analysed as part of this study, in all laboratory methods, demonstrating “calibration effectiveness” for the candidate material. This was a distinct improvement in commutability to that demonstrated with serum samples using either the estimates that were reported by kits (commutable in 12 from 18 laboratories) or estimates calculated relative to the IRR 65/093 (commutable in 13 from 18 laboratories). Although some of this improvement may be due to the use of individual serum samples that made up the candidate standard bulk material, the results highlight the current heterogeneity among TgAb immunoassays, and the ongoing requirement for the provision of a reference material to aid standardisation and harmonisation of these assays.

Stability of the candidate preparation was not directly assessed for this report, but results will be appended prior to the meeting of the WHO Expert Committee on Biological Standardisation (ECBS) at their meeting in October 2024.

It is proposed that the candidate preparation on ampoules coded 23/180 is established as the **First WHO International Standard for Thyroglobulin Antibodies**, with an assigned content of **735 IU per ampoule.**

Introduction

Autoimmune thyroid diseases are common disorders and include several inflammatory diseases with Graves' and Hashimoto's thyroiditis as the most frequent forms. Autoantibodies against thyroid antigens such as thyroglobulin (Tg) are often found in association with these autoimmune diseases and the detection of these autoantibodies is an important component in their diagnosis [1,2]. Thyroglobulin autoantibodies (TgAb) are specific to the thyroglobulin protein, a 660 KDa glycoprotein which plays a major role in thyroid hormone synthesis, storage and release [3]. Measurement of TgAb by immunoassay is used to aid the diagnosis of autoimmune and other hypo/hyper thyroid diseases [4,5]. TgAb may also be measured in tandem with thyroglobulin in the diagnosis and monitoring of differentiated thyroid cancers due to potential interference with the measurement of thyroglobulin antigen in sera [6,7].

A World Health Organisation (WHO) International Reference Reagent (IRR) for anti-thyroglobulin serum, coded 65/093, was established by the WHOs Expert Committee on Biological Standardisation (ECBS) in 1970, and it has been widely adopted for the calibration of immunoassays for TgAb. Stocks of the IRR are now depleted, and given the widespread adoption of the IRR, the proposal to prepare a replacement standard for TgAb was endorsed by the ECBS at their meeting in October 2021. It was anticipated that this replacement would form the 1st International Standard (IS) for Anti-Thyroglobulin Antibodies.

A new preparation of TgAb was filled into ampoules (NIBSC code 23/180), following procedures recommended by WHO [8], and an international collaborative study with expert laboratories was launched to aid in the value assignment of the candidate 1st International Standard by immunoassay. Human serum samples containing a range of TgAb concentrations were included in the study in order to assess commutability of the candidate standard with native samples. This report summarises the production and value assignment of the candidate 1st International Standard for TgAb in this collaborative study. The aims of the study were therefore:

1) To confirm the immunoreactivity of the candidate 1st IS by immunoassay and assess the relationship of this activity with the IRR, 65/093, and existing local standards.
2) To calibrate the candidate 1st IS, in units of IU/ampoule, for TgAb relative to the IRR, 65/093, by immunoassay.
3) To assess the commutability of the candidate 1st IS with native samples in immunoassays.
4) To assess the long-term stability of the candidate 1st IS by accelerated thermal degradation (ATD) study.

Participants

A total of eight laboratories comprising of expert laboratories and manufacturers of TgAb immunoassays participated in the collaborative study, from 7 countries. One laboratory was unable to return data due to resource constraints. The remaining seven laboratories are listed alphabetically by country in Table 1. Ten laboratories in China were co-ordinated through the kind offices of Dr Nan Sun, NIFDC.

Throughout the study, each method provided by the seven laboratories (eighteen in total) is referred to by a code number. The code numbers were randomly assigned and do not reflect the order of listing.

Table 1. List of participants in order of country

BELGIUM	Ms Maxime Rombouts, DiaSource, Rue du Bosquet 2, Louvain-la-neuve, 1348
CHINA	Dr Nan Sun, National Institutes for Food and Drug Control, In Vitro Diagnostic Laboratory 67095322, No 2. Tiantan Xili, Dongcheng District, Beijing, 100050.
FRANCE	Ms Nathalie Auberger Ripoll, BioMerieux, Chemin de l'Orme, 69280 Marcy l'etoile
ITALY	Dr Lorenzo Sangalli, Technogenetics S.p.A., Via della Filanda 24-26, Lodi, 26900
JAPAN	Ms Yuka Imai, Fujirebio Inc, Akasaka Intercity AIR, 1-8-1, Akasaka Minato-ku, Tokyo, 107-0052
UK	Dr Melanie Moore, Medicines and Healthcare products Regulatory Agency, South Mimms Laboratories, Blanche Lane, Potters Bar, EN6 3QG
USA	Dr Paul D'Agostino and Mariah Arcuri, Siemens Healthineers Core Lab Solutions, 511 Benedict Avenue, Tarrytown, 10591

Materials and Methods

Bulk materials and processing

A batch of human serum containing high titre TgAb was produced by pooling 60 mL serum from three donors, purchased from Logical Biological (Sandwich, Kent, UK) and diluting into 746 mL of pooled normal human serum, purchased from TCS Biosciences Ltd (Buckingham, UK) to give a final volume of 926 mL. Aliquots of 0.5 mL were dispensed into glass ampoules, lyophilised and sealed under nitrogen, according to procedures recommended by WHO [8]. The ampoules were assigned a NIBSC code of 23/180. Bulk material processing was carried out on 15th September 2023 at NIBSC (MHRA South Mimms Site, Blanche Lane, Potters Bar, EN6 3QG, UK). Ampoules are stored at -20°C in the dark at this address.

The high titre TgAb serum samples were tested and found negative for anti-HIV-1 and 2, anti-HCV and HBsAg. The bulk pooled normal human serum was tested and found negative for anti-HIV-1 and 2, HBsAg, syphilis and HIV-1 and HCV by PCR.

Product characterisation

A total number of 1719 ampoules, coded 23/180, were produced and are stored at -20°C under temperature-controlled conditions at NIBSC. All 1719 ampoules available are offered to WHO as the candidate 1st International WHO Standard for thyroglobulin antibodies. Check weights measured during filling demonstrated a mean fill weight of 0.515 g, with a low CV of fill of 0.78% (n=237), a mean dry weight of 0.042 g (CV 0.55%, n=6), mean residual moisture of 0.87% (CV 33.6%, n=12) as determined by Karl Fischer and mean oxygen head space of 0.37% (CV 29.7%, n=12). No microbial contamination was detected in the pre-fill or post-filled or post freeze-dried material.

Collaborative study design for the value assignment of 23/180 by immunoassay

Materials

The collaborative study was organised by MHRA. The materials provided to study participants are listed in Table 2. Each participant was allocated a minimum set of samples consisting of the IRR, 65/093, duplicates of 23/180 and a panel of 9 human serum samples.

Human serum samples for inclusion in the study were prepared from the individual serum donations that made up the bulk material (Logical Biological, Kent, UK) or from serum obtained from Abbaltis Ltd (Sittingbourne, Kent, UK) diluted into normal human serum (TCS Biosciences, Buckingham, UK) at varying concentrations. Samples were coded TgAbSerum1 to TgAbSerum9. All samples were tested and found negative for anti-HIV-1 and 2, HBsAg and HCV.

Table 2. Preparations provided to participants

TgAb preparation	Ampoule content
International Reference Reagent for TgAb, 65/093, coded CS731 sample C	1000 IU per ampoule
Coded preparations of the candidate 1st IS, 23/180, coded CS731 sample A and CS731 sample B	Nominally 1000 IU per ampoule

Nine human serum samples labelled CS731 TgAbSerum1 to CS731 TgAbSerum9	0.25 ml* aliquots human serum

*Aliquot volume varied according to assay requirements.

Methods contributed

Participants were requested to perform the immunoassay normally used in house for their measurement of anti-Tg. The immunoassay methods contributed, in alphabetical order, were: Abbott Alinity, Abbott Architect, Siemens Atellica IM, Autobio, Siemens Centaur XP, Elecsys COBAS, DiaSource, Hokapi, Diasorin Liaison XL, Fujirebio Lumipulse, Maccura, Maglumi, Mindray, Kehua Polaris, Tellgen, VIDAS, Shenzhen YHLO, ZENIT RA and Zybio.

Participants were requested to perform three independent runs of the assay method(s) normally used in their laboratory, and a copy of the study protocol provided to participants can be found in Appendix 1. As ampoules of 65/093 were in very limited supply, participants were asked to perform two independent runs with a full sample panel that included dilutions of 65/093, 23/180 and serum samples, using fresh ampoules/tubes for each assay run, and a third independent run with a reduced sample panel consisting of 23/180 and serum samples only. If the required assay dead volumes allow, participants were requested to freeze a stock solution of 65/093 for use in the third independent run.

Participants were asked to prepare a minimum of 5 dilutions of the ampouled preparations, including additional dilutions where needed to ensure that a minimum of 5 points in the linear part of the dose response curve were measured. An example core dilution series was provided.

A total of eighteen different immunoassay methods were contributed to the study from the 7 laboratories listed in Table 1.

Stability assessment

A thermally accelerated degradation (ATD) study has been designed to assess the stability of the candidate preparation 23/180. Coded samples of the candidate preparation, 23/180, are currently stored at elevated temperatures of +4º, +20º, +37º and +45º C and will be analysed once they have been stored at these temperatures for a period of 6 months and data will be appended to this report.

Data and Statistical Analysis

Participants were requested to provide all raw assay data for central computation at MHRA along with participants own estimates of activity as calculated by the method normally used in their laboratory.

Potency assignment of 23/180

Analysis was based on the results supplied by the participants, reported in IU/mL. To calculate laboratory reported potency estimates, the reported IU/mL values were corrected for the corresponding dilution factors. Then the geometric mean value at each dilution was calculated for each sample within an assay run, and a single estimate was then obtained for each sample within an assay run by taking the geometric mean value across the dilutions. A single estimate for each sample within the laboratory was then calculated as the geometric mean of the estimates of IU/mL across assay runs.

Relative potencies were estimated using a parallel line model with $\log_{10}$-transformed IU/mL values as responses [9]. Calculations were performed using the software package, R [10]. Estimates from all valid assays were combined to generate a geometric mean value for each laboratory. Overall sample estimates were calculated as the geometric means of all individual laboratories. Variation between laboratories (inter-laboratory) was expressed as % geometric coefficient of variation (GCV = $\{10^s-1\}\times100\%$ where s is the standard deviation of the $\log_{10}$ transformed estimates) of the estimates. To provide estimates that reduce the influence of more extreme laboratory potency estimates Huber's robust mean was calculated using the R package "WRS2" [11].

Assessment of commutability

Commutability of the candidate IS, 23/180, and the current IRR, 65/093, was assessed using a "calibration effectiveness" approach [12]. The analysis uses bias values calculated from reported estimates and from estimates recalculated relative to 65/093 or 23/180 using the fitted dose-response data for the WHO preparation. Bias was defined as the laboratory geometric mean estimate for the sample reported as a ratio to the study consensus value for the sample (the study median value for the samples were used as sample target values for the purposes of this analysis).

Assessment of stability

Samples of candidate standard are currently stored at elevated temperatures (+4, +20, +37, +45^{o}C) and a reference temperature (-20^{o}C), and will be analysed via immunoassay, with the intention of fitting an Arrhenius equation relating degradation rate to absolute temperature assuming first-order decay [13], and thus predict the degradation rate when stored at -20°C.

Results

Data returned for analysis

A total of 106 immunoassay measurements of 23/180 were performed and returned for central analysis from 7 laboratories performing a total of 18 different immunoassay methods. All assays included kit controls/standards and met the participant's associated acceptance criteria.

Validity criteria

Relative potencies were estimated using a parallel line model. Where relevant, non-parallelism was assessed by calculation of the ratio of fitted slopes for the candidate IS, 23/180, and the reference reagent 65/093. The samples were concluded to be non-parallel when the slope ratio was outside of the range 0.90 – 1.11 and in these cases, no estimates are reported (marked NP in tables). Slope ratios for all laboratories are shown graphically in Figure A2.1 in Appendix 2. Ratios ranged from 0.60 to 1.27, with 79% of assays within the range of 0.90 to 1.11. Using this range, all estimates from Lab 15 and Lab 17 were found to be non-parallel, and results were therefore excluded from overall calculations. In Lab 14, estimates of 23/180-A from Assay 1 and estimates of both coded duplicates in Assay 2 were excluded from overall calculations due to non-parallelism, as were estimates of both coded duplicates in Assay 1 of Lab 18.

Estimated potency of the candidate IS, 23/180

Potency estimates were calculated relative to kit standards from results supplied by participants in IU/mL, and calculated relative to the current IRR, 65/093, using a parallel line model. Individual laboratory estimates of 65/093 and 23/180, calculated relative to kit standards, are summarised in Table 3 along with individual laboratory estimates of 23/180, calculated relative to 65/093. Overall study results are summarised in Table 4, calculated as the geometric means of all individual laboratories, with Huber's robust geometric mean also calculated to provide estimates that reduce the influence of more extreme laboratory potency estimates.

The estimated content, as reported by kit standards, of the current IRR, 65/093, ranged from 323 – 4001 IU/amp, with a geometric mean of laboratory estimates of 1008 IU/amp and a robust geometric mean of 1003 IU/amp (where IU/mL is equivalent to IU/ampoule), which despite the variability apparent in lab estimates (GCV 77%), is very close to the assigned value of 1000 IU/ampoule for this reference reagent (Tables 3 and 4). A large proportion of this variability is likely due to the very high estimate of 65/093 found in 1 assay method (Lab 5), which gave a similarly high estimate of the content of the candidate standard. As reported by kit standards, 23/180 had an estimated content ranging from 474-3336 IU/amp, a geometric mean of laboratory estimates of 980 IU/amp, a robust geometric mean of 988 IU/amp and a GCV of 87% (Tables 3 and 4). These estimates are shown graphically in Figure 1.

When expressed relative to 65/093, estimates of the candidate standard ranged from 326 – 1251 IU/amp, with a GCV of 50%, a geometric mean of 712 IU/amp and a robust geometric mean of 735 IU/amp (Tables 3 and 4). These estimates are also represented graphically in Figure 2. This data excludes that from laboratories 15 and 17, in which 23/180 and 65/093 were reported as non-parallel, and also excludes Lab 14 and 18 in which those estimates that were not excluded due to non-parallelism, were found to be outliers by Rosner's Test [14]. Despite a reduction in variability when compared to the reported estimates, the GCV is still very high and therefore it is proposed that the Huber's robust geometric mean is used as the overall estimate for the potency of 23/180.

Overall, the data demonstrated the candidate standard 23/180 to be immunoreactive, and to behave in a similar manner to the current IRR 65/093 in the majority of the immunoassays included in this study. However, there are a few notable exceptions where there are differences between reported estimates for the current IRR 65/093 and the candidate IS 23/180 and their respective assigned contents. Labs 4 and 13 in particular have higher estimates of 65/093 compared to its assigned content of 1000 IU/ampoule but underestimate the content of 23/180. Also, although removed from further analysis, Labs 14, 17 and 18 on the other hand have much higher estimates of 23/180 compared to lower estimates of 65/093. A similar pattern is observed for Lab 11, albeit to a lesser extent.

The results indicate that the introduction of 23/180 as the 1st IS to replace the IRR 65/093 will enable continued calibration of the majority of immunoassays for anti-thyroglobulin but that there may be some discontinuity between the standards with some methods. To illustrate if this will have an impact on clinical measurements, estimates of the anti-thyroglobulin content of serum samples, that were measured by all laboratories, are shown in Figure 3, calculated relative to both the candidate standard 23/180 and the current reference reagent 65/093 and compared with estimates of content of these serum samples as reported from kit standards. It is interesting to note that across the serum samples included, estimates when calculated relative to the candidate 23/180 are more harmonised than those expressed relative to kit standards or the current reference reagent, and this appears more significant for those assay methods in which non-parallelism was noted (Labs 14, 15, 17 and 18). A caveat to this apparent harmonisation is that serum samples S2 through S9 are based on the individual donor material that made up the bulk material for 23/180, so are more likely to bear resemblance to the candidate standard than to the current reference reagent 65/093, which has been prepared from a different pool of donor material. Serum 1 however is an independent serum sample, and shows a similar pattern of improvement, giving confidence in the ability of 23/180 to harmonise anti-thyroglobulin measurements.

Table 3. Potency estimates (IU/mL) for 65/093 and 23/180 (NP = Non-parallel; GM = Geometric Mean; * = Not used in further summary calculations)

Lab	Calculated relative to kit standards											Calculated relative to 65/093						
	65/093				23/180							23/180						
	Assay 1	Assay 2	Assay 3	Lab GM	Assay 1		Assay 2		Assay 3		Lab GM	Assay 1		Assay 2		Assay 3		Lab GM
					A	B	A	B	A	B		A	B	A	B	A	B	
1	1117	1037	986	1045	1032	990	980	963	941	912	969	920	882	946	928	954	923	925
2	1076	1186	1192	1150	588	536	611	619	695	662	617	579	531	539	544	593	574	560
3	1543	1559	1578	1560	1213	1230	1247	1192	1203	1188	1212	776	788	790	754	752	743	767
4	1267	1398	1429	1363	480	519	567	568	577	565	545	354	378	371	369	372	364	368
5	4063	3941		4001	3394	3291	3337	3308	3347	3342	3336	785	753	799	790			782
6	1066	1073	1080	1073	663	671	646	639	648	649	653	621	630	601	594	598	598	607
7	926	932	927	928	540	543	548	554	541	549	546	556	560	557	564	558	565	560
8	976	1003		990	797	790	793	832	833	780	804	816	810	789	828			811
9	875	868		871	1045	1101	1049	1017	1001	998	1035	1188	1249	1202	1167			1201
10	999	1017		1008	1038	981	1023	985	1049	977	1008	1042	982	1006	967			999
11	985	994	984	987	1182	1203	1251	1255	1244	1242	1229	1204	1228	1267	1270	1271	1270	1251
12	1000	991	984	992	568	559	572	570	572	572	569	572	563	578	576	581	582	575
13	1914	1363*		1914	644	653	452*	208*	389	459	524	322	330	*NP*	*NP*			326
14	986	1009		998	3625	3954	3828	3940	3808	3827	3829	*NP*	3994	*NP*	*NP*			*3825**
15	317	328		323	480	466	476	460	514	451	474	*NP*	*NP*	*NP*	*NP*	*NP*	*NP*	*n/a*
16	1121	1021		1070	1093	1170	954	951			1043	984	1045	935	933			973
17	679	675	709	688	2069	2089	2007	1994	1965	1954	2018	*NP*	*NP*	*NP*	*NP*	*NP*	*NP*	*n/a*
18	269	326		296	1233	1247	1324	1253	1299	1217	1261	*NP*	*NP*	4057	3915			*3985**

Table 4. Summary calculations on laboratory estimates (IU/mL) for 65/093 and 23/180
(GM = Geometric Mean; CI = Confidence Interval; GCV = Inter-lab geometric coefficient of Variation (%))

	Relative to Kit Standards		Relative to 65/093	
	65/093	**23/180**	**23/180**	**23/180 (outliers excluded)**
GM	1008	980	883	712
95%	758 – 1341	723 – 1329	608 – 1282	563 – 900
GCV	77%	87%	101%	50%
Robust GM	1003	988	803	735

Figure 1 Histogram graph – potency estimates of 23/180 relative to kit standards

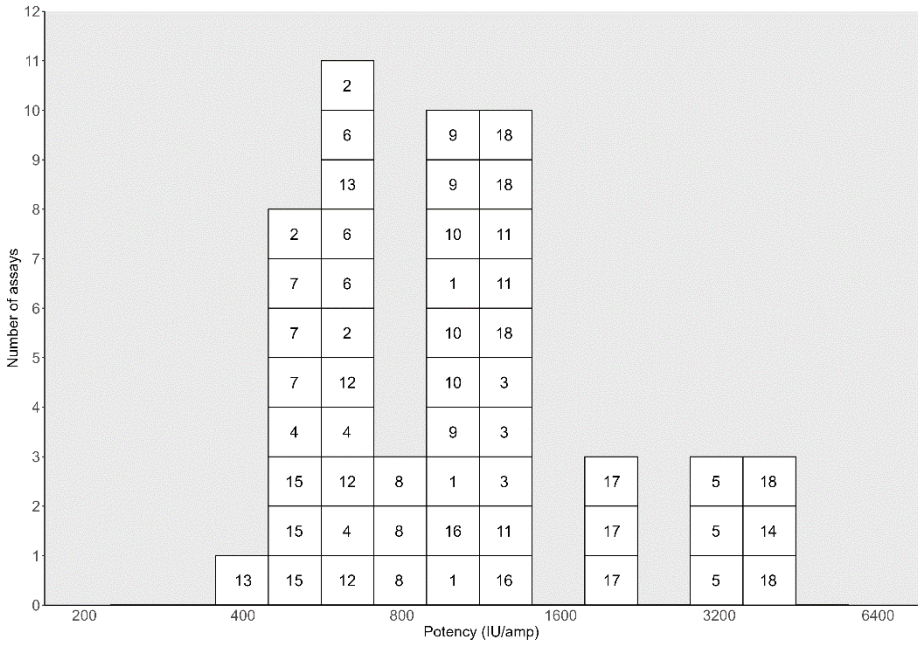


Figure 2 Histogram graph – potency estimates of 23/180 relative to 65/093

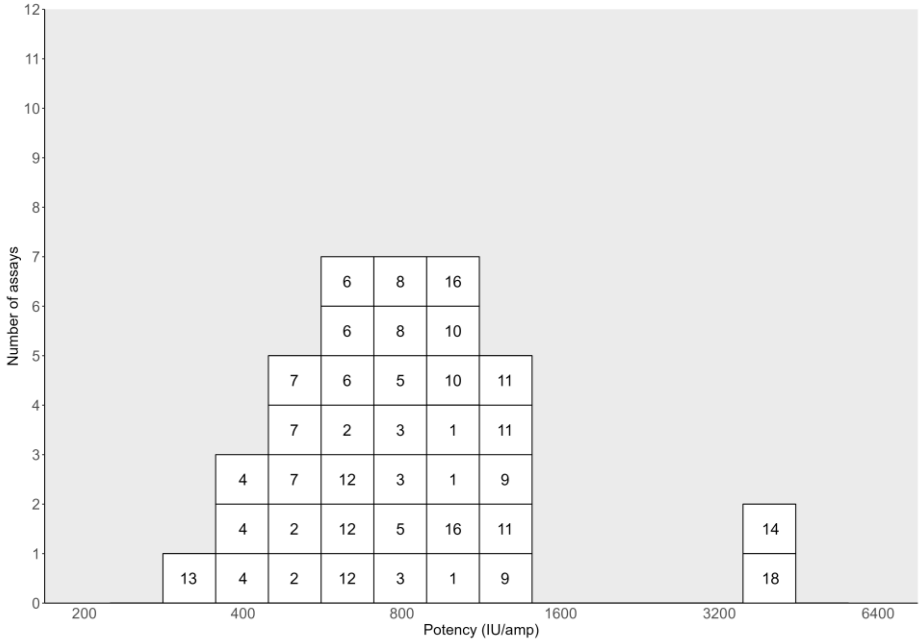
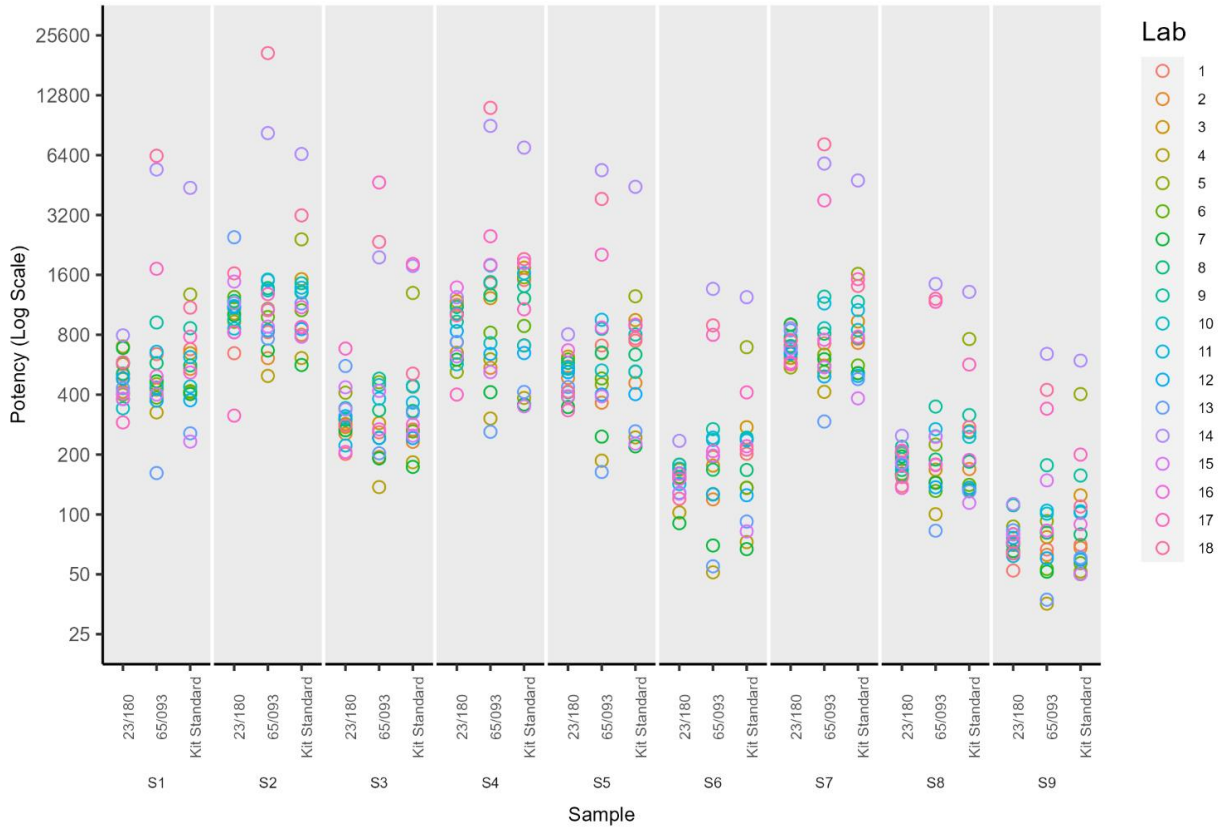


Figure 3 Potency estimates of serum and plasma samples as reported by laboratories and when expressed relative to either standard preparation



Commutability of the candidate IS, 23/180, and current IRR, 65/093

The commutability of the candidate IS, 23/180, and the current IRR, 65/093, with human serum and plasma samples was assessed for all methods included in the study using a difference in bias approach. This approach measures calibration effectiveness [12], by examining the observed method bias for clinical samples using results as reported in a range of laboratory methods, and the impact on the bias for these samples when results are re-calibrated relative to a reference material. Where calibration is effective at reducing bias for clinical samples (harmonising results), this indicates that the reference material is commutable with these samples in the methods used.

Data used for the assessment of commutability are shown in Appendix 2, Tables A2.1-A2.3. Table A2.1 shows the geometric mean reported estimates for the serum samples in each laboratory, taken as the reported concentrations from whichever assay kits they used. Geometric mean estimates calculated relative to 65/093 and 23/180 by parallel line analysis are shown in Tables A2.2 and A2.3 respectively, assuming a content of 1000 IU/amp for 65/093 and 735 IU/amp for 23/180 when used as standard. Overall study estimates for each sample are shown as geometric mean and median estimates (calculated using log transformed data), with inter-laboratory variation expressed as GCV and anti-logged MAD (Median Absolute Deviation) values.

Bias values were calculated as the laboratory geometric mean estimate as a ratio of the study median value for the sample and are shown in Table 5-7 and Figures 2-4. In order to derive an acceptable bias range (for analysis of this study only), the standard deviation of the log transformed bias values in Table 6 was calculated within each laboratory, and the value, $S_{M,}$ (median standard deviation) was calculated across log-transformed values for all laboratories. Criteria representing the maximum bias range were then set at $\pm 0.3 S_M$. The range determined in this study was ±0.241, or 0.575 to 1.740 on the untransformed scale, i.e the bias should be demonstrated to be not less than 58% and not more than 174% of the study median value. Reference standards were to be concluded as commutable if the observed difference in bias was within these commutability criteria. For this commutability assessment, the bias for plasma and serum samples has been assumed to be constant over the concentration range used.

Table 5. Bias in reported estimates for serum samples (lab GM estimate as ratio of study median value for sample); shaded cells are outside range 0.58-1.74

Sample	Lab																	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
S1	1.23	0.98	1.17	0.75	2.32	0.76	0.73	1.02	1.57	0.80	1.12	0.68	0.46	7.97	0.42	0.94	1.42	1.99
S2	0.77	0.71	1.35	0.54	2.15	0.95	0.50	1.16	1.29	1.17	1.23	0.76	1.02	5.76	0.70	0.98	0.78	2.83
S3	0.82	0.71	1.34	0.56	3.96	0.80	0.53	1.01	1.35	1.35	1.11	0.74	0.99	5.42	0.76	0.86	5.54	1.56
S4	1.32	0.62	1.52	0.34	1.36	0.77	0.31	1.06	1.23	0.62	1.42	0.57	0.36	6.10	0.31	1.60	0.94	1.68
S5	1.08	0.66	1.38	0.35	1.81	0.76	0.32	0.92	1.16	0.76	1.29	0.58	0.38	6.43	0.33	1.30	1.30	1.11
S6	0.98	0.66	1.33	0.35	3.35	0.66	0.32	0.81	1.18	1.16	1.14	0.60	0.45	5.98	0.40	1.07	1.99	1.02
S7	0.98	0.94	1.20	0.66	2.08	0.72	0.66	0.99	1.51	1.09	1.37	0.64	0.62	6.14	0.49	1.01	1.96	1.81
S8	0.99	0.91	1.41	0.73	4.08	0.75	0.71	0.99	1.69	1.31	1.39	0.73	0.70	7.04	0.61	1.01	3.04	1.47
S9	0.84	0.81	1.48	0.61	4.79	0.68	0.60	0.94	1.87	1.21	1.23	0.70	0.72	7.05	0.59	1.06	2.38	1.30
Median bias	0.98	0.71	1.35	0.56	2.32	0.76	0.53	0.99	1.35	1.16	1.23	0.68	0.62	6.14	0.49	1.01	1.96	1.56
GM bias	0.99	0.77	1.35	0.52	2.66	0.76	0.49	0.99	1.41	1.02	1.25	0.66	0.59	6.39	0.49	1.07	1.81	1.57
95% LCL	0.86	0.67	1.26	0.41	1.91	0.70	0.38	0.91	1.24	0.82	1.16	0.61	0.44	5.83	0.38	0.93	1.14	1.24
95% UCL	1.13	0.87	1.44	0.67	3.69	0.82	0.65	1.06	1.60	1.26	1.34	0.72	0.80	7.00	0.63	1.24	2.88	1.99
Slope	0.07	-0.04	-0.02	-0.10	-0.43	0.07	-0.11	0.07	-0.11	-0.18	0.03	-0.04	-0.09	-0.04	-0.13	0.07	-0.44	0.20

Table 6. Bias in estimates for serum samples calculated relative to 65/093 (lab GM estimate as ratio of study median value for sample); shaded cells are outside range 0.58 – 1.74

Sample	Lab																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
S1	1.39	0.92	0.94	0.71	0.98	0.84	1.02	1.26	2.01	0.95	1.43	0.81	0.35	11.80	0.87	1.08	3.74	13.83
S2	0.77	0.57	1.00	0.47	1.01	0.92	0.63	1.29	1.42	1.26	1.41	0.79	0.72	7.75	0.82	1.00	1.20	19.56
S3	0.83	0.63	0.93	0.44	1.49	0.78	0.62	1.08	1.55	1.44	1.23	0.78	0.65	6.32	1.35	0.86	15.05	7.56
S4	1.45	0.55	1.22	0.30	0.60	0.82	0.41	1.27	1.47	0.73	1.78	0.64	0.26	8.97	0.52	1.80	2.50	11.06
S5	1.21	0.62	1.10	0.32	0.77	0.82	0.42	1.11	1.46	0.91	1.63	0.68	0.28	9.19	0.67	1.49	3.46	6.59
S6	0.99	0.61	0.90	0.26	1.01	0.65	0.36	0.86	1.37	1.21	1.24	0.64	0.28	6.96	1.01	1.06	4.09	4.56
S7	1.07	0.82	0.94	0.61	0.94	0.77	0.89	1.20	1.83	1.27	1.69	0.73	0.43	8.56	0.81	1.11	5.58	10.69
S8	0.97	0.79	0.92	0.55	1.22	0.71	0.78	1.03	1.90	1.35	1.46	0.75	0.45	7.85	1.34	0.96	6.39	6.61
S9	0.81	0.76	0.94	0.43	1.13	0.65	0.63	0.99	2.16	1.23	1.27	0.73	0.45	7.83	1.81	1.01	4.15	5.16
Median bias	0.99	0.63	0.94	0.44	1.01	0.78	0.63	1.11	1.55	1.23	1.43	0.73	0.43	7.85	0.87	1.06	4.09	7.56
GM bias	1.03	0.69	0.98	0.43	0.99	0.77	0.60	1.11	1.67	1.12	1.45	0.73	0.41	8.24	0.96	1.12	4.14	8.56
95% LCL	0.86	0.60	0.91	0.33	0.81	0.70	0.46	1.00	1.46	0.95	1.31	0.68	0.31	7.18	0.71	0.94	2.44	5.93
95% UCL	1.23	0.79	1.06	0.56	1.21	0.84	0.80	1.23	1.90	1.34	1.61	0.78	0.54	9.45	1.29	1.34	7.01	12.37
Slope	0.11	-0.08	0.07	0.01	-0.18	0.12	-0.01	0.13	-0.10	-0.12	0.11	0.00	-0.02	0.07	-0.39	0.14	-0.38	0.44

Table 7. Bias in estimates for serum samples calculated relative to 23/180 (lab GM estimate as ratio of study median value for sample); shaded cells are outside range 0.58 – 1.74

Sample	Lab																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
S1	1.06	1.19	0.86	1.46	0.84	0.99	1.41	1.05	1.20	0.72	0.79	1.01	0.93	1.64	0.86	0.77	0.60	1.26
S2	0.62	0.78	0.97	1.03	0.88	1.17	0.94	1.10	0.90	1.06	0.81	1.04	2.08	1.08	1.43	0.75	0.30	1.61
S3	0.68	0.86	0.92	0.95	1.38	1.00	0.88	1.00	1.01	1.17	0.74	1.04	1.89	1.11	1.48	0.68	2.26	0.96
S4	1.30	0.84	1.34	0.74	0.59	1.17	0.67	1.24	1.05	0.66	1.15	0.95	0.83	1.40	0.72	1.52	0.45	1.19
S5	1.12	0.97	1.24	0.78	0.78	1.19	0.68	1.14	1.07	0.83	1.08	1.03	0.90	1.58	0.79	1.29	0.66	0.85
S6	1.00	1.02	1.10	0.66	1.17	1.00	0.59	0.99	1.10	1.19	0.92	1.05	0.96	1.50	0.84	1.02	0.98	0.80
S7	0.83	1.09	0.88	1.29	0.79	0.93	1.28	1.00	1.11	1.01	0.95	0.92	1.18	1.23	1.00	0.81	0.81	1.11
S8	0.76	1.03	0.87	1.10	1.10	0.86	1.05	0.90	1.17	1.01	0.84	0.95	1.21	1.32	0.99	0.72	1.12	0.85
S9	0.73	1.12	1.01	0.95	1.20	0.87	0.89	1.00	1.52	1.00	0.84	1.05	1.35	1.52	1.11	0.86	0.99	0.88
Median bias	0.83	1.02	0.97	0.95	0.88	1.00	0.89	1.00	1.10	1.01	0.84	1.03	1.18	1.40	0.99	0.81	0.81	0.96
GM bias	0.87	0.98	1.01	0.97	0.94	1.01	0.90	1.04	1.12	0.94	0.89	1.00	1.20	1.36	1.00	0.90	0.78	1.03
95% LCL	0.72	0.88	0.89	0.79	0.77	0.92	0.71	0.97	1.00	0.80	0.80	0.96	0.93	1.21	0.82	0.73	0.50	0.86
95% UCL	1.06	1.09	1.14	1.18	1.16	1.11	1.12	1.12	1.25	1.11	1.00	1.04	1.54	1.53	1.21	1.12	1.22	1.23
Slope	0.06	-0.08	0.03	0.05	-0.23	0.10	0.05	0.07	-0.12	-0.10	0.05	-0.02	0.00	-0.06	-0.02	0.07	-0.42	0.20

Figure 2. Bias in reported estimates for serum samples (lab GM estimate as ratio of study median value for sample)

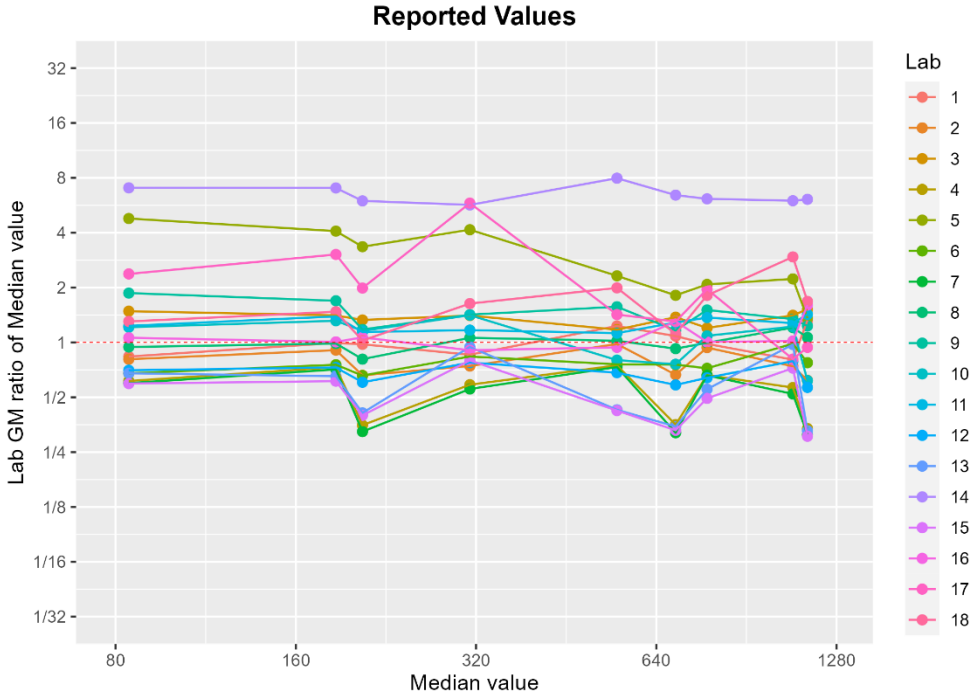


Figure 3. Bias in estimates for serum samples calculated relative to 65/093 (lab GM estimate as ratio of study median value for sample)

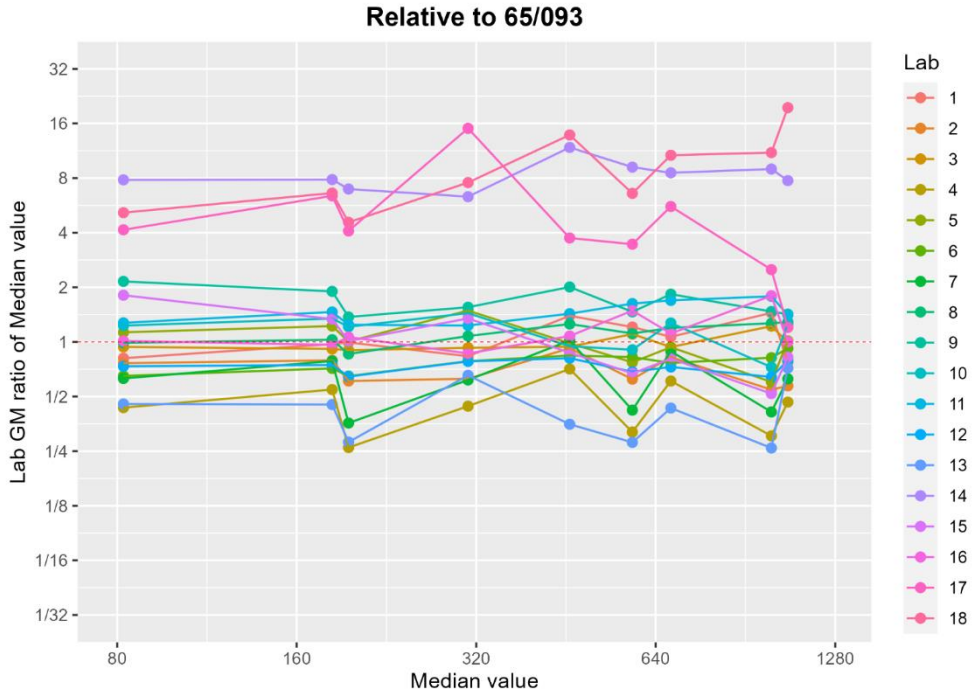
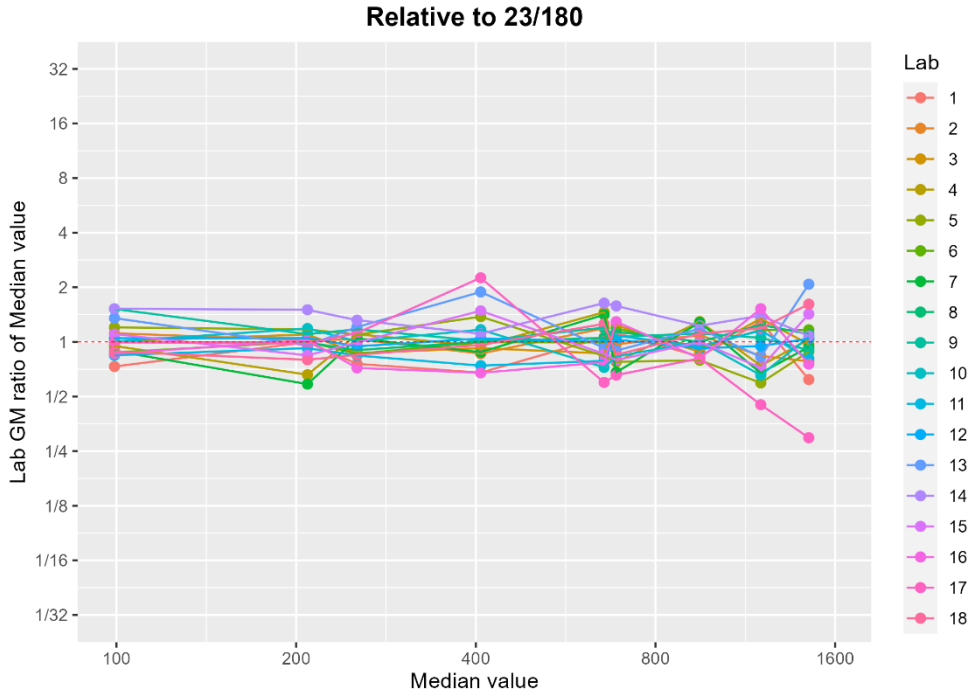


Figure 4. Bias in estimates for serum samples calculated relative to 23/180 (lab GM estimate as ratio of study median value for sample)



As illustrated in Table 5 and Figure 2, the study samples as reported by the majority of laboratory methods (12 out of 18 labs) fall within the commutability criteria of 0.58 – 1.74 (58 – 174%). There were a few exceptions, where all, or the majority, of samples were outside these criteria (Labs 4, 5, 7, 14, 15 and 17). Two labs showed a mixture of bias across the samples (Labs 13 and 18); however, the majority were commutable.

When calculated relative to 65/093, commutability was demonstrated with samples in 13 laboratory methods. A similar pattern of commutability was maintained, with 65/093 remaining either commutable with samples (Labs 1, 2, 3, 6, 8, 9, 10, 11, 12 and 16) or non-commutable with samples (Labs 4, 14, and 17), which is to an extent expected, as the assay methods in this study all state traceability to 65/093 (Table 6 and Figure 2). However, for some methods, an improvement in commutability was observed (Lab 5, 7 and 15) where the majority of samples were found commutable. However, for Labs 13 and 18, the majority of samples fell outside of the criteria when calculated relative to 65/093, indicating non-commutability with samples in these laboratory methods. Labs 14, 17 and 18 showed a very high positive bias for serum estimates. Overall, calibration effectiveness was demonstrated for 65/093 for the majority of laboratory assays.

When samples were calculated relative to 23/180 there a was a significant improvement in the number of samples that fell within the set commutability criteria, and in contrast to as reported and relative to 65/093, samples were found commutable with all 18 laboratory methods (Table 7 and Figure 3), indicating calibration effectiveness for 23/180. Full commutability was demonstrated with serum samples in 16 of the laboratory methods, and partial commutability in 2 methods (Labs 13 and 17) in which there was a minority of samples outside the criteria set. The positive bias shown by Labs 14, 17 and 18 was reduced when calculated relative to 23/180.

It should be reiterated here that serum samples 2 to 9 are prepared from the single donors that made up the bulk material. Although these samples can be considered to be individual patient samples, the antibody heterogeneity in these samples will be more reflective of that found in 23/180. Indeed, the bias noted for serum samples as reported by assays (Table 5) reflect the trend in bias observed for 23/180 in terms of kit standards (Table 3). Serum sample 1 however is an independent sample, and the trend to improved commutability of 23/180 in the methods in this study is also reflected with this sample, giving confidence in the introduction of 23/180 as the 1st international standard to enable continued calibration of TgAb immunoassays, and in its potential to aid harmonisation these assays.

Discussion

The current IRR, 65/093, has been widely adopted for the calibration of immunoassays for anti-thyroglobulin antibodies. Stocks of this standard are now exhausted, and a replacement standard is required. This study describes the preparation and evaluation of a candidate 1st International Standard for thyroglobulin antibodies, coded 23/180.

Immunoassay data from 7 laboratories (a total of 18 laboratory methods) were analysed to determine the potency of 23/180 in terms of the IRR 65/093. Estimates of the reference preparation 65/093, when calculated relative to kit standards, were in reasonable agreement with the assigned value of 1000 IU/ampoule, with an overall geometric mean of 1008 IU/amp and a robust geometric mean of 1003 IU/amp, although a large degree of variability was observed with a GCV of 77%. A similarly wide range of estimates for 23/180 was obtained relative to kit standards, with an overall potency estimate of 980 IU/amp, a robust geometric mean of 988 IU/amp and a GCV of 87%. When calculated relative to 65/093, the overall potency for 23/180 was 712 IU/amp, with a robust geometric mean of 735 IU/amp and a reduced GCV of 50%, indicating improved agreement in potency estimates of the standard when expressed relative to the IRR 65/093. Due to the high GCV of the estimated potencies for 23/180, the robust geometric mean of 735 IU/amp is recommended as the estimated potency of the candidate standard 23/180.

Overall, the candidate standard has been demonstrated to be immunoreactive and behaves in a similar manner to 65/093 in the majority of immunoassays included in this study, although discontinuity may be observed with some methods (Labs 4, 13, 14, 15, 17 and 18). This could be due to a number of reasons, including current calibration to the IRR, 65/093, or the use of different buffers to reconstitute and dilute the reference preparations that introduced matrix effects. In particular, the candidate material is not an identical material to that of 65/093. As a heterogenous mix of immunoglobulins with different specificities to different epitopes of the thyroglobulin protein, it is not possible to replace like for like in an autoantibody standard. This heterogeneity is also likely to be responsible for the variability in potency estimates observed for both the reference material preparations, where differences in epitope recognition and bias could occur across the methods, along with a number of other factors including different method signals, contamination with other autoantibodies or proteins (e.g. Tg) and different methods of antigen preparation that could influence the specificity and sensitivity of TgAb assays [15, 16]. It is particularly interesting to note that Labs 14, 17 and 18 are competitive assays, whereas the remaining laboratories are non-competitive, and the discontinuity observed for the candidate standard with these labs in this study suggests that the candidate material may behave differently in this assay format compared with the current IRR 65/093. A statement highlighting that 23/180 comprises a different bulk material would be an informative addition to the IFU, along with a recommendation that users perform their own assessment of the material. A draft IFU has been included in Appendix 3.

Both the candidate standard 23/180 and the NRR 65/093 were also analyzed in comparison with serum samples to determine the calibration effectiveness, as an indicator of commutability of the standards with patient samples, using a difference in bias approach. Commutability of the reference material is indicated where the majority of bias values for serum samples relative to that standard fall within the statistically-derived commutability limits. When expressed relative to kit standards, serum samples were found commutable in the majority of immunoassay methods performed in this study (12/18). When expressed relative to the IRR, 65/093, this pattern of commutability with samples was continued in most methods, with commutability demonstrated in 13/18 labs and non-commutability in 5/18 labs. Commutability was noted to improve in three laboratory methods (Labs 5, 7 and 15) and reduce in two laboratory methods (Labs 13 and 18), indicating a potential issue with current calibration of these assay method. This is particularly noticeable for Lab 5 which had very high bias in reported estimates which was also reflected in estimates of the IRR and 23/180. The assay methods used in this study all state traceability to the IRR, and the data reflects the variability of potency estimates obtained across the immunoassay methods for 65/093. Interestingly, although variability in potency estimates were also noted for 23/180, when serum samples were expressed relative to the candidate standard 23/180, there was a distinct improvement in comparison with those relative to 65/093 or as reported values, and commutability with samples was demonstrated in all methods, indicating calibration effectiveness of the candidate preparation. Although some of this improvement may be due to the use of serum samples that comprised the individual donor material from which 23/180 was prepared, these samples still

represent individual donors, and the inclusion of an independent TgAb donor sample (S1) was able to show a similar level of improvement in commutability with the candidate material.

An analysis of the long-term stability during storage at -20ºC will be performed in an accelerated thermal degradation (ATD) study, using samples of the candidate standard that have been stored at elevated temperatures of +4, +20, +37 and +45°C. These samples will be analyzed after 6 months storage at these temperatures, using 2 immunoassay methods and data will be appended to this report.

In summary, the candidate standard exhibited the expected immunoreactivity, and was found to be commutable with serum samples, in the majority of immunoassays performed in this study. Standardization of a heterogeneous analyte such as that found with autoantibodies is likely to be particularly challenging. Despite potential discontinuity with some methods, it is reassuring to observe improved harmonisation when expressing serum sample estimates relative to the candidate standard, indicating that the preparation 23/180 provides a suitable material to form the 1st International Standard for Thyroglobulin Antibodies, for continued calibration of TgAb immunoassays. It is anticipated that the continued provision of a reference preparation for TgAb is of particular importance to help improve assay harmonisation and that manufacturers will be encouraged to recalibrate their assays where necessary.

Proposal

It is proposed that the candidate preparation in ampoules coded 23/180 is established as the First International Standard for Thyroglobulin Antibodies with an assigned content of 735 IU/ampoule.

Comments from Participants

Minor edits to address details from one participant. No specific comments received from participants.

Acknowledgements

We gratefully acknowledge the important contributions of all the study participants. We would also like to acknowledge the important contributions and support of our MHRA colleagues: the Formulation Science team for preparation of trial fill materials and the Production and Dispatch for the preparation and dispatch of ampouled materials.

References

[1] Frohlich, E. and Wahl, R. (2017) Thyroid autoimmunity: Role of anti-thyroid antibodies in thyroid and extra-thyroidal diseases. *Front. Immunol. 8: 521.*

[2] Sinclair, D (2006) Clinical and laboratory aspects of thyroid autoantibodies. *Ann Clin Biochem, 43:173-183*

[3] Citterio, C.E., Targovnik, H.M. and Arvan, P. (2019) The role of thyroglobulin in thyroid hormonogenesis. *Nat Rev Endocrinology, 15: 323-338.*

[4] Vargas-Uricoechea, H., Nogueira, J.P., Pinzon-Fernandez, M.V. and Schwarzstein, D. (2023) The usefulness of thyroid antibodies in the diagnostic approach to autoimmune thyroid disease. *Antibodies (Basel), 12(3):48.*

[5] Soh, S-B. and Aw, T-C. (2019) Laboratory testing in thyroid condition – pitfalls and clinical utility. *Annals of Laboratory Medicine, 39:3-14*.

[6] Spencer, C. and Fatemi, S. (2013) Thyroglobulin antibody (TgAb) methods – strengths, pitfalls and clinical utility for monitoring TgAb-positive patients with differentiated thyroid cancer. *Best Practice & Research Clinical Endocrinology & Metabolism, 27: 701-712.*

[7] Latrofa, F., Ricci, D., Grasso, L., Vitti, P., Masserini, L., Basolo, F., Ugolini, C., Mascia, G., Lucacchini, A. and Pinchera, A. (2008) Characterisation of thyroglobulin epitopes in patients with autoimmune and non-autoimmune thyroid diseases using recombinant human monoclonal thyroglobulin autoantibodies. *Endocrine Research, 93 (2):591-596.*

[8] WHO Tech Rep Ser No 800, (1990) 181-214.

[9] Finney, D. (1978) Statistical Method in *Biological Assay. Third ed. London: Charles Griffin*

[10] R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[11] Mair, P., Schoenbrodt, F. and Wilcox, R. (2017) WRS2: Wilcox robust estimation and testing.

[12] IFCC Working Group Recommendations for Assessing Commutability Part 3: Using the Calibration Effectiveness of a Reference Material (2018). *Clin Chem.* 64(3):465-474

[13] Kirkwood, T.B. (1977). Predicting the stability of biological standards and products. *Biometrics 33:736-42*

[14] Rosner, Bernard (May 1983), Percentage Points for a Generalized ESD Many-Outlier Procedure,*Technometrics*, 25(2), pp. 165-172.

[15] Pickett, A.J., Jones, M and Evans, C (2012) Causes of discordance between thyroglobulin antibody assays. *Ann Clin Biochem, 49: 463-467.*

[16] D'Aurizio, F., Metus, P., Ferrari, A., Caruso, B., Castello, R., Villalta, D., Steffan, A., Gaspardo, K., Pesente, F., Bizzaro, N., Tonutti, E., Valverde, S., Cosma, C., Plebani, M. and Tozzoli, R. (2017) Definition of the upper reference limit for thyroglobulin antibodies according to the National Academy of Clinical Biochemistry guidelines: a comparison of eleven different automated methods. *Autoimmun Highlights, 8:8*

APPENDIX 1 PARTICIPANT STUDY PROTOCOL

INTERNATIONAL COLLABORATIVE STUDY TO ESTABLISH THE 1ST WHO INTERNATIONAL STANDARD FOR ANTI-THYROGLOBULIN

INTRODUCTION

Autoimmune thyroid diseases are common disorders and include several inflammatory diseases with Graves' disease and Hashimoto's thyroiditis as the most frequent forms. Autoantibodies against thyroid antigens such as thyroglobulin (Tg) are often found in association with these autoimmune diseases and the detection of these autoantibodies is an important component in their diagnosis [1,2]. Thyroglobulin autoantibodies (TgAb) are specific to the thyroglobulin protein, a 660KDa glycoprotein which plays a major role in thyroid hormone synthesis, storage and release [3]. Measurement of TgAb by immunoassay is used to aid the diagnosis of autoimmune and other hypo/hyper thyroid diseases [4,5]. TgAb may also be measured in tandem with thyroglobulin in the diagnosis and monitoring of differentiated thyroid cancers due to potential interference with the measurement of thyroglobulin antigen in sera [6,7].

A WHO International Reference Reagent (IRR) for anti-thyroglobulin serum, coded 65/093, was established by the World Health Organisation's (WHO) Expert Committee on Biological Standardization (ECBS) in 1970 and has been widely adopted for the calibration of immunoassays for TgAb. Stocks of this IRR are now depleted, and as a result there is now an urgent requirement to prepare a replacement. It is anticipated that this replacement will become the 1st International Standard for anti-Thyroglobulin antibodies.

A new preparation of TgAb has been filled into ampoules (NIBSC code 23/180), following procedures recommended by WHO [8]. It is now intended to initiate an international collaborative study with expert laboratories to aid in the value assignment of the candidate 1st International Standard by immunoassay in units of IU/ampoule in comparison with the International Reference Reagent (IRR), 65/093. Human serum samples containing a range of TgAb concentrations will be included in the study in order to assess commutability of the candidate standard with native samples.

The aims of the collaborative study are therefore:

1) To confirm the immunoreactivity of the 1st IS by immunoassay, and to assess the relationship of this activity with the IRR, 65/093, and existing local standards.

2) To calibrate the candidate 1st IS for TgAb relative to the IRR, 65/093, by immunoassay.

3) To assess the commutability of the candidate 1st IS with native samples in immunoassays.

MATERIALS

Preparations supplied to participants in the collaborative study

The materials for the study are listed in Table 1. Each participant will be allocated a minimum set of samples consisting of the IRR, 65/093, duplicates of 23/180 and a panel of 9 human serum samples.

Table 1 – preparations for participants

TgAb preparation	Ampoule content
International Reference Reagent for TgAb, 65/093, coded CS731 sample C	1000 IU per ampoule
Coded preparations of the candidate 1st IS, 23/180, coded CS731 sample A and CS731 sample B	Nominally 1000 IU per ampoule
Nine human serum samples labelled CS731 TgAbSerum1 to CS731 TgAbSerum9	0.25 ml* aliquots human serum

*Aliquot volume may differ depending on assay requirements. Please refer to cover letter for further details.

International Reference Reagent, 65/093

A plasma pool from multiple donors with TgAb was clotted with the addition of calcium and the resulting serum diluted 1:23 with an isotonic diluent and filtered with an 0.4 µm filter. Aliquots of 1 ml were then dispensed into glass ampoules, lyophilised and sealed according to procedures recommended by WHO [8] and stored at -20ºC in the dark at NIBSC.

The ampoules were tested and found negative for anti-HIV 1 and 2, HBsAg and HCV.

Candidate standard, 23/180

Serum from three donors with high titre TgAb was obtained from Logical Biological (Sandwich, Kent, UK) and pooled with normal human serum (TCS Biosciences, Buckingham, UK) to provide a bulk solution of approximately 2000 IU/ml. Aliquots of 0.5 ml were then dispensed into glass ampoules, lyophilised and sealed according to procedures recommended by WHO [8] and stored at -20ºC in the dark at NIBSC.

The bulk plasma were tested and found negative for anti-HIV 1 and 2, HbsAg and HCV.

Human samples

Serum samples for inclusion in this study have been prepared from the individual serum donations that make up the bulk (Logical Biological, Kent, UK) or from serum obtained from Abbaltis Ltd (Sittingbourne, Kent, UK) diluted into normal human serum (TCS Biosciences, Buckingham, UK) at varying concentrations. Samples are coded TgAbSerum1 to TgAbSerum9. All samples have been tested and found negative for anti-HIV1 and 2, HBsAg and HCV.

A table of approximate TgAb content of each sample is provided in Appendix 1. Participants are requested to pre-dilute samples based on the approximate concentrations as necessary and in accordance with their in-house protocols. Please provide all details of pre-dilution steps where necessary.

This material is to be used only for this study and in accordance with the Human Tissue Act or equivalent national legislation and is to be destroyed at the end of the collaborative study.

Handling of materials

Upon receipt, ampoules should be stored at -20ºC or below until use. Allow the contents to reach room temperature before opening. Reconstitute with a volume of appropriate assay diluent (e.g. your own assay buffer, PBS or saline, preferably with 0.05 – 0.1% added protein such as bovine or human serum albumin to prevent adsorption). Leave at room temperature for a minimum of 10 minutes to fully dissolve. Dilutions should be prepared from this stock using your own assay diluent or PBS with protein cover as defined in common test sample concentrations below. A detailed protocol for reconstitution and dilution of the standards is provided in Appendix 1.

Please provide details of the reconstitution of the ampoules and the dilutions used to prepare the test samples.

Upon receipt, test samples of human serum should be stored at -20ºC or below until use. Allow contents to thaw and reach room temperature. Mix contents gently before measuring. Please do not re-freeze. Use a fresh aliquot for each run.

All material of human origin should be considered as potentially hazardous and handled with appropriate care. It should be used and discarded according to your own laboratory's safety procedures.

TESTS REQUESTED

Participants are requested to perform three independent runs of the assay method(s) normally used in your laboratory.

As ampoules of 65/093 are in very limited supply, participants are asked to perform two independent runs with a full sample panel that includes dilutions of 65/093, 23/180 and serum samples, using fresh ampoules/tubes for each assay run, and a third independent run with a reduced sample panel consisting of 23/180 and serum samples only. If the required assay/dead volumes allow, participants are requested to freeze a stock solution of 65/093 for use in the third independent run. Please indicate in your reported results any freeze thaw steps that have been used.

A sample panel will therefore consist of:

Runs 1+2 Full sample panel:

- One set of dilutions prepared from a fresh ampoule of 65/093
- One set of dilutions prepared from a fresh ampoule of 23/180-A
- One set of dilutions prepared from a fresh ampoule of 23/180-B
- One set of TgAb serum samples (n=9), thawed specifically for each run.

Run 3 Reduced sample panel (or frozen stock 65/093):

- One set of dilutions prepared from a fresh ampoule of 23/180-A
- One set of dilutions prepared from a fresh ampoule of 23/180-B
- One set of TgAb serum samples (n=9), thawed specifically for run 3.
- (one set of dilutions prepared from a frozen stock of 65/093, if available)

Assay kit calibrators and controls should also be included in each run where applicable. An independent run will use a single calibrated kit, integral or 96 well plate as required for your method.

Participants are asked to prepare a minimum of 5 dilutions of the ampouled preparations, including additional dilutions where needed to ensure that a minimum of 5 points in the linear part of the dose response curve are measured. An example core dilution series is provided in Appendix 1. The TgAb concentration of these dilutions and the serum samples should be measured in triplicate in each independent run.

Data submission

Participants are requested to provide **all raw assay data** in an electronic spreadsheet format for central computation at NIBSC, along with participants' own estimates of activity as calculated by the method normally used in their laboratory, and all details of the reconstitution and dilution volumes used to prepare the test samples. A suggested reporting table is shown in Appendix 1, Table A3.

REPORT

A preliminary report will be prepared and circulated to all participants for comment before submission to the Expert Committee on Biological Standardisation of WHO. In the report, participating laboratories will be identified by a laboratory number only and any requests to treat information in confidence will be respected.

REFERENCES

[1] Frohlich, E. and Wahl, R. (2017) Thyroid autoimmunity: Role of anti-thyroid antibodies in thyroid and extra-thyroidal diseases. *Front. Immunol. 8: 521.*

[2] Sinclair, D (2006) Clinical and laboratory aspects of thyroid autoantibodies. *Ann Clin Biochem, 43:173-183*

[3] Citterio, C.E., Targovnik, H.M. and Arvan, P. (2019) The role of thyroglobulin in thyroid hormonogenesis. *Nat Rev Endocrinology, 15: 323-338.*

[4] Vargas-Uricoechea, H., Nogueira, J.P., Pinzon-Fernandez, M.V. and Schwarzstein, D. (2023) The usefulness of thyroid antibodies in the diagnostic approach to autoimmune thyroid disease. *Antibodies (Basel), 12(3):48.*

[5] Soh, S-B. and Aw, T-C. (2019) Laboratory testing in thyroid condition – pitfalls and clinical utility. *Annals of Laboratory Medicine, 39:3-14.*

[6] Spencer, C. and Fatemi, S. (2013) Thyroglobulin antibody (TgAb) methods – strengths, pitfalls and clinical utility for monitoring TgAb-positive patients with differentiated thyroid cancer. *Best Practice & Research Clinical Endocrinology & Metabolism, 27: 701-712.*

[7] Latrofa, F., Ricci, D., Grasso, L., Vitti, P., Masserini, L., Basolo, F., Ugolini, C., Mascia, G., Lucacchini, A. and Pinchera, A. (2008) Characterisation of thyroglobulin epitopes in patients with autoimmune and non-autoimmune thyroid diseases using recombinant human monoclonal thyroglobulin autoantibodies. *Endocrine Research, 93 (2):591-596.*

[8] WHO Tech Rep Ser No 800, (1990) 181-214.

For further information, please contact:
Dr Melanie Moore
Principal Scientist, Therapeutic Reference Materials team, Science, Research and Innovation (S,R&I)
Medicines and Healthcare products Regulatory Agency
Tel: +44 (0) 1707 641242
Email: Melanie.moore@mhra.gov.uk

Appendix 1

Assay buffer

For dilution steps below, please use your appropriate assay buffer, ensuring protein cover is provided to prevent adsorption. PBS plus 0.1% bovine serum albumin (BSA) or 0.1% human serum albumin (HSA) may be used as an alternative.

Standard and sample processing

The following provides details on the reconstitution of the standards, along with example dilution steps to generate the working stock solution of the standards for use in the assay. These example dilution steps, or your own in-house dilution methods may be used, but please provide all details of reconstitution and dilution steps taken in your report.

A fresh ampoule of 65/093 (coded CS731 sample C) and duplicates of 23/180 (coded CS731 sample-A or CS731 sample-B) should be used for each full independent run, and fresh ampoules of duplicates of 23/180 (coded CS731 sample-A or CS731 sample-B) should be used in the third reduce panel. If the dead volume of your assay permits a third run from an ampoule of 65/093 after reconstitution and dilution, then please freeze a stock solution of 65/093 (1000 IU/ml) and thaw for use in the third run.

.

A. Reconstitution and example dilution series for ampoules of IRR 65/093 (coded CS731 sample C) and 23/180 (coded CS731 sample A and CS731 sample B)

1. Before opening, ampoules should be brought to room temperature to minimize moisture uptake.
2. Reconstitute each ampoule in 1 ml diluent to provide a working stock solution of 1000 IU/ml. Leave this solution to reconstitute fully (approximately 10 mins, with periodic gentle agitation).
3. The working stock solution at 1000 IU/ml will form dilution 1 and the solution from which serial dilutions should be made. For ampoules of 65/093, please freeze remaining Dilution 1 stock solution after preparation of the dilution curve for use in the final run.

- Prepare serial dilutions (1:2) of this working stock solution to provide dilutions 2 to 7. For example:
    - Dilution 2: 500 µl dilution 1 to 500 µl diluent
    - Dilution 3: 500 µl dilution 2 to 500 µl diluent, etc.
- Table A1 below provides an example core dilution series and expected concentrations.

Table A1 65/093 and 23/180 dilution table

Dilution	65/093 and 23/180 concentration (IU/ml)
1	**1000**
2	**500**
3	**250**
4	**125**
5	**62.5**
6	**31.25**
7	**15.625**

B. Preparation of serum and plasma samples TgAbSerum1 to TgAbSerum9

Upon arrival, please store all serum and plasma samples at -20ºC or below until use. Serum and plasma samples should be thawed at 37ºC and mixed well. Where required, serum and plasma samples should be diluted in your own assay buffer according to your standard assay protocol for measurement of TgAb. To aid with these dilutions Table A2 below provides the approximate TgAb content of each sample.

Table A2 Approximate serum/plasma sample anti-TPO content

Sample	Approx. TgAb content (IU/ml)
TgAbSerum1	350
TgAbSerum2	700
TgAbSerum3	190
TgAbSerum4	1180
TgAbSerum5	590
TgAbSerum6	150
TgAbSerum7	510
TgAbSerum8	140
TgAbSerum9	65

C. Assay design and plate layout

Alongside local standards and controls, each assay/each plate should include dilutions of 65/093 (for runs 1 and 2 as a minimum) and coded preparations of the candidate standard 23/180, CS731 sample A and B, plus 1 set of serum/plasma samples, TgAbSerum1-9. All samples should be tested in triplicate if possible, according to the in-house method. If space does not permit triplicate readings, then duplicate readings may be performed.

To enable us to gather data regarding inter and intra-assay variability within each laboratory, participants are requested to perform a minimum of 2 independent assays with the serum samples, candidate standard ampoules and IRR 65/093 ampoules provided. A third run with only candidate standard ampoules and serum samples is also requested if possible. Due to limited stocks, aliquots of 0.25 ml of each human serum/plasma sample are provided and should be diluted in the appropriate assay buffer as per your usual assay protocol.

D. Data reporting

Estimates of the TgAb content of the candidate standard 23/180, CS731 sample A and B, and the IRR 65/093, (coded CS731 sample C), should be calculated in comparison

with the in-house assay kit standard. Participants are requested to provide details of the assay method used, including dilution steps, together with all the raw data e.g. counts for each sample, in electronic for (excel file) if possible. Please also clarify if serum sample dilutions have been taken into account when calculating the results. A sample reporting table is provided in the attached excel workbook (Table A3).

APPENDIX 2

Figure A2.1. Slope ratios of 65/093 and coded duplicates of 23/180 (A and B)

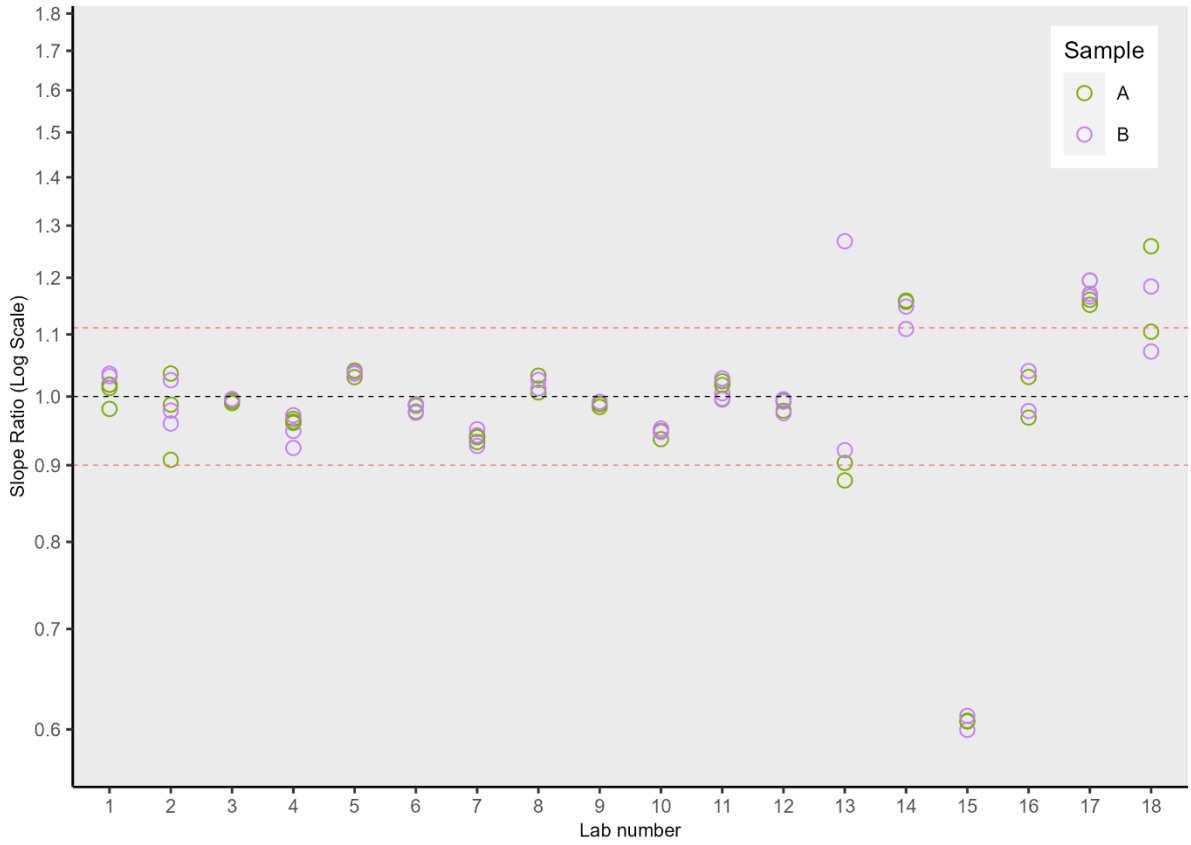


Table A2.1. Geometric mean reported estimates (IU/mL) for samples S1 to S9

Sample	Lab																		GM	GCV	Median	10^{MAD}
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18				
S1	677	539	646	412	1276	417	403	562	863	440	618	375	256	4383	232	517	782	1096	605	95%	550	1.35
S2	870	800	1524	611	2414	1063	564	1304	1452	1321	1379	854	1149	6483	782	1101	876	3187	1242	82%	1125	1.31
S3	267	232	439	183	1298	261	173	331	443	441	365	242	324	1775	249	283	1816	510	398	104%	328	1.35
S4	1510	709	1740	385	1556	886	357	1219	1409	706	1631	648	413	6980	349	1835	1075	1921	1031	115%	1144	1.61
S5	746	458	948	244	1250	521	220	637	802	523	886	403	263	4434	227	899	897	767	619	108%	689	1.35
S6	202	136	275	73	693	136	67	167	243	239	236	125	92	1236	82	220	411	212	196	113%	207	1.52
S7	763	728	934	514	1619	561	512	774	1172	845	1064	497	479	4774	383	782	1523	1408	865	82%	778	1.51
S8	185	169	263	136	762	141	133	184	316	245	259	136	131	1314	114	188	567	275	235	95%	187	1.40
S9	70	68	125	52	403	57	50	79	157	102	104	59	61	593	50	89	200	110	100	102%	84	1.45

Table A2.2. Geometric mean estimates (IU/mL) for samples S1 to S9 calculated relative to 65/093

Sample	Lab																		GM	GCV	Median	10^{MAD}
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18				
S1	640	424	432	326	451	387	467	576	922	435	657	374	161	5420	400	495	1717	6352	641	154%	459	1.24
S2	825	610	1063	497	1081	983	667	1373	1516	1339	1502	842	765	8259	879	1069	1281	20849	1295	149%	1066	1.29
S3	259	194	288	137	462	243	191	334	482	447	382	242	203	1960	418	268	4666	2343	416	159%	310	1.46
S4	1447	546	1222	304	602	819	411	1268	1472	726	1785	641	260	8979	519	1798	2505	11063	1100	178%	1000	1.79
S5	707	365	646	186	451	483	246	651	857	530	952	401	163	5378	392	870	2022	3854	651	154%	585	1.49
S6	194	119	176	51	197	127	70	168	268	237	243	126	55	1361	198	208	800	892	200	143%	196	1.46
S7	724	560	636	413	635	520	603	811	1244	863	1150	494	293	5805	549	755	3787	7254	915	144%	679	1.29
S8	178	145	168	100	224	131	144	189	349	247	268	137	83	1441	245	177	1172	1214	242	131%	184	1.34
S9	67	63	77	36	93	53	52	81	177	101	104	60	37	641	148	83	340	423	101	125%	82	1.45

Table A2.3. Geometric mean estimates (IU/mL) for samples S1 to S9 calculated relative to 23/180

Sample	Lab																		GM	GCV	Median	10^{MAD}
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18				
S1	513	573	417	703	403	479	681	508	578	349	379	485	447	789	414	373	288	606	483	30%	482	1.20
S2	659	830	1031	1094	938	1238	995	1173	957	1131	860	1105	2209	1148	1514	802	315	1714	1020	50%	1062	1.15
S3	203	258	277	284	413	299	263	300	303	350	222	313	565	332	444	202	676	288	316	38%	300	1.14
S4	1145	739	1186	653	525	1029	595	1096	929	579	1019	838	730	1232	636	1345	398	1048	829	41%	883	1.33
S5	566	490	625	392	393	601	344	575	538	420	548	521	453	798	401	653	332	429	491	27%	505	1.20
S6	153	157	169	101	180	154	90	152	169	182	142	161	147	231	129	157	150	122	150	24%	154	1.10
S7	579	759	615	902	553	647	893	698	777	702	660	643	825	856	697	567	566	775	698	17%	697	1.13
S8	141	192	161	205	204	160	194	168	218	187	156	176	225	245	185	134	208	158	182	18%	186	1.12
S9	53	81	73	69	88	64	65	73	111	73	62	76	99	111	81	63	73	64	75	22%	73	1.13

APPENDIX 3 DRAFT INSTRUCTIONS FOR USE

1st WHO International Standard for Thyroglobulin Antibodies, 23/180
(version 1, dated XX/XX/XXXX)

1. INTENDED USE

The 1st WHO International Standard for Thyroglobulin Antibodies, coded 23/180, is intended for use in the calibration of immunoassays for anti-thyroglobulin (TgAb). It replaces the International Reference Reagent, coded 65/093, for antithyroglobulin serum, produced in the 1970's as stocks are now exhausted. It should be noted that due to the inherent heterogeneity of autoantibodies, the replacement standard, 23/180, will not consist of identical autoantibody populations and it is therefore not a direct replacement of the material from which 65/093 was produced. Users are recommended to perform their own assessments to determine the impact of this new material in their own assays.

[The 1st IS was established by the Expert Committee on Biological Standardisation of the World Health Organisation in ___].

2. CAUTION

THIS PREPARATION IS NOT FOR ADMINISTRATION TO HUMANS OR ANIMALS IN TH HUMAN FOOD CHAIN:

The preparation contains material of human origin.

As with all materials of biological origin, this preparation should be regarded as potentially hazardous to health. It should be used and discarded according to your own laboratory's safety procedures. Such safety procedures probably will include the wearing of protective gloves and avoiding the generation of aerosols. Care should be exercised in opening ampoules or vials, to avoid cuts.

3. UNITAGE

Each ampoule of the International Standard contains **735 IU/ampoule** of anti-thyroglobulin antibodies**.**

4. CONTENTS

Country of original of biological material: United Kingdom

A batch of human serum containing high anti-thyroglobulin antibodies was produced by pooling serum from three donors containing high titre anti-thyroglobulin antibodies with normal human serum. Aliquots of 0.5 mL were then dispensed into glass ampoules, lyophilised and sealed according to procedures recommended by WHO.

Each ampoule therefore contains the residue after freeze-drying of 0.5 mL of human serum.

5. STORAGE

Unopened ampoules should be stored at -20ºC.
Please note: because if the inherent stability of lyophilized material, NIBSC may ship these materials at ambient temperature.

6. DIRECTIONS FOR OPENING

DIN ampoules have an "easy-open" coloured stress point, where the narrow ampoule stem joins the wider ampoule body. Tap the ampoule gently to collect the material at the bottom (labelled) end. Ensure that the disposable ampoule safety breaker provided is pushed down on the stem of the ampoule and against the shoulder of the ampoule body. Hold the body of the ampoule in one hand and the disposable ampoule breaker covering the ampoule stem between the thumb and first finger of the other hand. Apply a bending force to open the ampoule at the coloured stress point, primarily using the hand holding the plastic collar. Care should be taken to avoid cuts and projectile glass fragments that might enter the eyes, for example, by the use of suitable gloves and an eye shield. Take care that no material is lost from the ampoule and no glass falls into the ampoule. Within the ampoule is dry nitrogen gas at slightly less than atmospheric pressure. A new disposable ampoule breaker is provided with each DIN ampoule.

7. USE OF MATERIAL

No attempt should be made to weigh out any portion of the freeze-dried material prior to reconstitution. For all practical purposes each ampoule contains the same quantity of the substances listed above. Depending on the intended use, dissolve the total contents of the ampoule in a known volume of a suitable diluent. Users should make their own investigations into the type of diluent suitable for their use. If extensive dilutions are prepared, a carrier protein should be added. The ampoules do not contain bacteriostat and solutions of the material should not be assumed to be sterile.

8. STABILITY

Reference materials are stored at NIBSC within assured, temperature-controlled facilities. Reference materials should be stored on receipt as indicated on the label.

Data from an accelerated thermal degradation study performed as part of the collaborative study found that 23/180 is…… *when stored at -20ºC to serve as an International Standard.*

It is the policy of WHO not to assign an expiry date to their international reference materials. They remain valid with the assigned potency and status until withdrawn or amended.
Once reconstituted, diluted or aliquoted, users should determine the stability of the material according to their own method of preparation, storage and use. Users who have data supporting any deterioration in the characteristics of any reference preparation are encouraged to contact NIBSC.

9. CITATION

In any circumstance where the Recipient publishes a reference to NIBSC materials, it is important that the title of the preparation and any NIBSC code number, and the name and address of NIBSC are cited correctly.

10. LIABILITY AND LOSS

9.1 Unless expressly stated otherwise by NIBSC, NIBSC's Standard Terms and Conditions for the Supply of Materials (http://www.nibsc.org/terms_and_conditions.aspx) ("Conditions") apply to the exclusion of all other terms and are hereby incorporated into this document by reference.

9.2 Unless the context otherwise requires, the definitions in the Conditions shall apply.

9.3 Nothing in this document or the Conditions shall limit or exclude NIBSC's liability for fraud or fraudulent misrepresentation, death or personal injury caused by its negligence, or the negligence of its employees. Subject to clause 9.1:

9.3.1 NIBSC shall under no circumstances whatsoever be liable to the Recipient, whether in contract, tort (including negligence), breach of statutory duty, or otherwise, for any loss of data, loss of profit, loss of business or goodwill, or any indirect or consequential loss or damage suffered or incurred by the Recipient arising in relation to the supply of the Materials or the use, keeping, production or disposal of the Materials or any waste products arising from the use thereof by the Recipient or by any other person; and

9.3.2 NIBSC's total liability to the Recipient in respect of all other losses arising under or in connection with the Contract, whether in contract, tort (including negligence), breach of statutory duty, or otherwise, shall in no circumstances exceed 100% of the fees paid to NIBSC for the Materials.

9.4 The Recipient shall defend, indemnify and hold NIBSC, its officers, employees and agents harmless against any loss, claim, damage or liability including reasonable legal costs and fees (of whatsoever kind or nature) made against NIBSC which may arise as a result of the wilful act, omission or negligence of the Recipient or its employees, the breach of any of the terms of the Contract, or the use, keeping, production or disposal of the Materials or any waste products arising from the use thereof by the Recipient or on its behalf.

11. REFERENCES

[1] Moore, M., Hockley, J. Rigsby, P. and Cowper, B. (2024) WHO/BS/2024.XXXX: ECBS *report to be referenced/linked.*

12. MATERIAL SAFETY SHEET

Physical properties (at room temperature)			
Physical appearance	Yellowish powder		
Fire hazard	None		
Chemical properties			
Stable	Yes	Corrosive:	No
Hygroscopic	No	Oxidising:	No
Flammable	No	Irritant:	No
Other (specify)	Contains material of human origin, see caution, section 2		
Handling:	See caution, section 2		
Toxicological properties			
Effects of inhalation:	Not established, avoid inhalation		
Effects of ingestion:	Not established, avoid ingestion		
Effects of skin absorption:	Not established, avoid contact with skin		
Suggested First Aid			
Inhalation	Seek medical advice		
Ingestion	Seek medical advice		
Contact with eyes	Wash with copious amounts of water. Seek medical advice.		
Contact with skin	Wash thoroughly with water.		
Action on Spillage and Method of Disposal			
Spillage of ampoule contents should be taken up with absorbent material wetted with a virucidal agent. Rinse area with a virucidal agent followed by water. Absorbent materials used to treat spillage should be treated as biologically hazardous waste.			