Collaborative Study for the Validation of Alternative *in vitro* Potency Assays for Human Tetanus Immunoglobulin

S. Gross, S.W.J. Janssen, B. de Vries, E. Terao, A. Daas, K.-H. Buchheit

ABSTRACT

The European Pharmacopoeia (Ph. Eur.) monograph Human tetanus immunoglobulin (0398) gives a clear outline of the in vivo assay to be performed to determine the potency of human tetanus immunoglobulins during their development. Furthermore, it states that an in vitro method shall be validated for the batch potency estimation. Since no further guidance is given on the in vitro assay, every control laboratory concerned is free to design and validate an in-house method. At the moment there is no agreed in vitro method available.

The aim of this study was to validate and compare 2 alternative in vitro assays, i.e. an enzyme-linked immunoassay (EIA) and a toxoid inhibition assay (TIA), through an international collaborative study, in view of their eventual inclusion into the Ph. Eur.. The study was run in the framework of the Biological Standardisation Programme (BSP), under the aegis of the European Commission and the Council of Europe.

The collaborative study reported here involved 21 laboratories (public and industry) from 15 countries. Initially, 3 samples with low, medium and high potencies were tested by EIA and TIA. Results showed good reproducibility and repeatability of the 2 in vitro methods. The correlation of the data with the in vivo potency assigned by the manufacturers however appeared initially poor for high potency samples. Thorough re-examination of the data showed that the in vivo potencies assigned by the manufacturers had to be corrected: one for potency loss at the time of in vitro testing and one because of a reporting error. After these corrections the values obtained by in vivo and in vitro methods were in close agreement. A supplementary collaborative work was carried out to validate the 2 methods for immunoglobulin products with high potencies. Eight laboratories (public and industry) took part in this additional study to test 3 samples with medium and high potencies by EIA and TIA. Results confirmed that the 2 alternative methods are comparable in terms of assay repeatability, precision and reproducibility. In all laboratories, both methods discriminated between the low, medium and high potency samples.

Analysis of the data collected in this study showed a good correlation between EIA and TIA potency estimates as well as a close agreement between values obtained by in vitro and in vivo methods.

The study demonstrated that EIA and TIA are suitable quality control methods for polyclonal human tetanus immunoglobulin, which can be standardised in a quality control laboratory using a quality assurance system. Consequently, the Ph. Eur. Group of Experts 6B on Human Blood and Blood products decided in April 2009 to include both methods as examples in the Ph. Eur. monograph 0398 on Human Tetanus immunoglobulin.

KEYWORDS

Tetanus immunoglobulin, potency assay, alternative method, 3R, in vitro, EIA, TIA, batch control.

1. INTRODUCTION

Tetanus is a disease caused by Clostridium tetani, an organism that only flourishes in dead tissue and produces an exotoxin that passes into the central nervous system. Anti-tetanus immunoglobulin, produced by different manufacturers from human plasma, prevents tetanus intoxication by passive transfer of antibodies.

The potency of tetanus antibodies in these medicinal products is determined by the manufacturers according to the Ph. Eur. monograph *Human Tetanus Immunoglobulin* [1] by means of the toxin-neutralising assay in mice. According to this monograph, a satisfactory relationship shall be established between the potency determined by an immunoassay and that determined by the mouse assay.

Despite various efforts to develop an alternative *in vitro* assay [2, 3], no commonly accepted *in vitro* assay is yet available for the quality control of anti-tetanus immunoglobulin preparations. Consequently, every control laboratory concerned with the testing of potency of these products, e.g. the Official Medicines Control Laboratories

(OMCLs) and the manufacturers, is free to design and validate its own in-house potency assays. This might lead to discrepant results during Official Control Authority Batch Release of these products.

Although the Ph. Eur. monograph refers in its potency testing section to the general text on immunochemical assays and states that the potency is determined by comparing the antibody titre of the immunoglobulin to be examined with that of a reference preparation calibrated in International Units, using an immunoassay of suitable sensitivity and specificity, the mouse potency assay is still used by manufacturers, whereas OMCLs already use validated in vitro assay to resolve batch control issues.

Experience with *in vitro* potency assays is available in various laboratories and suggest that both EIA and TIA could be good alternative in vitro candidate methods to the mouse toxin neutralisation assay, provided they prove to be easily transferable, robust and accurate. The Paul-Ehrlich Institut (PEI, Germany) validated two different methods: an enzyme immunoassay (EIA, an ELISA-based assay originally

S. Gross. Paul-Ehrlich-Institut, 63225 Langen, Germany
S.W.J. Janssen. National Institute for Public Health and the Environment (RIVM), 3721 MA Bilthoven, The Netherlands
B. de Vries. National Institute for Public Health and the Environment (RIVM), 3721 MA Bilthoven, The Netherlands
E. Terao. (corresponding author: e-mail: eriko.terao@edqm.eu). European Directorate for the Quality of Medicines & HealthCare (EDQM), Council of Europe, Strasbourg, France
A. Daas. European Directorate for the Quality of Medicines & HealthCare (EDQM), Council of Europe, Strasbourg, France
K.-H. Buchheit. European Directorate for the Quality of Medicines & HealthCare, Council of Europe, Strasbourg, France

developed at the NIBSC and modified at PEI) and a tetanus toxoid inhibition assay (TIA, a modified Toxoid Binding Inhibition (ToBI) assay originally developed at the RIVM (The Netherlands)). Preliminary studies run at PEI have shown that both assays give equivalent results and showed both acceptable intra-laboratory precision and repeatability. Assay accuracy estimated by recovery of samples spiked with defined amounts of reference preparation was good [2]. Upon a proposal from the PEI and the RIVM based on this pre-validation, a collaborative study (coded BSP079) aiming to complete the validation of the 2 assays as potential alternative potency assays for batch release of Tetanus immunoglobulin products was launched in the framework of the Biological Standardisation Programme (BSP).

In order to avoid the use of animals by the laboratories participating in the collaborative study, it was decided not to estimate the accuracy of the alternative methods relative to the *in vivo* assay. This was deemed acceptable since each manufacturer has to perform a validation study for his own product, independently from the collaborative study.

The international collaborative study aimed at testing the precision and the reliability of the 2 methods with a panel of immunoglobulin samples covering the potency range of the approved products available on the European market (130 – 350 IU/ml).

A positive outcome of the collaborative study will allow the implementation of an animal-free method, which is in accordance with the European Convention on the protection of animals used for experimental and other scientific purposes [4] and with the Council Directive 86/609/EEC [5]. Furthermore, the establishment of a commonly accepted method would contribute to the standardisation of the potency determination of tetanus immunoglobulin preparations among manufacturers and OMCLs.

2. PARTICIPANTS

Twenty-one laboratories from 15 countries and the Council of Europe/EDQM laboratory participated in the first part of the collaborative study (Phase 2a: 1 Australia, 1 Austria, 1 Belgium, 1 Denmark, 2 France, 2 Germany, 1 Hungary, 2 Italy, 2 Netherlands, 1 Poland, 1 Portugal, 1 Spain, 1 Sweden, 1 Switzerland, 2 U.K.). Eight of these laboratories also participated in an additional part of the collaborative study (Phase 2b, see section 4.2.). Each laboratory is referred to in this report by an arbitrarily assigned number, not necessarily representing the order of listing in section 10. The coding of the laboratories was different in each sub-phase of this study, so that laboratories that participated in both parts are referred to by 2 codes in this report.

3. MATERIALS

All samples and key reagents (Tetanus immunoglobulin BRP, tetanus toxoid and coating antibody) provided by EDQM were shipped on ice. Participants were requested to store each product upon reception at 4°C as specified in the study protocol.

3.1. Tetanus immunoglobulin samples

Six commercially available tetanus immunoglobulin samples (labelled Sample A-F) with manufacturer stated in vivo potencies ranging from low to high were used for the study. The initially stated potency of the samples was: Sample A: 320 IU/ml; Sample B: 250 IU/ml; Sample C: 144 IU/ml, Sample D: 280 IU/ml, Sample E: 330 IU/ml, Sample F: 260 IU/ml. The potencies of Sample A and Sample B were

later corrected by the manufacturers to 285 IU/ml and 275 IU/ml respectively (see sections 6.1.4. and 7).

The starting concentrations of the dilution series were 1.2 IU/ml for EIA and 0.4 IU/ml for TIA.

Three vials of samples A, B and C were provided to the 21 laboratories participating in the first part (Phase 2a) of the collaborative study.

Two vials of samples D, E and F were shipped to the 8 laboratories that took part in the additional part of the collaborative study (Phase 2b).

3.2. Human tetanus immunoglobulin BRP

The Human tetanus immunoglobulin BRP batch 1 (EDQM Cat. Number: H1110000), used to calculate the potencies of the test samples, is a lyophilised preparation with a potency of 120 IU/vial, identical to the Anti-tetanus immunoglobulin, human, 1st International Standard (TE-3).

Each participant was provided with 3 vials of the Human tetanus immunoglobulin BRP. The content of each vial had to be reconstituted in 1 ml water and further diluted (1:100, to 1.2 IU/ml) in Phosphate Buffered Saline (PBS).

3.3. Tetanus toxoid

One vial of purified tetanus toxoid (chemically inactivated tetanus toxin, 730 Lf/ml, liquid, 1 ml/vial, donated by a manufacturer to EDQM) was provided to each participant. A pre-diluted solution at 7.3 Lf/ml in carbonate buffer was used to prepare the working solution of toxoid at 0.2 Lf/ml.

3.4. Coating antibody for TIA

One vial of a commercial human tetanus immunoglobulin, donated by a European manufacturer, was shipped to each participant performing TIA, to be used for the coating of the assay plates (270 IU/ml, liquid, 1 ml/vial).

3.5. Additional reagents (procured by the participants)

Peroxidase-coupled detection antibodies for EIA (eg. peroxidase-conjugated rabbit anti-human IgG (whole molecule) antibody, Sigma, Cat. No. A8792) and TIA (eg. peroxidase-conjugated donkey anti-mouse IgG(H+L) antibody, affinity pure F(ab)2 fragment, Jackson ImmunoResearch, Cat. No. 715-036-151) as well as the monoclonal anti-tetanus toxoid antibody for TIA (eg. from Antibody Shop, Cat. No. HYB 278-15-1; batch 251102, used at a dilution value around 1:5,000) as well as the peroxidase substrate (TMB solution for EIA and TIA, eg. DAKO, Cat. No. S159985) were procured by each participant.

3.6. Solutions

- Phosphate buffered saline (PBS without calcium and magnesium) : NaCl 137 mM, 8.1 mM disodium hydrogen phosphate, 2.7 mM potassium chloride, 1.47 mM potassium dihydrogen phosphate in water, pH 7.1 ± 0.1.

- PBS-T (washing buffer): PBS solution containing 0.05% polysorbate 20 (Tween 20).

- Blocking buffer: PBS containing 0.5% BSA.

- Dilution buffer: PBS-T containing 0.5% BSA.

- Carbonate buffer: 13.2 mM sodium carbonate, 35.7 mM sodium hydrogen carbonate, in water, pH 9.6 ± 0.1.

4. METHODS AND STUDY DESIGN

4.1. Methods

Two *in vitro* methods for the determination of the potency of tetanus immunoglobulin preparations were used: Enzyme immunoassay (EIA) and Toxoid binding inhibition assay (TIA). The potencies of the preparations were expressed in IU by comparison with the Human tetanus immunoglobulin BRP. Detailed standard operating procedures (SOPs), covering required equipment, reagents, chemicals and ELISA protocols, were elaborated in the laboratories of the Project Leaders for each of the methods and provided to the participants. These SOPs are available from EDQM on request. The following sections (4.1.1. and 4.1.2.) give a brief description of the procedures.

4.1.1. EIA

EIA is a direct method to determine the concentration of immunoglobulin bound to the tetanus toxoid, which is coated to a micro-titerplate. The bound immunoglobulin is detected by a peroxidase-conjugated polyclonal anti-human-IgG antibody, using TMB as chromogen. Participants were advised to coat the wells of a flat-bottomed microtitre plate (high protein-binding capacity, eg. Nunc Maxisorb, cat. No. 439454) with 100 µl of tetanus toxoid diluted to 0.2 Lf/ml in carbonate buffer, at 4°C for about 18 hours. The plates were washed 5 times in PBS-T (400 µl/well) and incubated in blocking buffer (200 µl/well) for 1h at 37°C. Tetanus immunoglobulin BRP had to be reconstituted according to instructions (see 3.2.). Two independent predilutions of 0.004 IU/ml in PBS were prepared for the BRP and each test sample. The potency assigned by the manufacturer was taken as a basis for the calculation of the dilution factor needed for the test samples. Each pre-diluted preparation was further diluted down with a factor of 1.5, resulting in dilution series of 6 dilutions in the range of 0.004 to 0.0005 IU/ml. After washing with PBS-T, 100 µl of sample or BRP were added and incubated for 2 hours at 37°C on a horizontal plate shaker set at 120 rpm. After washing in PBS-T, plates were incubated with 100 µl/well of a peroxidase-conjugated anti-Human IgG antibody diluted to a suitable concentration in dilution buffer, for 1 hour at 37°C on a plate shaker set at 120 rpm. The plates were washed with PBS-T before addition of 100 µl of TMB substrate to each well and incubated at room temperature for exactly 10 minutes in the dark. Detection reaction was stopped by addition of 100 µl of 2M sulfuric acid to the wells. Absorbances were immediately read at 450 nm, with a reference wavelength at 630 nm.

4.1.2. TIA

With TIA, the potency of specific tetanus immunoglobulin is indirectly estimated. TIA determines the concentration of unbound toxoid in a toxoid-antitoxoid mixture. Several dilutions of the test sample are incubated in a microplate with a defined amount of tetanus toxoid, leading to a dose-dependent toxoid binding. Subsequently, the amount of unbound toxoid is analyzed in a second microplate coated with polyclonal tetanus immunoglobulin, using an indirect detection system (a mouse monoclonal anti-toxoid antibody and a peroxidase-conjugated polyclonal anti-mouse antibody with TMB as chromogen).

The assay was run on 2 consecutive days. On the 1st day, pre-plates (round bottomed uncoated polystyrene microtiter plates, eg. Greiner PS/U-form/96 wells, Cat. No. 650001) were incubated with 200 µl/well of blocking buffer for 1 hour at 37°C, on a horizontal plate shaker set at 120 rpm. The plates were washed 5 times with PBS-T, taking care to remove the liquid completely to ensure good performance

of the assay. After the last washing step, plates were blotted against a clean paper towel to dry the wells.

The samples and the BRP were pre-diluted in PBS, to reach a concentration of 0.4 IU/ml. From these pre-diluted preparations, two independent series of 7 dilutions (0.2, 0.18, 0.16, 0.14, 0.12, 0.10 and 0.04 IU/ml) were further prepared in PBS. Each dilution was directly prepared from the 0.4 IU/ml solution. Hundred µl of each dilution were added to the pre-plate and incubated with 50 µl of tetanus toxoid (0.2 Lf/ml in carbonate buffer). Negative control wells were added with 150 µl of PBS and no toxoid solution. Positive control wells defining the maximum binding values were added with 100 µl PBS and 50 µl of toxoid. Plates were sealed and incubated overnight at 37°C on a horizontal plate shaker set at 120 rpm.

On the same first day, test-plates (high binding capacity flat-bottomed microtitre test-plates, eg. Nunc Maxisorp Cat. No. 439454) were coated with a human tetanus immunoglobulin diluted to 1 IU/ml in carbonate buffer (100 µl/well) and incubated for approximately 18 hours at 37°C on a shaker set at 120 rpm.

On the second day, the test-plates were washed 5 times with PBS-T and incubated in blocking buffer (250 µl/well) for 1 hour at 37°C on a shaker set at 120 rpm. After washing in PBS-T, 100 µl of the preparations from the pre-plates were transferred to the test-plates and incubated for 2 hours at 37°C on a plate shaker set at 120 rpm. The plates were further washed 5 times with PBS-T before addition of 100 µl/well of an anti-tetanus toxoid antibody diluted in PBS at the appropriate concentration. Plates were covered and incubated for 1 hour at 37°C on a shaker. After washing with PBS-T, the plates were incubated for 1 hour at 37°C on a shaker with 100 µl/well of peroxidase-conjugated secondary antibody diluted at the appropriate concentration in dilution buffer. The plates were washed in PBS-T before incubation with the chromogenic substrate (TMB, 100 µl/well) at room temperature for 10 minutes, in the dark. Detection reaction was stopped by addition of 100 µl/well of 2M sulfuric acid. The absorbances were immediately read at 450 nm, with a reference wavelength at 630 nm.

4.2. Study design

The study initially included a single collaborative study (Phase 2a), requiring participants to estimate the potency of 3 samples (samples A, B, C) using EIA and/or TIA. Three independent assays carried out at least 2 days apart, using a fresh vial of BRP and of each of the samples, had to be performed. For each assay, 2 independent dilutions series had to be prepared according to the provided SOPs.

After preliminary analysis of the participants' datasets, it was decided to launch an additional collaborative study (Phase 2b) to test 3 new samples (samples D, E, F). A subset of 8 participants was requested to perform duplicate assays with both EIA and TIA, using the provided study protocol and SOPs.

Participants were asked to report all results and modifications made to the SOPs to EDQM using the electronic datasheets provided.

5. STATISTICAL EVALUATION

All statistical analyses were performed at EDQM as detailed below, using the Human tetanus immunoglobulin BRP batch 1 to express the potencies in IU/ml.

5.1. EIA

The raw data were submitted to central calculations at the EDQM using the statistical software package CombiStats [6].

Initially, the parallel line model was intended to be used after log-transformation of the raw data:

$$\ln(y) = a_i + bx + \varepsilon$$

where y are the observed optical densities (OD), a_i are the intercepts of the preparations, b is the common slope, x is the ln(dose) and ϵ is an error term with constant variance and expectation 0. However, in some cases this did not lead to satisfactory linearity. It was noted that in most of these cases the ODs for low doses tended to level off at a small positive value and that linearity could be improved by subtracting this baseline value from the ODs before applying the log-transformation. Indeed, some laboratories had used similar approaches in their own calculations, for example by subtracting the average response for blanks. To avoid problems with negative values, the adopted approach for the central calculations was to fit an exponential model to the untransformed ODs:

$$y = d + e^{a_i + bx} + \varepsilon$$

where d denotes the lower asymptote and all other symbols are as before. To be consistent with the assumption of constant variance of the log-transformed ODs, the variance of ϵ must be assumed to be proportional to the squared response, corrected for d. In consequence an iteratively reweighted regression was carried out using weights inversely proportional to the estimated variance:

$$\mathrm{Var}\,(\varepsilon) \cong \left(e^{a_i + bx} \right)^2$$

5.2. TIA

A 4-parameter logistic curve model was used:

$$y = d + \frac{a - d}{1 + e^{-b(x - c_i)}}$$

where y are the ODs, d is the lower asymptote, a is the upper asymptote, b is the common slope-factor, c_i are the points of inflexion of the curves, x is the ln(dose) and ϵ is an error term with constant variance and expectation 0.

6. RESULTS

The feasibility phase of the project was performed by the PEI and the RIVM to define standard operating procedures and to identify critical reagents and parameters for the collaborative study. The *in vitro* potencies of most of the 7 commercially available human immunoglobulin samples tested were consistent with the manufacturers' values. The study also confirmed the equivalence between the Human tetanus immunoglobulin biological reference preparation (BRP, batch 1) and the anti-tetanus immunoglobulin, human, international standard (IS, TE-3). Furthermore, similar results were obtained with 2 different coating antibodies for TIA.

6.1. Phase 2a (samples A, B and C)

6.1.1. EIA

20 laboratories carried out the method of EIA. Laboratory 1 submitted results from 5 assays. Laboratory 18 submitted results from 4 assays, but reported only calculated potencies because the format of their raw data did not conform to the requested assay design. All other laboratories carried out 3 assays as requested per protocol. Laboratory 12 carried out 3 assays using the provided SOP and 3 assays according to their in-house SOP. The 2 sets of results are coded as Lab 12a and 12b respectively. Laboratories 12 and 19 did not provide potency estimates from own calculations.

The statistical model described earlier (see Section 5) produced in most cases satisfactory parallelism and linearity although it was in many cases necessary to exclude one or two of the lowest or highest dose-levels. For laboratory 10 it was necessary to exclude three dose levels. A summary of results is given in Table 1. Shown are the P-values for deviations from parallelism and linearity together with the weighted correlation coefficient. Significant P-values (<0.05) are printed on a grey background and highly significant P-values (<0.01) are printed on a black background. Correlation coefficients less than 0.99 are printed on a grey background and less than 0.98 are printed on a black background. Table 1 also shows the estimated potencies as calculated by the participants and those as calculated at the EDQM with associated 95% confidence limits. Confidence intervals wider than 90 to 111 per cent are printed on a grey background and intervals wider than 80 to 125 per cent are printed on a black background.

Table 1 – *Overview of the EIA results (samples A, B and C; potency estimates in IU/ml)*

Lab	Assay	Par	Lin	Corr	Own (A)	EDQM (A)	(Conf. Lim.) (A)	Own (B)	EDQM (B)	(Conf. Lim.) (B)	Own (C)	EDQM (C)	(Conf. Lim.) (C)
1	1	0.980	0.963	0.995	244	222	(94% - 107%)	263	258	(94% - 107%)	148	133	(94% - 107%)
	2	0.890	0.991	0.996	258	226	(94% - 106%)	260	260	(94% - 106%)	134	130	(94% - 106%)
	3	0.922	0.956	0.997	234	258	(95% - 105%)	294	272	(95% - 105%)	131	136	(95% - 105%)
	4	0.275	0.214	0.994	221	254	(94% - 106%)	222	230	(94% - 106%)	140	141	(94% - 106%)
	5	0.870	0.430	0.999	311	290	(98% - 102%)	336	304	(98% - 102%)	156		
2	1	0.841	0.982	0.996	304	301	(95% - 105%)	279	294	(95% - 105%)	164	173	(95% - 105%)
	2	0.699	0.991	0.996	281	284	(95% - 106%)	274	296	(95% - 106%)	146	153	(95% - 106%)
	3	0.524	0.994	0.996	290	288	(94% - 106%)	304			166	176	(94% - 106%)
3	1	0.054	0.999	0.996	278	278	(94% - 106%)	283	284	(94% - 106%)	154	153	(94% - 106%)
	2	0.346	0.446	0.997	263	264	(95% - 105%)	292	290	(95% - 105%)	146	146	(95% - 105%)
	3	0.700	0.786	0.997	269	273	(96% - 105%)	265	264	(96% - 105%)	142	142	(96% - 105%)
4	1	0.344	0.533	0.998	271	267	(97% - 103%)	237			135	134	(97% - 103%)
	2	0.000	0.000	0.994	284	288	(97% - 103%)	225	225	(97% - 103%)	153	159	(97% - 103%)
	3	0.000	0.000	0.986	274	269	(96% - 104%)	231	234	(96% - 104%)	135	132	(96% - 104%)
5	1	0.563	0.917	0.993	263	277	(93% - 108%)	253	270	(93% - 108%)	126	135	(93% - 108%)
	2	0.846	1.000	0.977	262	265	(86% - 116%)	330	334	(86% - 116%)	152	156	(86% - 116%)
	2	0.717	0.983	0.993	287			321	324	(93% - 108%)	151	149	(93% - 108%)
6	1	0.073	0.396	0.997	207	242	(96% - 104%)	223	234	(96% - 104%)	111	125	(96% - 104%)
	2	0.443	0.129	0.994	222	249	(95% - 105%)	330	278	(95% - 105%)	135	138	(95% - 105%)
	3	0.167	0.778	0.995	240	265	(95% - 106%)	242	243	(95% - 106%)	102	119	(95% - 106%)
7	1	0.413	1.000	0.998	315	286	(96% - 105%)	272	272	(96% - 105%)	161	157	(96% - 105%)
	2	0.944	0.990	0.993	335	303	(92% - 109%)	320	323	(92% - 109%)	154	153	(92% - 109%)
	3	0.858	1.000	0.992	337	272	(91% - 110%)	298	298	(91% - 110%)	163	162	(91% - 110%)
8	1	0.000	0.050	0.942	212	171	(92% - 109%)	207	226	(92% - 109%)	73	98	(92% - 108%)
	2	0.000	0.193	0.994	267	206	(97% - 103%)	225	231	(97% - 103%)	132	97	(97% - 103%)
	3	0.003	0.422	0.988	231	214	(94% - 107%)	273	254	(94% - 107%)	146	126	(94% - 107%)
9	1	0.167	0.585	0.973	328	337	(87% - 115%)	279	289	(87% - 115%)	176	178	(87% - 115%)
	2	0.145	0.904	0.991	327	326	(92% - 109%)	298	298	(92% - 109%)	182	182	(92% - 109%)
	3	0.795	0.774	0.975	284	275	(86% - 116%)	295	292	(86% - 116%)	179	172	(86% - 116%)
10	1	0.085	0.597	0.945	318	286	(85% - 117%)	284	270	(86% - 117%)	137	113	(84% - 117%)
	2	0.481	0.334	0.979	268	215	(89% - 111%)	252	232	(90% - 111%)	173	161	(90% - 111%)
	3	0.864	0.058	0.969	340	308	(90% - 111%)	302	316	(90% - 112%)	178	168	(90% - 112%)
11	1	0.025	0.056	0.996	273	270	(96% - 104%)	287	290	(96% - 104%)	156	157	(96% - 104%)
	2	0.444	0.749	0.993	260	257	(93% - 107%)	269	273	(93% - 107%)	155	156	(93% - 107%)
	3	0.218	0.703	0.997	227	221	(96% - 105%)	249	251	(96% - 105%)	145	146	(96% - 105%)
12a	1	0.638	0.550	0.954		285	(84% - 119%)		269	(84% - 120%)		137	(84% - 119%)
	2	0.468	0.552	0.975		272	(87% - 114%)		268	(88% - 114%)		187	(88% - 114%)
	3	0.619	0.973	0.964		301	(84% - 119%)		247	(84% - 119%)		166	(84% - 119%)
12b	1	0.227	0.676	0.997		262	(95% - 105%)		269	(95% - 105%)		140	(95% - 105%)
	2	0.754	0.119	0.998		290	(97% - 103%)		263	(97% - 103%)		151	(97% - 103%)
	3	0.866	0.293	0.996		292	(95% - 105%)		252	(95% - 105%)		149	(95% - 105%)
13	1	0.984	0.991	0.930	300	293	(75% - 133%)	333	251	(75% - 134%)	128	138	(75% - 133%)
	2	0.057	0.564	0.994	263	283	(94% - 106%)	218	221	(94% - 106%)	135	135	(95% - 106%)
	3	0.964	1.000	0.961	398			241	234	(82% - 123%)	144	147	(82% - 123%)
14	1	0.210	0.997	0.965	221	194	(82% - 121%)	251	234	(83% - 121%)	161	150	(83% - 121%)
	2	0.000	0.001	0.998	171	169	(97% - 103%)	219	213	(97% - 103%)	120	118	(97% - 103%)
	3	0.000	0.000	0.993	167	159	(97% - 103%)	232	227	(97% - 103%)	142	140	(97% - 103%)
15	1	0.076	0.750	0.995	248	245	(94% - 106%)	276	280	(94% - 106%)	141	142	(94% - 106%)
	2	0.080	0.914	0.995	285	282	(94% - 106%)	290	289	(94% - 106%)	164	162	(94% - 106%)
	3	0.066	0.034	0.994	251	242	(95% - 105%)	276	267	(95% - 105%)	163	153	(95% - 105%)
16	1	0.618	0.039	0.996	253	253	(95% - 105%)	261	262	(95% - 105%)	147	145	(95% - 105%)
	2	0.871	1.000	0.982	264	255	(87% - 115%)	214	200	(87% - 115%)	113	124	(86% - 116%)
	3	0.273	0.995	0.994	277	280	(94% - 106%)	241	251	(94% - 106%)	140	146	(94% - 106%)
17	1	0.772	0.616	0.994	280	280	(93% - 107%)	251	251	(93% - 107%)	167	167	(93% - 107%)
	2	0.193	0.215	0.991	267	259	(93% - 107%)	273	264	(93% - 107%)	153	150	(93% - 107%)
	3	0.261	0.937	0.996	265	262	(95% - 105%)	263	262	(95% - 105%)	154	155	(95% - 106%)
18	1				328			330			138		
	2				292			357			140		
	3				253			267			132		
	4				248			284			130		
19	1	0.051	1.000	0.960		336	(82% - 122%)		268	(82% - 122%)		227	(82% - 123%)
	2	0.984	0.957	0.986		296	(89% - 113%)		332	(89% - 113%)		159	(89% - 113%)
	3	0.967	0.741	0.975		295	(86% - 116%)		272	(86% - 116%)		143	(86% - 116%)
21	1	0.221	0.994	0.983	277	268	(89% - 112%)	239	233	(89% - 112%)	138	135	(89% - 112%)
	2	0.671	1.000	0.951	277	247	(78% - 128%)	255	230	(78% - 128%)	184	180	(78% - 129%)
	3	0.253	0.997	0.989	247	244	(91% - 110%)	217	210	(91% - 110%)	135	134	(91% - 110%)

Explanations:

Par=p-value for deviations from parallelism. Lin=p-value for deviations from linearity. Corr=weighted coefficient of correlation.

Assays with significant deviations from parallelism and/or linearity are printed on a black (p<0.01) or grey (p<0.05) background.

Correlation coefficients below 0.98 are printed on a black background and below 0.99 are printed on a grey background.

Own=Potency in IU/ml as calculated by participants. EDQM=Potency as calculated at EDQM with 95% confidence limits between brackets.

Confidence limits wider than 90-111% are printed on a grey background and wider than 80-125% on a black background.

Most assays showed satisfactory linearity and parallelism. Significant deviations were observed for 8 assays in a total of 4 laboratories. In the case of laboratories 4, 11 and 14 these deviations were probably not relevant because the residual error was rather low compared to the overall error. Indeed, graphical inspection of the plots did not give any reason to suspect fundamental problems with validity conditions. In the case of laboratory 8 there were reasons to believe that problems with assay consistency were likely to be the cause of the observed non-parallelism rather than true lack of similarity of the dose response curves. For example, the slope of Sample C is 21 per cent shallower than the BRP in assay 1 whereas it is 24 per cent steeper than the BRP in assay 2.

The majority of the assays had a correlation coefficient higher than 0.99 and confidence intervals within 90 to 111 per cent. Only in 2 assays did the confidence limits exceed a width of 80 to 125 per cent, demonstrating a satisfactory overall precision of the method.

The unweighted geometric means (GM) and geometric coefficients of variation (GCV) of the potency estimates per laboratory as calculated at the EDQM are displayed in Table 2. Since no central calculations were possible for

laboratory 18 the calculated results as reported are listed. The median intra-laboratory RSD is between 6 and 8 per cent which can be considered a satisfactory method repeatability.

Also shown in Table 2 are the overall GM, GCV and median. This is shown for the case that all potency estimates are included, for the case that assays with significant deviations from linearity and/or parallelism are excluded, and for the case that assays with a correlation coefficient below 0.99 are excluded. It can be seen that inclusion or exclusion of lower quality assays did not affect the overall outcome much. The overall mean potencies are about 265 IU/ml, 265 IU/ml and 150 IU/ml for the respective samples. The inter-laboratory GCV of the laboratory means was between 10 and 15 per cent which can be considered satisfactory method reproducibility.

6.1.2. TIA

11 laboratories carried out the method by TIA (Lab 1, 2, 3, 5, 8, 10, 11, 16, 19, 20 and 21). Laboratory 1 submitted results from 5 assays. All other laboratories carried out 3 assays as requested per protocol. Laboratories 19 and 20 did not provide potency estimates from own calculations.

Table 2 − Geometric means and coefficients of variation per laboratory for EIA (samples A, B and C; potency estimates in IU/ml)

Lab	Including all results						Excluding assays with significant non-linearity and/or non-parallelism						Excluding assays with correlation coefficient less than 0.99					
	Sample A		Sample B		Sample C		Sample A		Sample B		Sample C		Sample A		Sample B		Sample C	
	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV
1	249	11	263	11	135	4	249	11	263	11	135	4	249	11	263	11	135	4
2	291	3	295	1	167	8	291	3	295	1	167	8	291	3	295	1	167	8
3	271	3	279	5	147	4	271	3	279	5	147	4	271	3	279	5	147	4
4	275	4	229	3	141	11	267				134		278	5	225		146	13
5	271	3	308	12	146	8	271	3	308	12	146	8	277		296	14	142	7
6	252	5	251	10	127	8	252	5	251	10	127	8	252	5	251	10	127	8
7	287	6	297	9	157	3	287	6	297	9	157	3	287	6	297	9	157	3
8	196	13	236	7	106	16							206		231		97	
9	311	12	293	2	177	3	311	12	293	2	177	3	326		298		182	
10	267	21	270	17	145	24	267	21	270	17	145	24						
11	249	11	271	8	153	4	239	11	261	6	151	5	249	11	271	8	153	4
12a	286	5	261	5	162	17	286	5	261	5	162	17						
12b*	281	6	261	3	147	4	281	6	261	3	147	4	281	6	261	3	147	4
13*	288	2	235	7	140	4	288	2	235	7	140	4	283		221		135	
14	173	10	225	5	135	13	194		234		150		164	4	220	5	128	13
15	256	9	279	4	152	7	263	10	285	2	152	10	256	9	279	4	152	7
16	263	6	236	16	138	10	268	7	224	18	134	12	266	8	257	3	145	0
17	267	4	259	3	157	6	267	4	259	3	157	6	267	4	259	3	157	6
18*	261	1	263	1	152	2	261	1	263	1	152	2	261	1	263	1	152	2
19	308	8	289	13	173	27	308	8	289	13	173	27	308	8	289	13	173	27
21	253	5	224	6	148	18	253	5	224	6	148	18	253	5	224	6	148	18
GM	262		262		147		267		265		150		262		261		146	
GCV	14		10		12		11		10		9		16		11		14	
Median	267		263		147		268		263		149		267		263		147	
GM*	260		264		147		266		267		150		260		263		146	
GCV*	15		11		13		11		10		10		17		11		16	
Median*	267		267		147		267		267		150		267		267		147	

* Excluding results from laboratories 12b, 13 and 18.

Explanations: GM = Geometric mean. GCV = Geometric Coefficient of Variation.

The raw data were submitted to central calculations at the EDQM, using the 4-parameter logistic curve model detailed in section 5. The assumption that the error term ϵ is constant over the entire range of responses appears to be violated in most of the assays. Indeed, the variance is in general more important for higher ODs at lower doses than for lower ODs at higher doses. A weighted regression where the weights are taken to be inversely proportional to the expected response would therefore seem to be more appropriate. However, in many cases this resulted in unstable convergence paths or convergence could not be reached at all. It was therefore decided to apply an unweighted regression instead, in accordance with what most laboratories had used. This should not influence the potency estimate much, but it should be borne in mind that this may bias the estimation of confidence limits and P-values for nonlinearity and non-parallelism.

A summary of results is given in Table 3. Most assays showed satisfactory linearity and parallelism. Visual inspections of plots where significant deviations from linearity and/or parallelism are indicated revealed no major problems with the data. This was mainly thanks to the rather steep regression over small dose intervals near the point of inflexion. Under such circumstances, slight but significant deviations from parallelism and/or linearity can occur without compromising the potency estimate. Table 3 reveals that the large majority of the assays had correlation coefficients below 0.98. However, the very steep regression of the responses largely compensates for this fact, resulting in overall good precision. Since, in addition, the quality of the data upon visual inspection appears to be satisfactory, it must be concluded that the correlation coefficient is not a suitable quality marker for this assay method.

Table 3 − *Overview of the TIA results (samples A, B and C; potency estimates in IU/ml)*

Lab	Assay	Goodness of fit			Sample A			Sample B			Sample C		
		Par	Lin	Corr	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)
1	1	0.264	0.902	0.938	292	293	(96% - 104%)	305	305	(97% - 104%)	149	149	(97% - 104%)
	2	0.007	0.960	0.964	274	275	(98% - 102%)	276	277	(98% - 102%)	147	147	(98% - 102%)
	3	0.106	0.997	0.877	355	356	(94% - 107%)	305	307	(94% - 107%)	145	145	(94% - 106%)
	4	0.770	1.000	0.922	329	330	(96% - 104%)	367	369	(96% - 105%)	184	185	(96% - 104%)
	5	0.681	1.000	0.889	283	283	(94% - 106%)	287	287	(94% - 107%)	170	170	(94% - 107%)
2	1	0.627	0.683	0.945	221	250	(97% - 104%)	279	267	(96% - 104%)	143	146	(96% - 104%)
	2	0.167	1.000	0.798	228	260	(93% - 108%)	283	266	(93% - 108%)	162	155	(93% - 108%)
	3	0.951	0.002	0.913	297	291	(97% - 103%)	342	296	(97% - 103%)	180	160	(97% - 103%)
3	1	0.000	0.069	0.972	258	257	(98% - 102%)	253	253	(98% - 102%)	145	145	(98% - 102%)
	2	0.004	0.857	0.967	263	262	(97% - 103%)	283	284	(97% - 103%)	160	160	(97% - 103%)
	3	0.533	0.000	0.970	264	264	(98% - 102%)	293	295	(98% - 102%)	157	157	(98% - 102%)
5	1	0.082	1.000	0.910	307	307	(95% - 104%)	215	315	(94% - 107%)	170	169	(94% - 106%)
	2	0.932	0.998	0.946	283	283	(96% - 104%)	287	287	(96% - 104%)	149	149	(96% - 105%)
	3	0.481	0.172	0.965	291	291	(97% - 103%)	311	311	(97% - 103%)	172	172	(97% - 103%)
8	1	0.442	0.999	0.885	259	260	(95% - 105%)	252	259	(95% - 106%)			
	2	0.344	0.000	0.949	275	273	(98% - 102%)	295	295	(98% - 102%)	135	154	(98% - 102%)
	3	0.000	0.000	0.963	293	293	(99% - 101%)	270	270	(99% - 102%)	150	149	(99% - 101%)
10	1	0.933	1.000	0.964	266	260	(97% - 103%)	250	278	(95% - 105%)	147	158	(96% - 104%)
	2	0.460	0.107	0.898	255	258	(96% - 104%)	282	289	(94% - 107%)	168	164	(95% - 106%)
	3	0.259	0.891	0.975	299	335	(98% - 102%)	305	342	(97% - 104%)	169	178	(98% - 102%)
11	1	0.040	0.301	0.972	285	285	(97% - 103%)	290	292	(97% - 103%)	159	159	(97% - 103%)
	2	0.223	0.008	0.974	311	310	(98% - 102%)	342	340	(97% - 103%)	167	167	(98% - 102%)
	3	0.053	0.320	0.964	257	258	(97% - 103%)	273	274	(97% - 104%)	146	146	(97% - 103%)
16	1	0.887	0.992	0.949	300	300	(96% - 104%)	314	314	(95% - 106%)	189	183	(95% - 106%)
	2	0.325	1.000	0.965	306	306	(97% - 104%)	324	323	(96% - 104%)	175	175	(96% - 104%)
	3	0.657	0.975	0.973	303	303	(97% - 103%)	310	310	(97% - 103%)	180	180	(97% - 103%)
19	1	0.376	1.000	0.824		685	(90% - 113%)		701	(88% - 116%)		326	(90% - 113%)
	2	0.000	0.780	0.899		783	(96% - 104%)		596	(96% - 104%)		319	(96% - 104%)
	3	0.087	0.847	0.798		406	(92% - 109%)		432	(90% - 113%)		184	(92% - 109%)
20	1	0.796	0.134	0.895		283	(96% - 104%)		288	(96% - 104%)		159	(96% - 104%)
	2	0.000	0.000	0.987		271	(99% - 101%)		291	(99% - 101%)		156	(99% - 101%)
	3	0.001	0.031	0.983		272	(98% - 102%)		282	(98% - 102%)		157	(98% - 102%)
21	1	0.711	1.000	0.967	259	258	(94% - 106%)	235	234	(94% - 106%)	159	163	(94% - 106%)
	2	0.288	0.809	0.961	233	240	(96% - 104%)	231	229	(96% - 104%)	133	135	(96% - 104%)
	3	0.209	0.604	0.938	211	210	(95% - 105%)	222	220	(95% - 105%)	140	139	(95% - 106%)

Explanations:

Par=p-value for deviations from parallelism. Lin=p-value for deviations from linearity. Corr=weighted coefficient of correlation.

Assays with significant deviations from parallelism and/or linearity are printed on a black (p<0.01) or grey (p<0.05) background.

Correlation coefficients below 0.98 are printed on a black background and below 0.99 are printed on a grey background.

Own=Potency in IU/ml as calculated by participants. EDQM=Potency as calculated at EDQM with 95% confidence limits between brackets.

Confidence limits wider than 90-111% are printed on a grey background and wider than 80-125% on a black background.

Table 3 also shows that the confidence limits of the potency estimate are in all cases within 80 to 125 per cent and in only 2 assays they exceed 90 to 111 per cent. Again, this is mostly thanks to the very steep regression of the curves but it should also be kept in mind that the confidence limits may be somewhat optimistic due to the unweighted fit of the model. It is therefore somewhat doubtful to draw the conclusion that the precision of the method would be better than EIA based on the smaller confidence limits alone. However, as the repeatability of both methods is comparable (see next paragraph) and the precision is by definition better than the repeatability, it is justified to state that TIA is at least as precise as EIA. Laboratory 19 clearly had a problem with this assay: not only are the confidence limits wider than in any other laboratory, also the potency estimates were very poorly repeatable and much higher than in other laboratories. The reason is not fully clear but it would seem that the BRP in this laboratory had considerably lost potency. Further calculations are performed excluding the results from laboratory 19.

The unweighted GMs and GCVs of the potency estimates are shown in Table 4. The median intra-laboratory GCV is between 6 and 8 per cent which is comparable to that of EIA and can be considered satisfactory method repeatability.

Also shown in Table 4 are the overall GM, GCV and median. This is shown for the case that all potency estimates are included and for the case that assays with significant deviations from linearity and/or parallelism are excluded. Again, it can be seen that inclusion or exclusion of lower quality assays does not affect the overall outcome much. The overall mean potencies are about 275 IU/ml, 285 IU/ml and 160 IU/ml for the respective samples. The interlaboratory GCV of the laboratory means is between 6 and 10 per cent which can be considered a satisfactory method reproducibility.

6.1.3. Comparisons between EIA and TIA (samples A, B and C)

A graphical impression of the distribution of potency estimates is provided in Figure 1. Shown are histograms for all samples and methods. Visually both methods seem to give comparable results, but Student's unpaired *t*-test gives *p*-values of 0.23, 0.02 and 0.06 for the 3 respective samples which would be a slight indication that the methods do not always yield equivalent results. When the *t*-test is applied to

estimates calculated by the participants, the *p*-values are 0.71, 0.15 and 0.05 respectively. Although the indication is weak there seems to be a tendency for TIA to give somewhat higher potency estimates than EIA in this part of the study. The importance may not be relevant though.

6.1.4. Comparisons with assigned potency from the manufacturers (samples A, B and C)

Table 5 shows a summary of the results compared with the assigned potencies from the manufacturers, which were obtained using *in vivo* methods. The initially reported assigned potency of Sample B was 250 IU/ml. It was later discovered that a mistake occurred and that the correct value was 275 IU/ml. The results for Samples B (corrected value) and C are quite close to the assigned potency with a difference of less than 8 per cent. Given the fact that the assigned potency is based on 1 assay only and the observed inter-laboratory GCV for EIA and TIA is about 10 per cent, this difference can be considered non-significant. For Sample A the difference was found to be 19 per cent. Although this difference could still be explained as normal statistical variation, it was found desirable to confirm the assigned potency by repeating the *in vivo* test because it was thought possible that the batch might have lost potency in the time between the initial test and the current study. The repeat *in vivo* assay was carried out by the manufacturer under the same conditions as the initial assay and confirmed a loss of potency, the new result being 285 IU/ml. This brings the difference with the *in vitro* methods down to less than 9 per cent.

Table 5 - *Summary table of results in comparison with assigned potencies (IU/ml, samples A, B and C)*

Sample	Assigned potency (manufacturer's *in vivo* method)	Potency estimated with EIA	Potency estimated with TIA
A	285 (initially 320)	260	277
B	275 (initially 250)	264	286
C	144	147	158

However, sample A was included in this study for its supposed high potency so as to cover a wide range of

Table 4 − *Geometric means and coefficients of variation per laboratory for TIA (samples A, B and C; potency estimates in IU/ml)*

Lab	Including all results						Excluding assays with significant non-linearity and/or non-parallelism					
	Sample A		Sample B		Sample C		Sample A		Sample B		Sample C	
	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV
1	306	12	307	12	159	11	314	11	316	12	161	12
2	266	8	276	6	153	5	255	3	267	0	150	4
3	261	1	277	8	154	5						
5	293	4	304	5	163	8	293	4	304	5	163	8
8	275	6	274	7	151	3	260		259			
10	282	16	302	12	166	6	282	16	302	12	166	6
11	284	10	301	12	157	7	258		274		146	
16	303	1	316	2	179	2	303	1	316	2	179	2
19	601	42	565	28	267	38	527	45	550	41	245	50
20	275	2	287	2	157	1	283		288		159	
21	235	11	227	3	145	11	235	11	227	3	145	11
GM	277		286		158		275		282		158	
GCV	8		10		6		10		11		7	
Median	279		294		157		282		288		160	

Explanations: GM = Geometric mean. GCV = Geometric Coefficient of Variation.

Figure 1 - *Histograms of estimated potencies by EIA and TIA (samples A, B and C; potency estimates in IU/ml)*
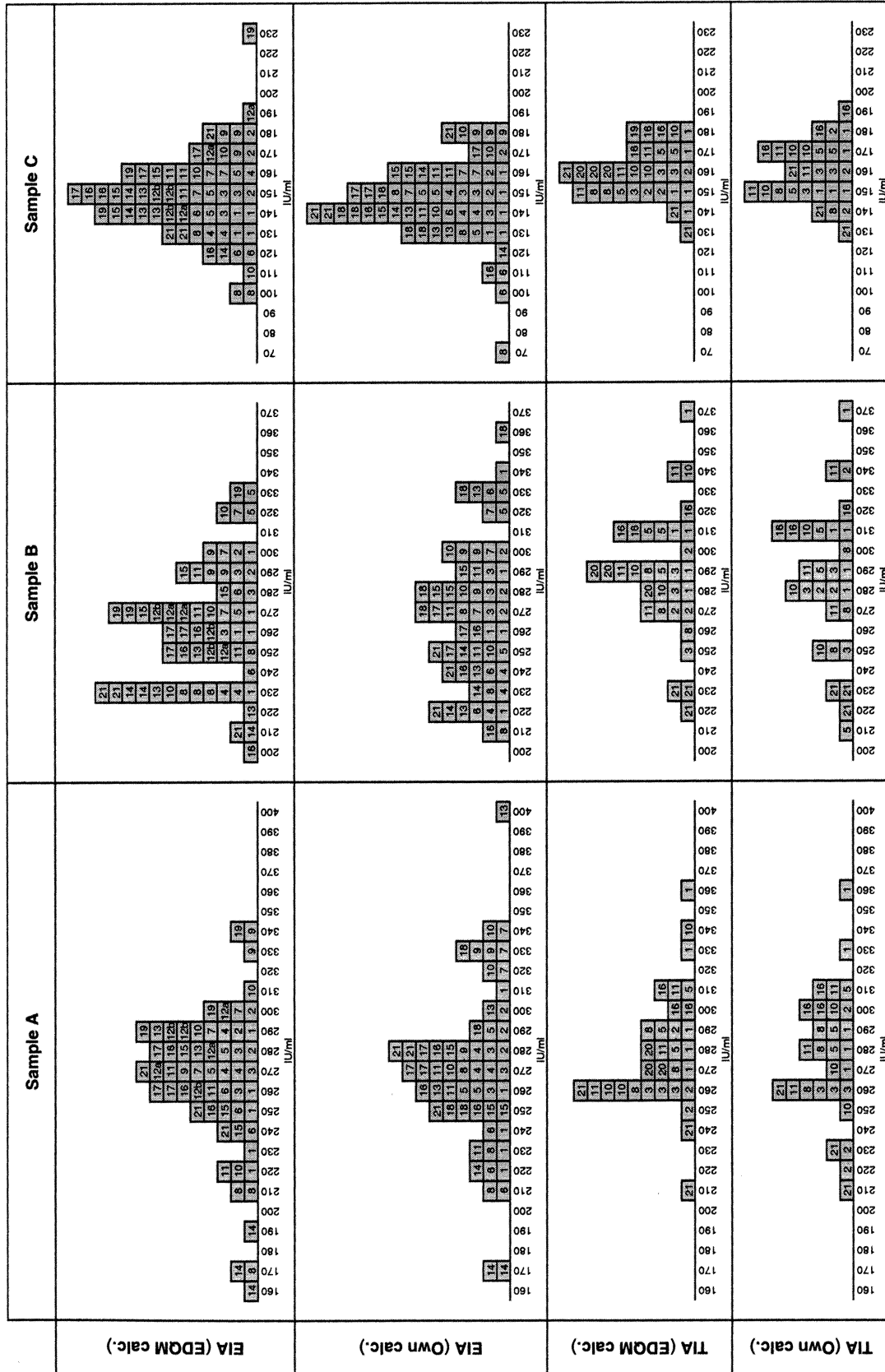
Explanations: Numbers in the boxes are the laboratory codes. The outlying results from laboratory 19 for TIA are not shown. The first and third row of histograms are according to calculations performed at the EDQM. The second and fourth row are according to participants' own calculations.

potencies. The observed loss of potency led to the situation that this condition was no longer fulfilled. It was therefore decided to run a small additional phase with 3 samples (coded D, E and F) including batches with high potency values that were recently estimated by the *in vivo* method.

6.2. Phase 2b – samples D, E and F

A subset of 8 laboratories participated in this additional phase. They are referred to by their code numbers (1 to 8) allocated at random and not corresponding to the coding for the first phase. The laboratories were selected for the fact that they carried out both methods.

6.2.1. EIA

All 8 laboratories carried out the method of EIA. Laboratory 8 submitted results from 3 assays. All other laboratories carried out 2 assays as requested per protocol. Laboratory 7 reported that they used a centrifuge (200 rpm) instead of a microplate shaker (120 rpm).

The raw data were submitted to central calculations at the EDQM using the same statistical models as the phase 2a (see section 5). Again, the applied statistical model produced in most cases satisfactory parallelism and linearity (Table 6) although it was in many cases necessary to exclude one or two of the lowest or highest dose-levels. For laboratory 4 it was necessary to exclude three dose levels.

Most assays showed satisfactory linearity and parallelism. Significant deviations were observed for 4 assays in a total of 4 laboratories. In the case of laboratories 3, 6 and 8, the deviations were probably not relevant because the residual error was rather low compared to the overall error. Indeed, graphical inspection of the plots did not give any reason to suspect fundamental problems with validity conditions. In the case of one laboratory there are reasons to believe that these deviations are relevant due to the rather high residual error.

The majority of the assays had a correlation coefficient higher than 0.99 and confidence intervals within 90 to 111 per cent. Only in 2 assays (Lab 7) did the confidence limits exceed a width of 80 to 125 per cent, demonstrating a satisfactory overall precision of the method.

The unweighted geometric means (GM) and geometric coefficients of variation (GCV) of the potency estimates per laboratory as calculated at the EDQM are displayed in Table 7. Also shown in Table 7 are the overall GM, GCV and median. This is shown for the case that all potency estimates are included, for the case that assays with significant deviations from linearity and/or parallelism are excluded, and for the case that assays with a correlation

Table 6 - Overview of the EIA results (samples D, E and F; potency estimates in IU/ml)

Lab	Assay	Par	Lin	Corr	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)
						Sample D			Sample E			Sample F	
1	1	0.233	0.981	0.991	325	264	(90% - 111%)	360	344	(90% - 111%)	340	279	(90% - 111%)
	2	0.492	0.996	0.988	264	280	(92% - 109%)	304	348	(92% - 109%)	298	324	(91% - 109%)
2	1	0.784	0.723	0.993	314	313	(93% - 107%)	366	363	(93% - 107%)	297	296	(93% - 107%)
	2	0.993	0.971	0.974	190	191	(85% - 117%)	253	266	(85% - 117%)	231	231	(85% - 117%)
3	1	0.271	0.110	0.996	258	257	(95% - 105%)	304	304	(95% - 105%)	275	276	(95% - 105%)
	2	0.001	0.817	0.995	225	233	(94% - 106%)	265	272	(94% - 106%)	216	223	(94% - 106%)
4	1	0.956	0.277	0.933	297	286	(84% - 119%)	315	301	(84% - 119%)	298	304	(84% - 119%)
	2	0.031	0.880	0.943	315	279	(84% - 119%)	326	261	(84% - 119%)	298	249	(84% - 118%)
5	1	0.603	0.999	0.992	259	243	(91% - 110%)	291	286	(91% - 110%)	278	264	(91% - 110%)
	2	0.471	1.000	0.993	280	264	(91% - 109%)	312	295	(91% - 109%)	245	235	(91% - 109%)
6	1	0.993	0.785	0.992	273	273	(92% - 108%)	322	322	(92% - 108%)	278	279	(92% - 108%)
	2	0.136	0.016	0.996	248	248	(96% - 104%)	291	291	(96% - 104%)	244	245	(96% - 104%)
7	1	0.430	1.000	0.903	301	322	(71% - 142%)	355	380	(71% - 143%)	245	289	(70% - 142%)
	2	0.869	1.000	0.891	268	276	(68% - 145%)	353	358	(69% - 146%)	321	345	(69% - 147%)
8	1	0.827	0.030	0.994	260	276	(95% - 105%)	299	313	(95% - 105%)	290	295	(95% - 105%)
	2	0.821	1.000	0.988	285	305	(90% - 111%)	333	353	(90% - 111%)	292	307	(90% - 111%)
	3	0.176	0.505	0.999	318	335	(97% - 103%)	280	318	(97% - 103%)	302	312	(97% - 103%)

Explanations:

Par=p-value for deviations from parallelism. Lin=p-value for deviations from linearity. Corr=weighted coefficient of correlation.

Assays with significant deviations from parallelism and/or linearity are printed on a black (p<0.01) or grey (p<0.05) background.

Correlation coefficients below 0.98 are printed on a black background and below 0.99 are printed on a grey background.

Own = Potency in IU/ml as calculated by participants. EDQM = Potency as calculated at EDQM with 95% confidence limits between brackets.

Confidence limits wider than 90-111% are printed on a grey background and wider than 80-125% on a black background.

* = Value considered to be an outlier and not included in further calculations.

Table 7 - Geometric means and coefficients of variation per laboratory for EIA (samples D, E and F; potency estimates in IU/ml)

Lab	Including all results						Excluding assays with significant non-linearity and/or non-parallelism						Excluding assays with correlation coefficient less than 0.99					
	Sample D		Sample E		Sample F		Sample D		Sample E		Sample F		Sample D		Sample E		Sample F	
	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV
1	272	4	346	1	301	11	272	4	346	1	301	11	264		344		279	
2	245	42	311	24	261	19	245	42	311	24	261	19	313		363		296	
3	244	7	288	8	248	16	257		304		276		244	7	288	8	248	16
4	283	2	280	10	275	15	286		301		304							
5	253	6	290	2	249	9	253	6	290	2	249	9	253	6	290	2	249	9
6	260	7	306	7	261	10	273		322		279		260	7	306	7	261	10
7	298	12	369	4	316	13	298	12	369	4	316	13						
8	304	10	327	7	305	3	320	7	335	8	310	1	304	15	315	1	304	4
GM	269		313		276		274		321		286		272		317		272	
GCV	9		10		10		9		8		9		11		10		9	
Median	266		308		268		272		316		290		262		311		270	

Explanations: GM = Geometric mean. GCV = Geometric Coefficient of Variation.

coefficient below 0.99 are excluded. The overall mean potencies are about 270 IU/ml, 320 IU/ml and 280 IU/ml for the respective samples. The inter-laboratory GCV of the laboratory means is between 9 and 11 per cent which can be considered a satisfactory method reproducibility.

6.2.2. TIA

8 laboratories carried out the TIA method. Laboratory 8 submitted results from 4 assays. All other laboratories carried out 2 assays as requested per protocol. Laboratory 7 reported that they used a centrifuge (200 rpm) instead of a microplate shaker (120 rpm).

The raw data were submitted to central calculations at the EDQM using the same statistical models as in the previous phase. A summary of results is given in Table 8.

Most assays show satisfactory linearity and parallelism. Visual inspection of plots where significant deviations from linearity and/or parallelism are indicated revealed no major problems with the data. Table 8 also shows that the confidence limits of the potency estimate are in all cases within 80 to 125 per cent and in only 2 assays they exceed 90 to 111 per cent. Laboratory 7 had a very high estimate for sample F in assay 2. The plots show that the ODs for this sample were all at the baseline making the estimate of the activity unstable. This value was therefore considered to be an outlier and further calculations were performed excluding this result from laboratory 7.

The unweighted GMs and GCVs of the potency estimates are shown in Table 9. Also shown are the overall GM, GCV and median. This is shown for the case that all potency estimates are included and for the case that assays with

Table 8 - *Overview of the TIA results (samples D, E and F; potency estimates in IU/ml)*

Lab	Assay	Goodness of fit			Sample D			Sample E			Sample F		
		Par	Lin	Corr	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)	Own	EDQM	(Conf. Lim.)
1	1	0.032	1.000	0.941	294	295	(97% - 103%)	345	345	(97% - 103%)	333	333	(97% - 103%)
	2	0.994	0.820	0.908	265	265	(93% - 107%)	344	344	(93% - 108%)	260	260	(93% - 107%)
2	1	0.953	0.078	0.919	227	216	(96% - 104%)	269	257	(96% - 104%)	215	214	(96% - 104%)
	2	0.329	0.986	0.957	261	259	(96% - 104%)	288	289	(96% - 104%)	244	245	(96% - 104%)
3	1	0.119	0.511	0.937	254	254	(95% - 105%)	318	318	(95% - 105%)	274	273	(95% - 105%)
	2	0.001	0.074	0.969	279	278	(98% - 103%)	329	328	(98% - 102%)	282	280	(98% - 102%)
4	1	0.008	0.013	0.981	307	287	(98% - 102%)	312	303	(98% - 102%)	272	259	(98% - 102%)
	2	0.148	1.000	0.986	294	300	(98% - 102%)	347	348	(98% - 102%)	278	293	(98% - 102%)
5	1	0.533	0.837	0.920	278	280	(96% - 104%)	318	320	(96% - 104%)	267	268	(96% - 104%)
	2	0.205	0.989	0.811	311	310	(95% - 106%)	368	369	(95% - 106%)	310	310	(95% - 106%)
6	1	0.537	0.344	0.875	263	263	(94% - 106%)	290	290	(94% - 106%)	247	246	(94% - 106%)
	2	0.195	0.111	0.959	245	245	(97% - 103%)	298	298	(97% - 103%)	241	241	(97% - 103%)
7	1	0.862	1.000	0.936	249	233	(92% - 109%)	361	412	(87% - 114%)	317	334	(89% - 112%)
	2	0.552	1.000	0.873	307	309	(94% - 107%)	360	385	(92% - 110%)	338	434*	(85% - 116%)
8	1	0.600	1.000	0.739	282	282	(90% - 112%)	287	289	(90% - 111%)	278	279	(90% - 112%)
	2	0.990	1.000	0.926	294	294	(93% - 108%)	309	309	(93% - 107%)	276	276	(93% - 107%)
	3	0.972	1.000	0.949	282	283	(92% - 109%)	298	298	(92% - 109%)	254	254	(92% - 109%)
	4	0.993	1.000	0.916	302	302	(93% - 108%)	333	333	(93% - 107%)	298	297	(93% - 108%)

Explanations:

Par=p-value for deviations from parallelism. Lin=p-value for deviations from linearity. Corr=weighted coefficient of correlation.

Assays with significant deviations from parallelism and/or linearity are printed on a black (p<0.01) or grey (p<0.05) background.

Correlation coefficients below 0.98 are printed on a black background and below 0.99 are printed on a grey background.

Own = Potency in IU/ml as calculated by participants. EDQM = Potency as calculated at EDQM with 95% confidence limits between brackets.

Confidence limits wider than 90-111% are printed on a grey background and wider than 80-125% on a black background.

* = Value considered to be an outlier and not included in further calculations.

Table 9 - *Geometric means and coefficients of variation per laboratory for TIA (samples D, E and F; potency estimates in IU/ml)*

Lab	Including all results						Excluding assays with significant non-linearity and/or non-parallelism					
	Sample D		Sample E		Sample F		Sample D		Sample E		Sample F	
	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV	GM	GCV
1	280	8	345	0	294	19	265		344		260	
2	236	14	273	9	229	10	236	14	273	9	229	10
3	265	7	323	2	277	2	254		318		273	
4	294	3	325	10	275	9	300		348		293	
5	295	8	343	11	289	11	295	8	343	11	289	11
6	254	5	293	2	243	2	254	5	293	2	243	2
7	268	22	398	5	334		268	22	398	5	334	
8	290	3	307	6	276	7	290	3	307	6	276	7
GM	272		324		276		270		326		273	
GCV	8		12		12		9		13		12	
Median	274		324		277		267		331		275	

Explanations: GM = Geometric mean. GCV = Geometric Coefficient of Variation.

significant deviations from linearity and/or parallelism are excluded. The overall mean potencies were about 270 IU/ml, 325 IU/ml and 275 IU/ml for the respective samples. The inter-laboratory GCV of the laboratory means was between 8 and 13 per cent which can be considered a satisfactory method reproducibility.

6.2.3. Comparisons between EIA and TIA (samples D, E and F)

A graphical impression of the distribution of potency estimates is provided in Figure 2.

Shown are histograms for all samples and methods. The previous phase showed a slight indication that TIA yields somewhat higher estimates than EIA. However, this observation was not confirmed in this additional phase.

6.2.4. Comparisons with assigned potency from the manufacturers (samples D, E and F)

Table 10 shows a summary of the results compared with the assigned potencies from the manufacturers, which were obtained using *in vivo* methods. The results for all samples are close to the assigned potency with a difference of less than 6 per cent. Given the fact that the assigned potency is based on 1 assay only and the observed inter-laboratory GCV for EIA and TIA is about 10 per cent, this difference was considered non-significant.

Figure 2 - Histograms of estimated potencies by EIA and TIA (samples D, E and F; potency estimates in IU/ml)
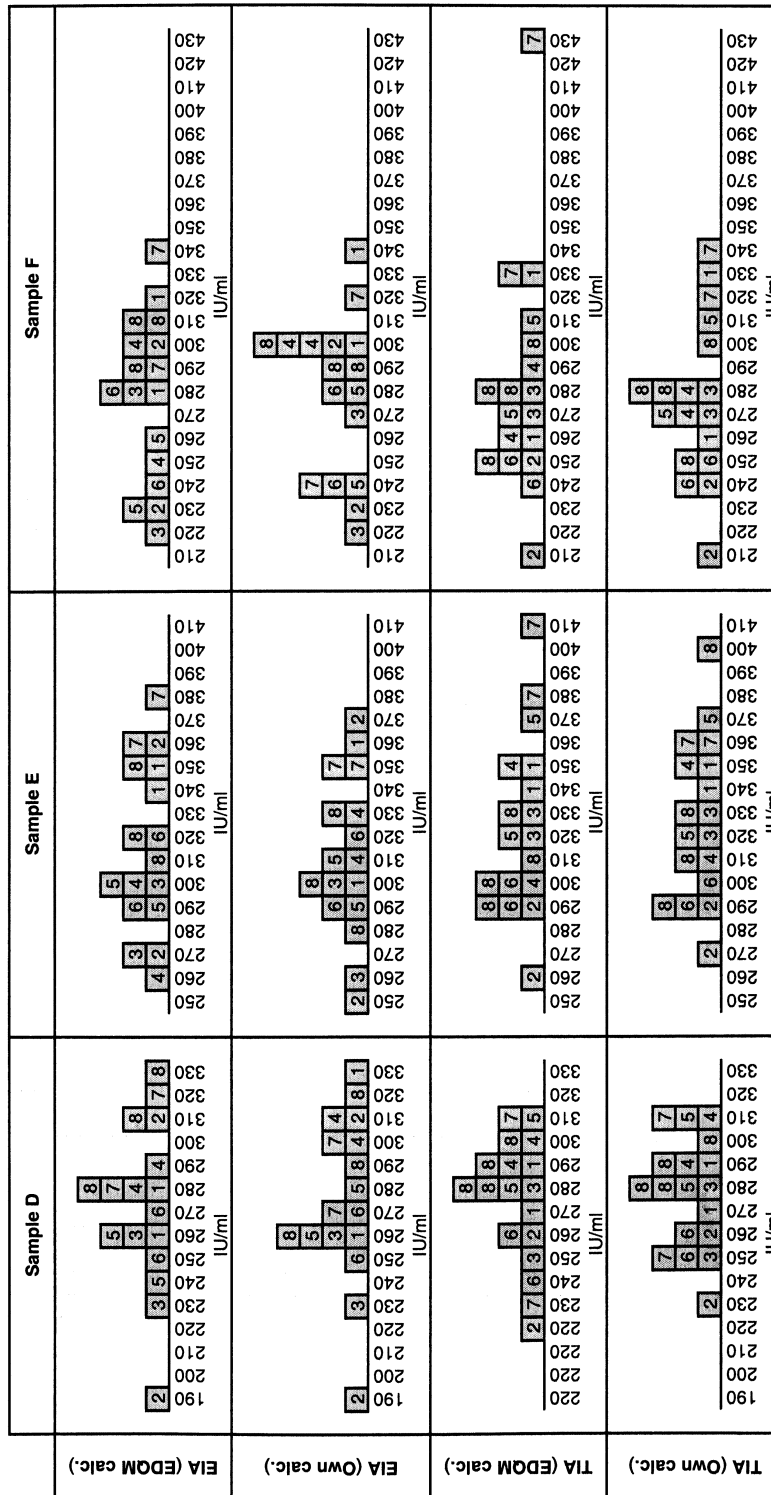
Explanations: Numbers in the boxes are the laboratory codes.
The first and third row of histograms are according to calculations performed at the EDQM. The second and fourth row are according to participants' own calculations

Table 10 - *Summary table of results in comparison with assigned potencies (IU/ml, samples D, E, and F).*

Sample	Assigned potency (manufacturer's *in vivo* method)	Potency estimated with EIA	Potency estimated with TIA
D	280	269	272
E	330	313	324
F	260	276	276

6.3. Overall concordance between methods (samples A to F)

The Lin's concordance correlation coefficient (ρ) is a measure of concordance between 2 methods [7]. A value of 1 indicates perfect agreement and a value of -1 indicates perfect reversed agreement. A value of 0 indicates that there is no correlation between both methods. Figures 3a to 3c show plots for all pairs of methods. The concordance coefficient between EIA and TIA is 0.968. The concordance between EIA and the assigned potencies is 0.960. The concordance between TIA and the assigned potencies is 0.979. These values seem satisfactory.

Figure 3a - *Concordance between TIA and EIA potencies (all samples). Lin's concordance correlation coefficient ρ = 0.968*
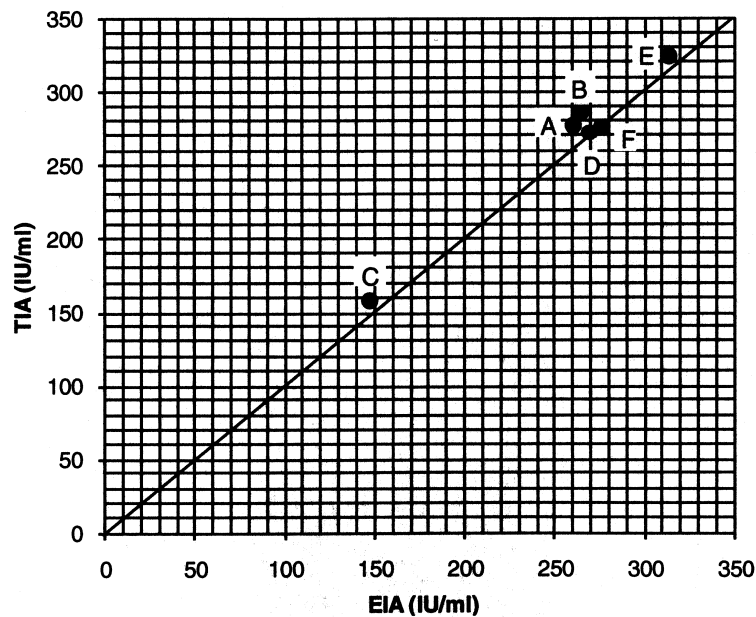
Figure 3b - *Concordance between EIA and assigned (in vivo) potencies (all samples). Lin's concordance correlation coefficient ρ = 0.960*
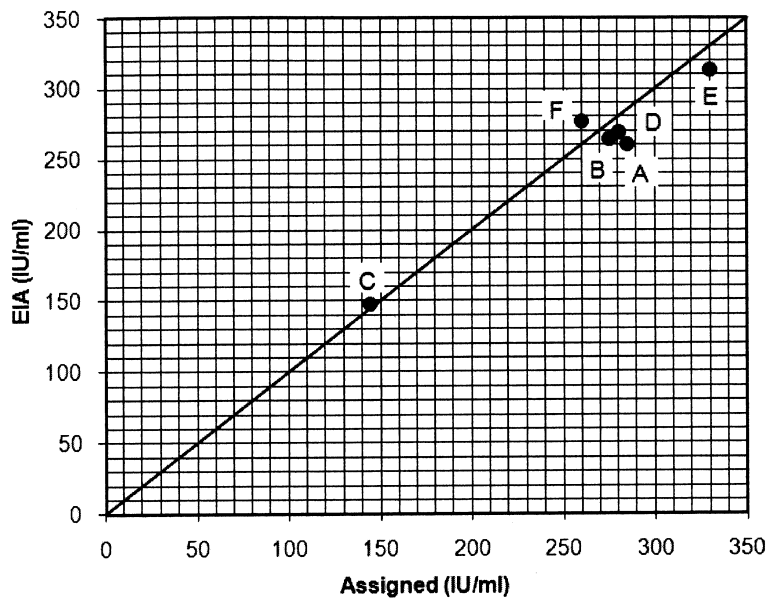
Figure 3c - *Concordance between TIA and assigned (in vivo) potencies (all samples). Lin's concordance correlation coefficient ρ = 0.979*
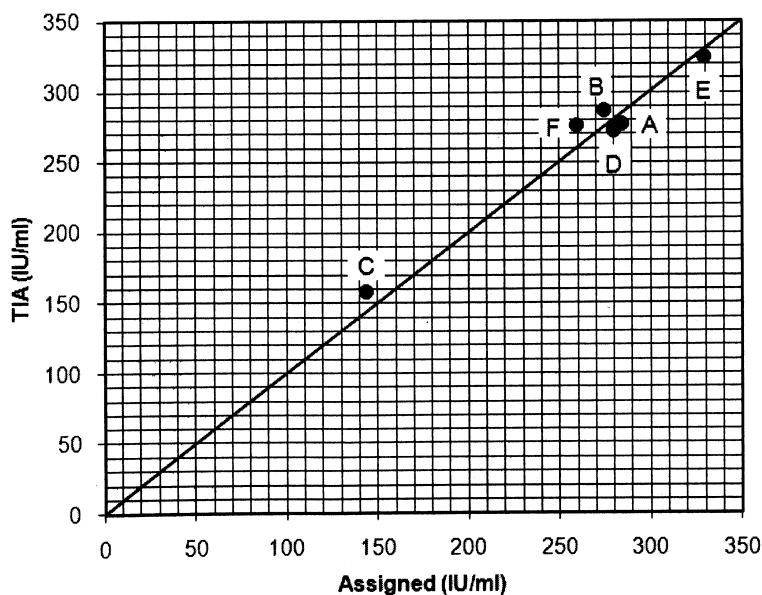
7. DISCUSSION

The establishment of a common alternative method to test the potency of tetanus immunoglobulin products is highly desirable to ensure an improved quality control through method harmonisation amongst manufacturers and Official Medicines Control Laboratories. Agreement on such method would help prevent discrepant results between OMCLs and manufacturers that require re-testings using *in vitro* and *in vivo* methods. The reduced use of the *in vivo* method would further contribute to the better compliance to the European Convention on the protection of animals used for experimental and other scientific purposes [4] and the Council Directive 86/609/EEC [5].

The optimal method for the routine potency estimation of tetanus immunoglobulin batches must be relatively fast, reproducible, accurate (relative to the *in vivo* method) and require general laboratory equipment as well as key reagents for which the supply is ensured.

Both EIA and TIA used in this collaborative study fulfil several of these pre-requisites. *In vivo* testing requires a minimum of 3 days and costly animal handling. Each *in vitro* assay requires less than 2 days of laboratory work and necessitates only standard laboratory equipment. The study also showed that both methods are robust, as commercially available key reagents such as coating antibodies and detection systems can be used and that different sources give comparable potency estimates.

Most participants to this collaborative study were not familiar with the 2 *in vitro* methods used in this study and it was their first experience with the provided SOPs. The results however confirmed an acceptable precision, repeatability and reproducibility of EIA and TIA. Although in a few cases significant deviations were observed for some assays, these events were rare. For both methods, intra-laboratory variations were within acceptable confidence intervals and inter-laboratory variations were about 10 per cent.

The 2 *in vitro* methods also produced comparable results, as confirmed by the satisfactory Lin's concordance correlation coefficient of 0.968 between EIA and TIA.

Products of at least 7 manufacturers are currently available on the European market. The range of their potencies varies between approximately 130 IU/ml and 350 IU/ml. The results of this study show that EIA and TIA are reliable

methods covering the entire range of these approved products.

Due to ethical reasons aiming at avoiding animal use by the participants, it was decided at the start of the project not to estimate the accuracy of the *in vitro* methods relative to the *in vivo* method. The results of this study however demonstrate that there is a good correlation and a good concordance between the potency estimates obtained by the *in vitro* methods and the *in vivo* potency assigned by manufacturers (Lin's concordance correlation coefficients of 0.960 and 0.979 for EIA and TIA respectively).

In all laboratories, both EIA and TIA allowed to discriminate similarly between the low and medium (Phase 2a) as well as between the medium and high (Phase 2b) potency samples. The distinction between the 4 samples (A, B, D, F) assigned with medium potencies was weakened by the uncertainty of these values due to the intrinsic low precision of the in vivo assay, and by the fact that the assigned potencies were based on 1 *in vivo* assay only.

Furthermore the results obtained by the *in vivo* method always represent a range of potency due to the nature of the procedure (challenge to a graded series of volumes of the test toxin). In contrast, the results of the *in vitro* assays are data that fall in a continuous values scale. Estimates are calculated by fitting an exponential model and a 4-parameter logistic curve model for EIA and TIA respectively, thereby allowing a more accurate determination of product potency.

During the collaborative study certain discrepancies between *in vivo* and *in vitro* data were observed for some samples. The closer evaluation of the data demonstrated that both EIA and TIA methods allowed the detection of inconsistencies in assigned potency estimates that were later confirmed, either by the discovery of a reporting error or a loss of potency of the product that was supported by data from *in vivo* re-testing.

8. CONCLUSIONS

The collaborative study showed that both EIA and TIA methods perform satisfactorily and are comparable in terms of assay precision, repeatability and reproducibility. There is a good concordance between the *in vitro* and *in vivo* methods that covers the potency range of the products currently marketed in Europe.

Specific criteria to verify the validity of each assay are important pre-requisites for running a potency assay in the context of a quality assurance system. Such criteria have to be defined in each laboratory during the validation of the assay in view of the implementation of the method. Any other basic components of a common quality assurance system, such as standard operational procedures for handling, cleaning, maintenance and calibration of the instruments, as well as trend control and quality control charts of the results of internal controls have to be established in parallel in each laboratory. It is also advised to ensure that the manufacturers' batch release and Official Medicines Control Laboratories' testings are performed not too far apart in time in order to avoid any possible shelf-life stability issue.

Given the results of this collaborative study, the 2 methods (EIA and TIA) tested in this collaborative study have been considered in 2009 by the Ph. Eur. Group of Experts on human blood and blood products for addition as examples to the Ph. Eur. monograph 0398.

Implementation of a commonly accepted *in vitro* alternative method for human tetanus immunoglobulin batch release purposes will ensure better implementation of the 3R policy (refinement, reduction, replacement of animal use). Acceptance of the EIA/TIA methods tested in this study would imply that individual in-house method development is no longer required and that controls would be performed using an animal-free scheme. A commonly accepted test method would also contribute to the harmonisation of the potency determinations by manufacturers and OMCLs for human tetanus immunoglobulin.

9. ACKNOWLEDGEMENTS

10. PARTICIPANTS (LISTED IN ALPHABETICAL ORDER BY COUNTRY AND NAME)

Balasz C., CSL Bioplasma, Australia

Schindl H., AGES PharmMed, Austria

Mathys E. & Carabin O., Institute of Public Health (IPH), Belgium

Sandberg E., Danish Medicines Agency (DKMA), Denmark

Bagot S., LFB Biomédicaments, France

Mouillot L., Agence Française de Sécurité Sanitaire et Produits de Santé (AFSSAPS), France

Rautmann G. & Cozic G., EDQM, Council of Europe, France

Gross S., Paul-Ehrlich-Institut (PEI), Germany

Rapp M., CSL Behring, Germany

Pogany C., TEVA Pharmaceutical Works, Hungary

Alfano I., Hardis SpA, Italy

Pisani G. & Wirz M., CRIVIB, Istituto Superiore di Sanita (ISS), Italy

Janssen S. & de Vries B., National Institute for Public Health and the Environment (RIVM), The Netherlands

Oerlemans C., Sanquin Plasma Products, The Netherlands

Bucholc B., National Institute of Public Health (NIH), Poland

Ferreira da Silva A., Infarmed, Portugal

Lopez R. & Hosta N., Instituto Grifols, Spain

Carlin G., Medical Products Agency (MPA), Sweden

Sänger M., Swissmedic, Switzerland

North J., Bio Products Laboratory (BPL), UK

Sesardic D., National Institute for Biological Standards and Control, NIBSC, Health Protection Agency, U.K.

11. ABBREVIATIONS

BRP: Biological Reference Preparation; BSA: Bovine Serum Albumin; BSP: Biological Standardisation Programme; EDQM: European Directorate for the Quality of Medicines & HealthCare; EIA: Enzyme-linked ImmunoAssay; GCV: Geometric Coefficient of Variation; GM: Geometric Mean; IS: International Standard; IU: International Unit; NIBSC: National Institute for Biological Standards and Control; OMCL: Official Medicines Control Laboratory; PEI: Paul-Ehrlich-Institut; Ph. Eur.: European Pharmacopoeia; RIVM: National Institute for Public Health and the Environment; SOP: Standard Operating Procedure; TIA: Toxoid Inhibition Assay; TMB: TetraMethylBenzidine; WHO: World Health Organization.

12. REFERENCES

[1] Human tetanus immunoglobulin, monograph 0398. Ph. Eur. 6th edition. Strasbourg, France: Council of Europe; 2008(vol 1).

[2] Gross S, Volkers P, Eckert-Ziem M, Kuschel S, Schäffner G. Validation of *in vitro* potency assays for tetanus immunoglobulin. *Pharmeuropa Bio* 2006;(1):1-6.

[3] Sesardic D, Wong MY, Gaines Das RE, Corbel MJ. The First International Standard for antitetanus immunoglobulin, Human: pharmaceutical evaluation and international collaborative study. *Biologicals* 1993;**21**:67-75.

[4] European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes. Council of Europe Treaty Series (CETS) No. 123. 1986.

[5] Council Directive (86/609/EEC) of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of the European Communities* 1986:L358, 1-29.

[6] CombiStats v4.0 EDQM – Council of Europe; [www.combistats.eu].

[7] Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;**45**(1):255–68.