SciMedCentral

Journal of Immunology & Clinical Research

Research Article

Relevance and Criticality in an External Quality Assessment for the Determination of Diphtheria Antitoxin

Christina von Hunolstein[1]*, Luisa Ralli[1], Antonella Pinto[1], Paul Stickings[2], Androulla Efstratiou[3], Ida Czumbel[4], EDSN EQA Participants[5] and Andrea Gaggioli[1]

[1]*Istituto Superiore di Sanità, Italy*
[2]*National Institute for Biological Standards and Control - South Mimms, Potters Bar, UK*
[3]*Public Health England, Microbiology Reference Services, UK*
[4]*European Centre for Disease Prevention and Control, Sweden*
[5]*European Diphtheria Surveillance Network EQA Participants: A. Decheva, R. Alexiev (Bulgaria), T. Genneou, G. Mitis (Cyprus), N. Guiso, E. Badell (France), A. Sing (Germany), A. Pangalis, H. Alexandrou-Athanassoulis (Greece), A. Pinto, L. Ralli, C. von Hunolstein (Italy), R. Paberza (Latvia), G. Berbers, P. van Gageldonk (Netherlands), M. Steinbakk, A. Aase (Norway), A. Zasada (Poland), M.A. Santos (Portugal), M. Damian, C. Dragomirescu (Romania), V. Lengyelová, A. Belyová (Slovakia), V. Mioč, M. Paragi (Slovenia), K. Broughton, A. Efstratiou, S. Neal (Public Health England, UK), P. Stickings, L. Coombes, D. Sesardic (National Institute for Biological Standards and Control, UK)*

***Corresponding author**
Christina von Hunolstein, Istituto Superiore di Sanità, viale Regina Elena, 00161 Rome, Italy, Tel: +39-06-4990-2036; Fax: +39-06-387115; Email: christina.vonhunolstein@iss.it

Submitted: 27 February 2014

Accepted: 05 March 2014

Published: 15 March 2014

ISSN: 2333-6714

Keywords

- Diphtheria
- Diphtheria antitoxin antibody
- ELISA
- Serological methods
- Pearson's correlation coefficient
- Lin's concordance-correlation coefficient
- Cohen's kappa coefficient

Abstract

Accurate determination of diphtheria antitoxin is of value in determining the rates of immunity within broad populations or the immune status of individuals who may be at risk of infection, assessing responses to vaccination and immunization schedule efficacy. An External Quality Assessment (EQA) study for diphtheria serology was carried out within the European Diphtheria Surveillance Network (EDSN). Sixteen national laboratories from 15 European countries participated by testing a standard panel of 150 sera using their current routine method: VERO cell Toxin Neutralization Test (TNT), Dual Double Antigen time-resolved fluorescence immunoassay (dDA-DELFIA), fluorescent bead-based multiplex assay (MIA), in-house or commercial Enzyme-Linked Immunosorbent Assays (ELISA). The objective of the study was not to identify the best assay but to verify whether laboratories using their current method would categorize (negative, equivocal or positive) a serum sample in the same way. Performance of each laboratory was determined by comparison of results on a quantitative and qualitative basis *versus* TNT results from a single reference laboratory, as this test is considered the *in vitro* "gold standard". Quantitative data were analyzed using Pearson's correlation coefficient as well as Lin's concordance-correlation coefficient. Qualitative diagnostic concordance was established by Cohen's kappa (*k*) coefficient. Performance of laboratories using TNT was very good, whilst the performance using other *in vitro* methods was variable. Laboratories that used ELISA performed less well than those using DELFIA or MIA. EQA is important both for laboratories that use *in vitro* non standardized methods or commercial ELISA kits.

ABBREVIATIONS

dDA- DELFIA: Dual Double-Antigen Time-Resolved Fluorescence Immunoassay; ECDC: European Centre for Disease Prevention and Control; EDSN: European Diphtheria Surveillance Network; ELISA: Indirect Enzyme-Linked Immunosorbent Assays; EQA: External Quality Assurance; IU: International Units; Lab: Laboratory; MIA: Fluorescent Bead-Based Multiplex Immunoassay; TNT: VERO cell Toxin Neutralization Test; ToBI: Toxin Binding Inhibition Test; WHO: World Health Organization

INTRODUCTION

Diphtheria has become a well controlled disease and at present only sporadic cases are observed in Europe and western countries [1,2]. Also in Russia, Latvia and Ukraine the disease appears to be under control [1,2]. Diphtheria is endemic in Asia (particularly in India, Indonesia, Nepal), Africa (Angola, Sudan). However, as the circulation of toxigenic strains is observed in all parts of the world, a constant threat to populations with low levels of seroprotection is posed [1,2]. Cases occurred recently and for the first time for many decades in New Zealand [2].

Clinical diphtheria is caused by toxin-producing corynebacteria. Three species, *Corynebacterium diphtheriae, Corynebacterium ulcerans*, and *Corynebacterium pseudotuberculosis* have the potential to produce diphtheria toxin and hence cause classic res-

piratory diphtheria [3]. More recently, *Corynebacterium ulcerans* has been increasingly isolated as an emerging zoonotic agent of diphtheria from both humans and also companion animals such as cats or dogs, indicating the enduring threat of this thought-to-be controlled disease [1,4,5].

As the morbidity of diphtheria is almost entirely due to diphtheria toxin, protection against disease is dependent on antibodies against the toxin [6,7]. Therefore, accurate determination of anti diphtheria toxin antibodies is essential to establish susceptibility of clinical laboratory workers, to obtain reliable information on the immune status of a person or a given population, to evaluate the immunogenicity of diphtheria vaccines in clinical trials, as well as to monitor long term immunity and thus provide recommendations for vaccination policy. It is therefore, of critical importance to have serological methods that are accurate, reproducible, specific and sensitive.

The *in vivo* toxin neutralization test using guinea pigs or rabbits is regarded as the gold standard method for determining protective levels of serum antitoxin [8]. However, this test requires animals and specialized facilities, is labour intensive, expensive and requires relatively large volumes of test serum. It is therefore, not practical for routine use in serological diagnosis and seroepidemiological studies. Tests using cells in culture have been developed as reliable alternatives to the *in vivo* test for detection of diphtheria toxin and for toxin neutralization [9]. The VERO cell Toxin Neutralization assay (TNT) is also the recommended World Health Organization (WHO) [10] and European Pharmacopeia [11] *in vitro* alternative method, as it provides comparable results to guinea pig protection models for potency testing of vaccines [9,12,13]. However, because this assay is also time consuming and requires cell culture facilities, diagnostic laboratories prefer to use simple format indirect Enzyme-Linked Immunosorbent Assays (ELISA), that offer significant advantages in terms of cost, speed, ease of use and adaptability to automation. Other *in vitro* methods are available, such as the double antigen ELISA (DAE) [14], the double antigen, time-resolved fluorescence immunoassay (dDA-DELFIA) [15], multiple Toxin Binding Inhibition Assay (ToBI) [16] and the fluorescent bead-based multiplex assay (MIA) [17], but none of these are as easy to perform as an indirect ELISA.

As diphtheria has become an uncommon disease for the majority of western countries, from time to time an External Quality Assurance (EQA) study for laboratory diagnosis of diphtheria, molecular typing and serology to strengthen diphtheria reference laboratories assurance is carried out. The last EQA was carried out in 2010 by the European Diphtheria Surveillance Network [18].

In this study, we report the results of the EQA study undertaken in 2012 amongst 16 national laboratories (labs) from 15 European countries that participated using their current routine method for assaying human diphtheria toxin antibodies.

To assess the extent of quantitative diagnostic agreement for diphtheria antitoxin between labs, the data were analyzed by applying Lin's concordance-correlation coefficient *(ρ)* [19] in addition to Pearson's correlation coefficient r. To measure the qualitative diagnostic agreement (concordance) between the different labs the Cohen's kappa (*k*) [20] coefficient was used.

MATERIALS AND METHODS

Study design

Each of the participating laboratories received from the coordinating center, the Istituto Superiore di Sanità (ISS), Rome, Italy, a panel of 150 samples to be tested for diphtheria antitoxin antibodies using an assay of their choice (Table 1). Each laboratory used its own standard curve and included a positive and a negative control sample normally used in the assay.

The standard panel was tested twice by each participant.

The results from the specific diphtheria antitoxin concentration, expressed in IU/ml, were calculated by each centre according to their standard operating procedures and sent by e-mail to ISS.

The study was approved by the Ethical committee of the ISS, Rome, Italy (CE/11/314) and by the Ethical committee of the UOC Immunoematologia e Medicina Trasfusionale, Università degli Studi "la Sapienza", Rome, Italy (C.E Prot. 122/11).

Standard panel construction

The standard panel was prepared using sera kindly donated by blood donors of the UOC Immunoematologia e Medicina Trasfusionale, Università degli Studi "la Sapienza", Rome, Italy.

ISS tested the panel twice by dDA-DELFIA and used the average of the two values to reduce the interassay variability. The panel, containing 150 sera with 300 µl of each specimen, was sent frozen by courier post to each participant and stored at –20°C until testing. Due to casual events, lab II and XV tested 149 samples, Lab IX tested 148 samples.

Assays

In the present EQA study, various assays were used to measure specific human diphtheria toxin antibodies. These included VERO cell TNT, dDA-DELFIA, MIA, and in-house or commercial ELISA methods. The commercial ELISA kits specific for the determination of diphtheria antitoxin antibodies were: Serion ELISA classic, Diphtheria IgG (Serion), Diphtheria ELISA IgG Testkit (Sekisui Virotech GmbH), Anti-Diphtheria toxoid ELISA IgG (Euroimmun), Diphtheria toxoid IgG (The Binding Site Group, product code MK014), VaccZyme™ Diphtheria Toxoid IgG (The Binding Site Group, product code MK114), NovaLisa™ *Corynebacterium diphtheriae* toxin IgG ELISA (NovaTec Immunodiagnostica GmbH), Novagnost™ Diphtheria toxin 5S IgG (Siemens Novagnost), and Diphtheria Ab ELISA (IBL International GmbH). Commercial ELISAs were performed according to the manufacturers' specifications, using reagents that were supplied with the kits. Details of the assays used by each participating lab, including toxin/toxoid and reference antitoxin used and the reported limit of detection are listed in Table 1.

Reference assay

The assay selected as a reference to evaluate the performance of the other assays was the VERO cell TNT from lab I. Using this assay, diphtheria antitoxin levels in individual serum samples were classified as follows, based on WHO guidelines [3], and the work performed by Ipsen [21]: positive, i.e., the full protective level of circulating antitoxin (≥ 0.1 IU/ml); equivocal, partial

protective level of circulating antitoxin (0.01 and 0.09 IU/ml), or negative, providing no protection (< 0.01 IU/ml). Results obtained by the reference lab I showed that of the panel of 150 sera, 63 samples were positive, 54 samples were equivocal, 33 samples were negative.

Data analysis

Raw data comprised estimates for anti diphtheria concentration in IU/ml for each sample in the panel. For those laboratories reporting two results for the same serum sample, the geometric mean was used for all subsequent analyses. In the case of ELISA kits, for serum samples with concentrations reported as 0 IU/ml, an arbitrary value of 0.009 IU/ml, that is considered a negative value of diphtheria antitoxin, was imputed (lab XI, n. of sera 10; lab. XV, n. of sera 25). For values reported as "lower than" or "higher than" the imputed value was decreased or increased by 0.01 IU/ml accordingly (Lab XI, XII, XIII).

Participants' results were compared on a quantitative and qualitative basis. Data were analyzed using IBM SPSS statistics Version 20.0 and Microsoft Excel 2007.

To assess the extent of quantitative diagnostic agreement, the TNT-derived values from the reference laboratory (lab I) and the results obtained with all of the other tests were compared using a scatter plot of antibody measurements on a $\log_{10}$ scale. Pearson's correlation coefficient r and Lin's concordance-correlation coefficient *(ρ)* [19], as well as the slope (β) and the intercept (α) of the regression line, with their corresponding 95% confidence intervals, were calculated for each scatter plot. The data (TNT assay) from the reference laboratory was considered as independent variable (X axis) and data from the other laboratories as dependent variables (Y axis). At present, there is no common opinion on the strength of agreement criteria for Lin's coefficient. Thus, concordance between assays was considered very good when $\rho > 0.90$ and good when $\rho > 0.80$.

To assess the extent of qualitative diagnostic agreement, the panel sera were classified as negative, equivocal and positive as described earlier. In the case of in-house ELISA, sera were qualified as positive when ≥ 0.1 IU/ml [7,22,23]. In the case of commercial ELISA kits, the test results were interpreted according to the manufacturer's indications. Cut offs for all

Table 1: Tests and reference preparations for participant laboratories.

Lab.	Assay	Lowest level of detection (IU/ml)	Diphtheria toxin or toxoid/producer/Lf	Diphtheria reference serum (antitoxin)
I	Vero cell (TNT)	0.0008	Toxin/EDQM/1 Lf/ml	NIBSC-batch 00/496 (human)
I	ELISA in-house method	0.015	Toxoid/NIBSC/1100 Lf/ml	NIBSC-batch 00/496 (human)
II	Vero cell (TNT)	0.016	Toxin/ RIVM/ 1000Lf/Ampoule	NIBSC-batch 66/153 (equine)
III	Vero cell (TNT)	0.00125	Toxin/List Biological Laboratories/GI50 0.94 ng/ml	WHO IS, DI (equine)
IV	Vero cell (TNT)	0.0016	Toxin/ RIVM/ 1000Lf/Ampoule	NIBSC-batch 66/153 (equine)
V	dDA-DELFIA	0.0004	Toxoid/SSI/1911 Lf/ml	WHO, batch DI09-204 (equine)
V	ELISA VaccZyme	0.012	Toxoid	NIBSC-batch 00/496 (human)
VI	Multiplex Immunoassay (MIA)	0.0001	Microspheres conjugated with Toxoid/NVI Bilthoven/4500 Lf	In-house reference serum calibrated in IU (human)
VII	ELISA in-house method	0.0009	Toxoid/BulBio-NCIPD, Bulgaria/1000 Lf	In-house reference serum calibrated in IU (human)
VIII	ELISA Binding Site	0.004	Toxoid	NIBSC-batch 00/496 (human)
IX	ELISA Virotech	0.065	Toxoid	NIBSC-batch 00/496 (human)
X	ELISA Siemens Novagnost	<0.01	Toxin	NIBSC-batch 91/534 (human)
XI	ELISA NovaLisa	0.01	Toxoid	NIBSC-batch 00/496 (human)
XII	ELISA Serion	0.05	Toxoid	1st International standard (Equine), Statens Serum Institute Copenhagen, Denmark
XIII	ELISA IBL	0.08	Toxoid/IBL international GmbH	NIBSC-batch 00/496 (human)
XIV	ELISA Virotech	0.04	Toxoid/Virotech	NIBSC-batch 00/496 (human)
XV	ELISA Euroimmun	0.0004	Toxoid	NIBSC-batch 00/496 (human)
XVI	ELISA Siemens Novagnost automated system	< 0.01	Toxin	NIBSC-batch 91/534 (human)

assays are reported in Table 3 and 4. For each lab, the measure of the qualitative diagnostic agreement (concordance) *vs* lab I, was estimated by Cohen's kappa (*k*) [20]. Landis and Koch's [24] interpretation of *k* coefficient was used to evaluate the strength of agreement: <0.20=poor, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=good, and 0.81–1.00=very good.

RESULTS AND DISCUSSION

EQA studies, proficiency testing studies or inter-laboratory comparison are important studies that allow labs to identify testing problems, compare methods, evaluate and eventually improve their performance. Qualified labs, according to ISO/IEC 17025 (International Standard Organization/International Electronic Committee), are required to participate in these kinds of studies on a regular basis. The goal of an EQA study is not to identify the best assay, as advantages and drawbacks of all methods used by the participants are well known, but to verify if the lab using a specific method would have categorized a serum sample in agreement with the reference VERO cell TNT and to allow participating laboratories to compare performance of the assay they use with those used in other laboratories.

Correlation between assays performed in different labs was not only determined by the Pearson's correlation coefficient r, used in previous studies [18,25] but also by using the Lin's concordance-correlation coefficient ρ, which is a reproducibility index [19].

Four labs (II, III, IV and reference lab I) performed the VERO cell TNT assay. The inter-laboratory comparison of this method showed a regression line close to the line of identity for all labs (Figure 1). Pearson's correlation coefficient r, indicated a very high correlation being all values above 0.94. Lin's concordance-correlation coefficient ρ was above 0.90 in each case.

Lab II had used a lower limit of detection of 0.016 IU/ml and considered negative all those sera that were below this value. Thus, its quantitative concordance correlation (ρ=0.90) (Table 2) with the lab I TNT was slightly lower than TNT of the other labs. Diagnostic agreement, measured by Cohen's kappa, showed that all three TNTs had a very good agreement with the reference TNT (k>0.80) (Table 3). It is worth noting that the performance of the labs using VERO TNT was good even when different protocols and key reagents (toxin, reference antiserum) were used.

Table 2: Quantitative correlation for the reference TNT and all other assays by Pearson's correlation coefficient (r) (slope (α) and intercept (β) of the regression line with their 95% confidence intervals) and by Lin's concordance-correlation coefficient (ρ).

Lab.	Assay	r	α (95 % IC)	β (95 % IC)	ρ (95 % CI)
II	TNT	0.94	-0.09 (-0.16; 0.02)	0.78 (0.74; 0.83)	0.90 (0.87; 0.93)
III	TNT	0.98	0.08 (0.02; 0.14)	1.02 (0.98; 1.05)	0.97 (0.96; 0.98)
IV	TNT	0.98	-0.22 (-0.26; 0.18)	0.85 (0.83; 0.88)	0.97 (0.96; 0.98)
V	dDA-DELFIA	0.92	-0.29 (-0.38; 0.20)	0.82 (0.76; 0.88)	0.91 (0.88; 0.93)
VI	MIA	0.95	-0.31 (-0.38; 0.24)	0.89 (0.85; 0.94)	0.93 (0.91; 0.95)
V	ELISA VaccZyme	0.76	0.07 (-0.08; 0.21)	0.68 (0.59; 0.78)	0.65 (0.56; 0.72)
VII	ELISA In-house	0.55	-0.34 (-0.54; 0.13)	0.54 (0.40; 0.67)	0.52 (0.40; 0.63)
VIII	ELISA Binding Site	0.77	-0.19 (-0.26; 0.12)	0.34 (0.29; 0.38)	0.38 (0.31; 0.44)
IX	ELISA Virotech	0.74	-0.42 (-0.50; 0.35)	0.33 (0.28; 0.37)	0.44 (0.37; 0.51)
X	ELISA Novagnost	0.92	-0.13 (-0.19; 0.06)	0.59 (0.55; 0.63)	0.72 (0.67; 0.77)
XI	ELISA NovaLisa	0.85	-0.66 (-0.75; 0.58)	0.56 (0.51; 0.62)	0.78 (0.72; 0.82)
XII	ELISA Serion	0.92	-0.13 (-0.19; 0.08)	0.54 (0.50; 0.58)	0.66 (0.60; 0.71)
XIII	ELISA IBL	0.83	-0.15 (-0.23; 0.07)	0.49 (0.43; 0.54)	0.57 (0.50; 0.64)
XIV	ELISA Virotech	0.91	0.08 (-0.002; 0.15)	0.70 (0.65; 0.75)	0.74 (0.68; 0.79)
XV	ELISA Euroimmun	0.82	-0.25 (-.038; 0.14)	0.68 (0.60; 0.75)	0.79 (0.72; 0.84)
XVI	ELISA Novagnost automated system	0.93	-0.25 (-0.32; 0.18)	0.73 (0.68; 0.77)	0.90 (0.87; 0.92)
I	ELISA In-house	0.93	-0.28 (-0.33; 0.22)	0.55 (0.52; 0.59)	0.75 (0.70; 0.79)

Table 3: Measure of diagnostic agreement by Cohen's kappa (*k*) for the reference TNT *vs* TNT, dDA-DELFIA, MIA.

Lab Test	TNT			Lab Test	TNT		
	Positive	Equivocal	Negative		Positive	Equivocal	Negative
	≥0.1	0.01-0.09	<0.01		≥0.1	1.01- 0.09	<0.01
Lab. II TNT				Lab. III TNT			
P[a] ≥0.1	58	6	0	P[a] > 0.1	58	10	0
E[b] 0.016-0.09	5	48	2	E[b] 0.01-0.1	5	44	3
N[c] < 0.016	0	0	30	N[c] < 0.01	0	0	30
k		0.86		*k*		0.81	
Lab. IV TNT							
P[a] ≥0.1	48	0	0				
E[b] 0.01-0.09	15	53	3				
N[c] < 0.01	0	1	30				
k		0.81					
Lab. V dDA-DELFIA				Lab. VI MIA			
P[a] ≥0.1	47	7	0	P[a] ≥0.1	39	2	0
E[b] 0.015-0.09	16	39	2	E[b] 0.01-0.09	24	42	3
N[c] < 0.015	0	8	31	N[c] < 0.01	0	10	30
k		0.66		*k*		0.61	

[a] According to the reference lab I TNT, 63/150 samples were positive, 54 samples were equivocal and 33 samples were negative. Lab II tested 149 samples. [a] P, Positive; [b] N, negative; [c] E, Equivocal

Table 4: Measure of diagnostic agreement by Cohen's kappa (*k*) for the reference TNT and in-house or commercial ELISAs kits[a].

Lab Test	TNT			Lab Test	TNT		
	Positive	Equivocal	Negative		Positive	Equivocal	Negative
	≥0.1	0.01-0.09	<0.01		≥0.1	1.01- 0.09	<0.01
Lab. I In house ELISA[b]				**Lab. VII In house ELISA[b]**			
P[c] ≥0.1	62	9	0	P[c] ≥0.1	58	42	10
N[d] <0.1	1	45	33	N[d] <0.1	5	12	23
Lab. VII In house ELISA				**Lab. XII Serion**			
P[c] ≥0.1	58	42	10	P[c] >1.0	8	0	0
E[e] 0.01-0.09	2	6	8	E[e] 0.1-1.0	55	25	2
N[d] < 0.01	3	6	15	N[d] < 0.1	0	29	31
k	0.24			*k*	0.17		
Lab. V VaccZyme				**Lab. XIII IBL Kit**			
P[c] >0.149	59	24	9	P[c] > 1.0	17	0	0
E[e] 0.1-0.149	4	5	0	E[e] 0.1-1.0	43	18	0
N[d] < 0.1	0	25	24	N[d] < 0.1	3	36	33
k	0.36			*k*	0.22		
Lab. VIII Binding Site				**Lab. XIV Virotech**			
P[c] >0.149	59	35	12	P[c] > 1	14	0	0
E[e] 0.1-0.149	4	11	11	E[e] 0.1-1.0	49	27	2
N[d] < 0.1	0	8	10	N[d] < 0. 1	0	27	31
k	0.24			*k*	0.25		

Lab. IX Virotech				**Lab. XV Euroimmun**			
P[c] > 1	7	0	0	P[c] > 1	6	0	0
E[e] 0.1-1.0	47	19	1	E[e] 0.1-1.0	48	14	1
N[d] < 0. 1	8	34	32	N[d] < 0. 1	9	39	32
k	0.14			*k*	0.09		
Lab. X Novagnost™				**Lab. XVI Novagnost™**			
P[c] ≥0.1	63	28	2	P[c] ≥0.1	51	3	0
E[e] 0.01-0.09	0	26	30	E[e] 0.01-0.09	12	49	18
N[d] < 0.01	0	0	1	N[d] < 0.01	0	2	15
k	0.34			*k*	0.63		
Lab. XI NovaLisa							
P[c] ≥0.1	48	4	0				
E[e] 0.01-0.09	15	46	8				
N[d] < 0.01	0	4	25				
k	0.68						

[a] According to the reference lab I TNT, 63/150 samples were positive, 54 samples were equivocal and 33 samples were negative. Lab IX and XV tested 148 and 149 samples, respectively.

[b] For in-house ELISA, as it use only two categories (N and P) the diagnostic agreement by Cohen's kappa (*k*) cannot be calculated; [c] P, Positive; [d] N, negative; [e] E, Equivocal.

The correlation coefficient between the lab V dDA-DELFIA and lab I TNT was r=0.92 (Table 2) and the concordance-correlation ρ=0.91 (Table 2). Performance of lab V in this EQA was in line with the previous study [18]. Lab V had set the cutoff for negative sera at < 0.015 IU/ml; equivocal sera were therefore those included in the range 0.015 – 0.09 IU/ml (Figure 2). Amongst a total of 150 samples, the dDA-DELFIA test identified 47 out of 63 samples as positive, 39 out of 54 samples as equivocal, and 31 out of 33 as negative (Table 3). Thus, 2 TNT negative sera were identified as equivocal, and 16 TNT positive as equivocal. Fifteen TNT equivocal sera were classified as positive or negative. Therefore, the estimated diagnostic agreement of the lab V by dDA-DELFIA in respect to the lab I TNT was *k*= 0.66 (Table 3).

The correlation coefficient between the lab VI MIA and lab I TNT was r=0.95 and the concordance-correlation ρ=0.93 (Table 2). Cut offs used by MIA are the same as those of the TNT (Figure 2). Amongst a total of 150 samples, the MIA identified as positive 39 out of 63 samples as positive, 42 out of 54 samples as equivocal, and 30 out of 33 as negative (Table 3). Thus, 3 TNT negative sera were identified as equivocal and 24 TNT positive as equivocal. Twelve TNT equivocal sera were classified as positive or negative. Therefore, the estimated diagnostic agreement of the lab VI by MIA in respect to the lab I TNT was *k*=0.61 (Table 3).

The majority of the EQA participants performed an indirect ELISA. Two in-house and nine different commercial ELISA kits were used (Table 1). The results obtained by the laboratories using the different ELISAs are shown in Fig. 3. All showed a great deviation from the identity line, usually overestimating the content at low levels of antibodies (< 0.01 IU/ml).

Lab I tested the standard panel by an in-house ELISA, validated against the TNT VERO cell assay using the same human reference as calibrator serum, i.e. NIBSC 00/496; lab VII used an in-house human serum calibrated against WHO IS (equine).

Testing the standard panel by the two in-house methods, lab I and lab VII obtained a r=0.93 and r=0.55, respectively, while ρ corresponded to 0.75 and 0.52, respectively (Table 2). For the two in-house ELISAs, the qualitative agreement with TNT tests was determined using a diagnostic threshold cutoff value of 0.1 IU/ml. Thus, the sera were divided only in two categories: negative (<0.1 IU/ml) and positive (≥0.1 IU/ml). The ELISA performed by lab VII showed poor agreement with the reference TNT assay: 10 samples that were categorized as negative by TNT were positive in the ELISA and 5 positive sera resulted false negative (Table 4). The ELISA performed by lab I categorized one TNT positive serum as negative in ELISA, but no TNT negative samples were categorized as positive in the ELISA (Table 4). Samples that were categorized as equivocal in the TNT assay (i.e. likely to offer some degree of protection) were mostly reported as negative (45/54 samples) by lab I and positive (42/54) by lab VII.

Lab VII categorized the sera also according to the cut offs that are usually used for TNT (Table 4). In this case, the *k* value was calculated and the result obtained was fair (*k*= 0.24).

Lab V tested the panel also using the VaccZyme ELISA kit MK114. The instruction of this kit, contrary to the others, does not give a precise indication on how to interpret the level of antibodies obtained for diagnostic purposes. The manufacturer recommends the use of an equivocal zone of 0.1 - 0.149 IU/ml, where samples falling within the zone should be repeated to confirm that protective levels of anti diphtheria antibodies are present or not. If the level of protection cannot be confirmed, the sample should be referred to a reference laboratory for further testing or a second sample requested. On this basis, therefore, the sera were classified as positive when >0.149 IU/ml and negative when <0.1 IU/ml. Sera within the equivocal range (i.e. 0.1 – 0.149

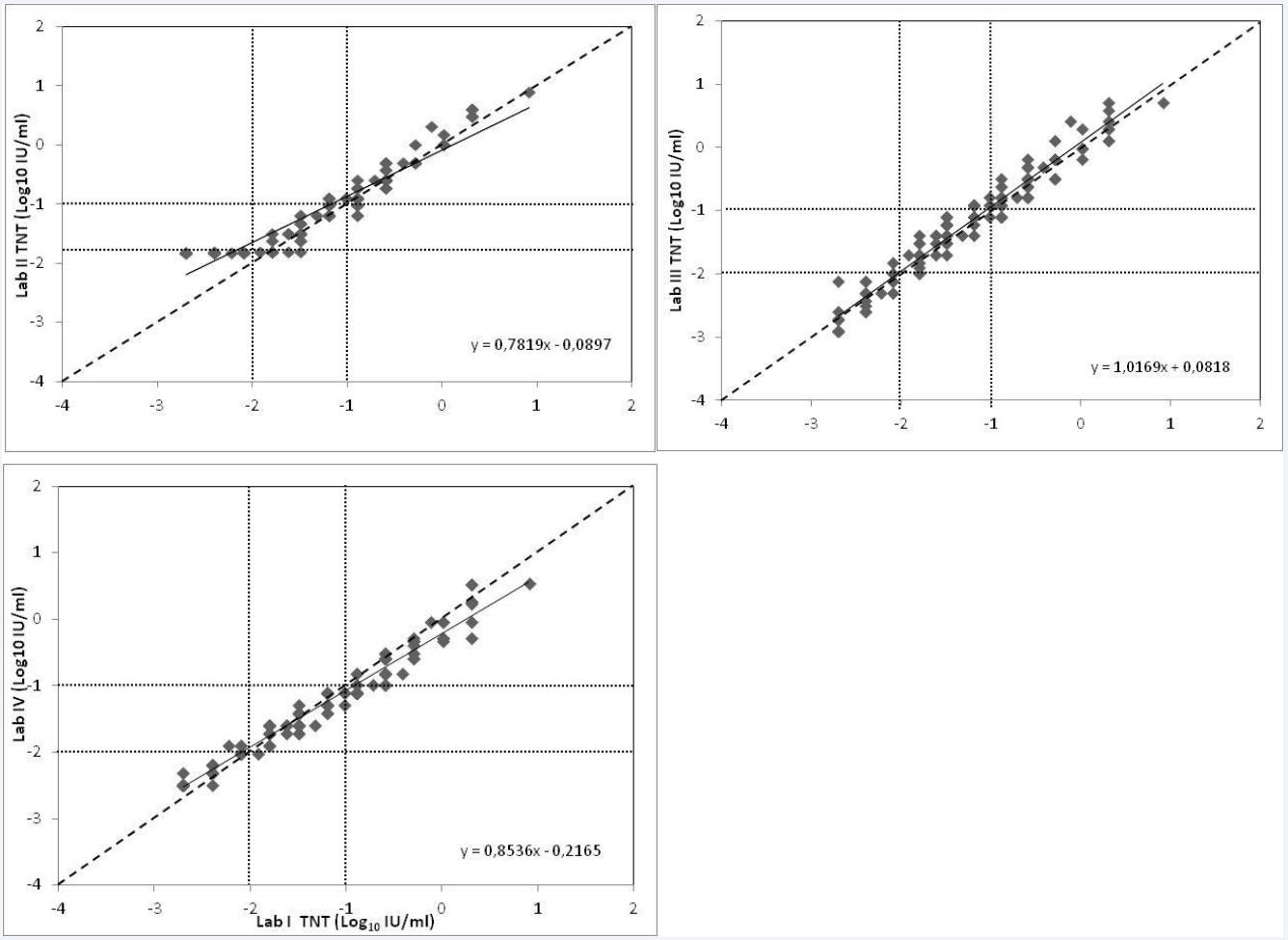


Figure 1 Interlaboratory comparison of the diphtheria antitoxin levels (IU/ml) of the standard panel tested by TNT in laboratory II, III and IV vs Laboratory I. Equation of regression line, regression line (solid), and line of identity (dashed) are shown. Vertical and horizontal lines indicate the cutoffs used by to determine negative (<0.01 IU/ml), equivocal (0.01-0.09 IU/ml) and positive (≥0.1 IU/ml) sera. Lab II used different cutoffs for negative (<0.016 IU/ml) and equivocal (0.016-0.09 IU/ml).

IU/ml) were classified as either equivocal or positive sera by the reference TNT assay; qualitative agreement was fair as k=0.36 and in fact only 27% of TNT negative sera were classified incorrectly by the ELISA (Table 4). According to Pearson's coefficient, the test showed with TNT an r=0.76, while Lin's concordance-correlation coefficient corresponded to ρ=0.65 (Table 2). Results of lab V using this kit were in line with those obtained in a previous EQA [18].

Lab VIII tested the panel only once using the Binding Site ELISA kit MK014 that is indicated only for research use. The test showed with TNT a correlation of r=0.77 (Table 2). Lin's concordance-correlation was fair (ρ=0.38) (Table 2), indicating much better than r the real situation and the differences between the two Binding Site kits (MK104 and MK114). Regarding the qualitative classification of the sera, the kit instructions (Insert Code E014, Version 20th, January 2010) state that "....., it is recommended that each laboratory determines its own normal range". Thus, the same criteria as for the kit MK114 used by lab V was applied. The measure of diagnostic agreement was fair k=0.24: 12/33 TNT negative sera were classified as positive, i.e. false positive and 11/33 as equivocal (Table 4).

Two labs, IX and XIV, tested the standard panel serum by the Virotech ELISA kit. Correlation coefficient comparing the results from lab IX and lab XIV *vs* lab I TNT were different (r=0.74 for lab IX and r=0.91 for lab XIV) as well as the concordance-correlation coefficient (ρ=0.44 and ρ=0.74) (Table 2). Both labs classified the negative sera almost in agreement with TNT test (Figure 3 and Table 4). However, the concordance for positive and equivocal sera was poor, due to the majority of TNT positive sera being classified as equivocal and TNT equivocal as negative, resulting in low k values for each laboratory (Table 4).

Two labs, X and XVI, tested the standard panel serum by the Siemens Novagnost ELISA kit. Lab XVI used an automated system. Also in this case the use of the Pearson's correlation coefficient is misleading as the values are very similar, while the Lin's concordance correlation coefficient is much more realistic (Table 2). In fact, the concordance-correlation coefficient obtained by comparing the results from lab X and lab XVI *vs* lab I TNT corresponded to ρ=0.72 and ρ=0.90, respectively. Also the application of Cohen's k for diagnostic agreement with TNT,

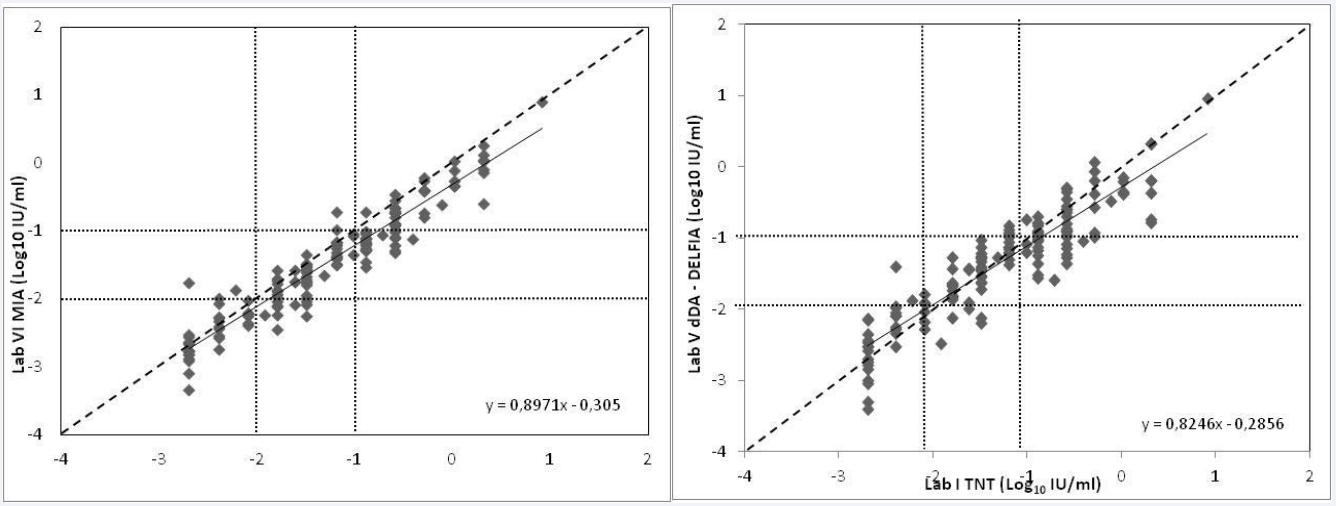


Figure 2 Interassay comparison of the diphtheria antitoxin levels (IU/ml) of the standard panel tested by TNT and dDA-DELFIA and MIA. Equation of regression line, regression line (solid) and line of identity (dashed) are shown. Vertical and horizontal dotted lines indicate the cut-offs used by the laboratories to determine negative (<0.01 IU/ml), equivocal (0.01-0.09 IU/ml) and positive (≥0.1 IU/ml) sera. dDA-DELFIA used a different cutoff for negative (<0.015 IU/ml) and equivocal (0.015-0.09 IU/ml).

underline that the assays carried out by the two labs performed very differently: *k*=0.34 and *k*=0.63 for lab X and XVI, respectively (Table 4); Lab X identified correctly only one TNT negative serum.

Lab XI used the NovaLisa ELISA kit. All sera with a value of IU/ml >0.16 were not further processed. The correlation coefficient with lab I TNT corresponded to r=0.85, while the concordance-correlation was ρ=0.78 (Table 2). The cut offs indicated by this kit to classify the sera in terms of diagnostic interpretation are equivalent of those used for TNT (Table 4). The measure of diagnostic agreement corresponded to *k*=0.68.

Lab XII tested the panel using the Serion ELISA kit. The test showed a high r=0.92; the concordance-correlation with TNT was ρ=0.66 (Table 2). Applying the criteria of results interpretation reported in the kit instruction (Table 4), TNT negative sera were identified correctly, while positive sera were considered equivocal and many equivocal identified as negative. The diagnostic agreement was very poor (*k*=0.17, Table 4 and Figure 3).

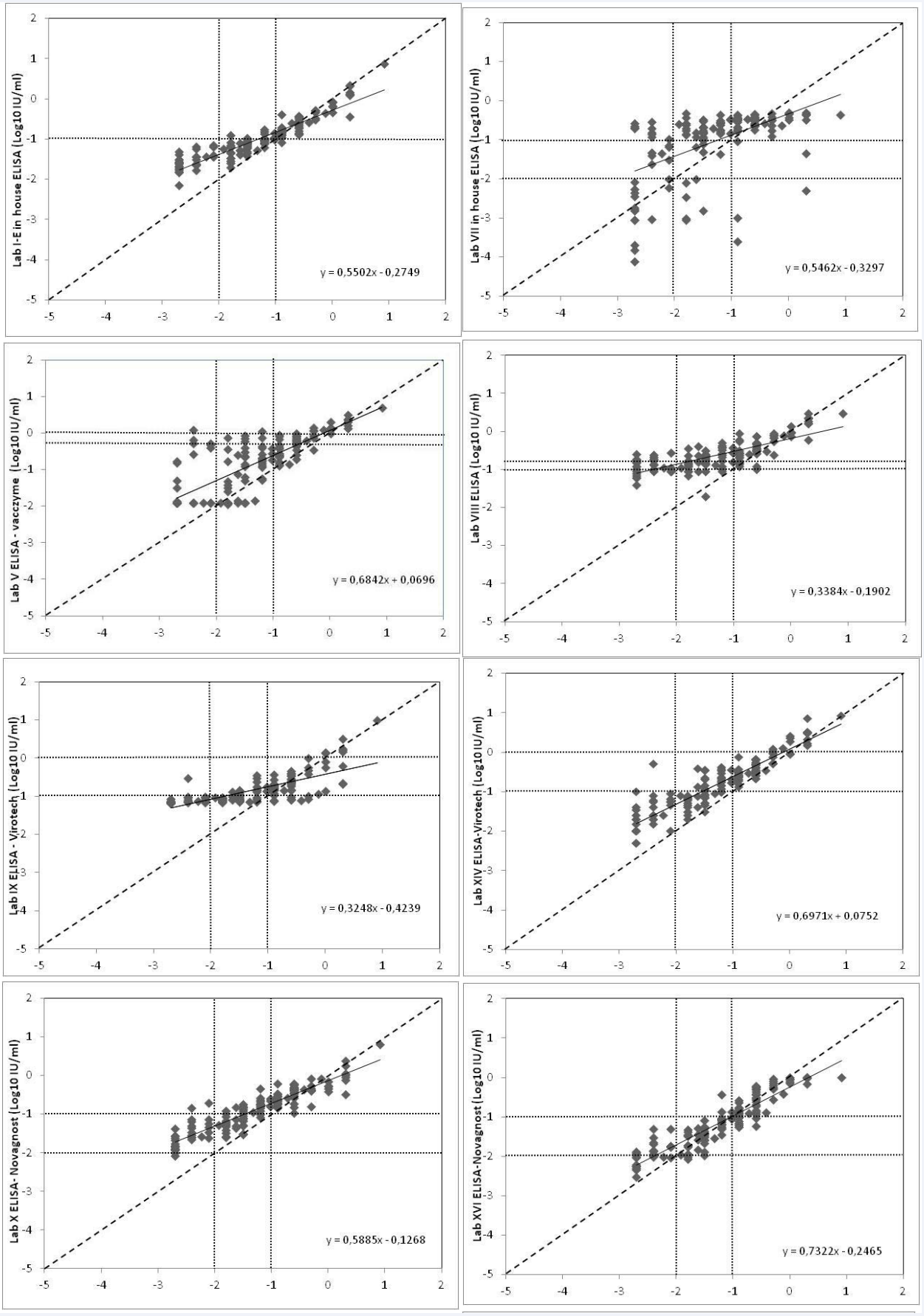
Lab XIII tested the standard panel by the IBL ELISA kit. The test showed *vs* TNT an r=0.83 and a ρ=0.57 (Table 2). The lab classified the negative sera in agreement with the TNT test (Figure 3 and Table 4), but the concordance for positive and equivocal sera was fair due to the majority of TNT positive sera being classified as equivocal and TNT equivocal as negative (*k*=0.22) (Table 4).

Lab XV used the Euroimmun ELISA kit which showed vs TNT an r=0.82 (Figure 3) and a concordance-correlation of ρ=0.79 (Table 2). Even if the lab classified 97% of the negative sera in agreement with TNT test (Figure 3 and Table 4) there was no qualitative agreement between Lab I TNT and lab XV ELISA (*k*=0.09). Many of the TNT positive sera were classified as negative (14%) or equivocal (76%) (Table 4). Seventy-four percent of TNT equivocal sera were categorized as negative.

From the results of this EQA it is evident that this kind of study is extremely important. This is underlined, for example, by the results obtained by the labs using the same commercial ELISA kit. ELISA kit are sold as validated methods inclusive of all key reagents and reference sera to allow calculation of diphtheria antitoxins concentrations in human serum samples. However, it is not always the case that by using such a validated kit the performance of the labs is the same.

When assessing the performance of a lab it is important to graph the data. When the antitoxin values obtained for each serum by the reference lab and the participant lab are very similar, the deviation from the identity line is minimal. In this case the two coefficients, Pearson and Lin are very similar. On the contrary, when there is a deviation from the identity line, particularly as in the case of ELISA, the two coefficients are very dissimilar. In fact, the Pearson's correlation coefficient fails to detect any departure from the line of identity, while Lin's concordance-correlation coefficient, measuring both accuracy and precision of the relationship, is a more reliable coefficient [26].

Furthermore, in this study, and in contrast to the previous EQA [18], the measure of the qualitative diagnostic agreement (concordance) *vs* the reference lab, was estimated by Cohen's kappa (*k*) [20]. Diagnostic agreement was particularly critical between ELISA and TNT. The difficulties derive from the setting of the cutoff. ELISA usually overestimates the antibodies levels in the range 0.001 to 0.01 IU/ml [27,28], measuring not only functional antibodies, but also IgG binding a variety of epitopes of the diphtheria toxin/toxoid. For this reason cutoff for positive sera should be 10 times higher than those applied for TNT. Therefore, sera with antibodies levels of <0.1 IU/ml are considered negative and all those with antibodies levels of ≥ 0.1 IU/ml are to be considered positive [22, 23]. Using kits cutoffs some laboratories obtained low Cohen's *k* values depending on the division of sera in the three ranges of negative, equivocal and positive. Several laboratories also classifying correctly TNT negative sera, however categorized TNT equivocal or positive sera as negative. The clinical implication of over- or under-estimation


Lab I-E in house ELISA (Log10 IU/ml)
y = 0,5502x - 0,2749
Lab VII in house ELISA (Log10 IU/ml)
y = 0,5462x - 0,3297
Lab V ELISA - vaccyme (Log10 IU/ml)
y = 0,6842x + 0,0696
Lab VIII ELISA (Log10 IU/ml)
y = 0,3384x - 0,1902
Lab IX ELISA - Virotech (Log10 IU/ml)
y = 0,3248x - 0,4239
Lab XIV ELISA-Virotech (Log10 IU/ml)
y = 0,6971x + 0,0752
Lab X ELISA- Novagnost (Log10 IU/ml)
y = 0,5885x - 0,1268
Lab XVI ELISA-Novagnost (Log10 IU/ml)
y = 0,7322x - 0,2465

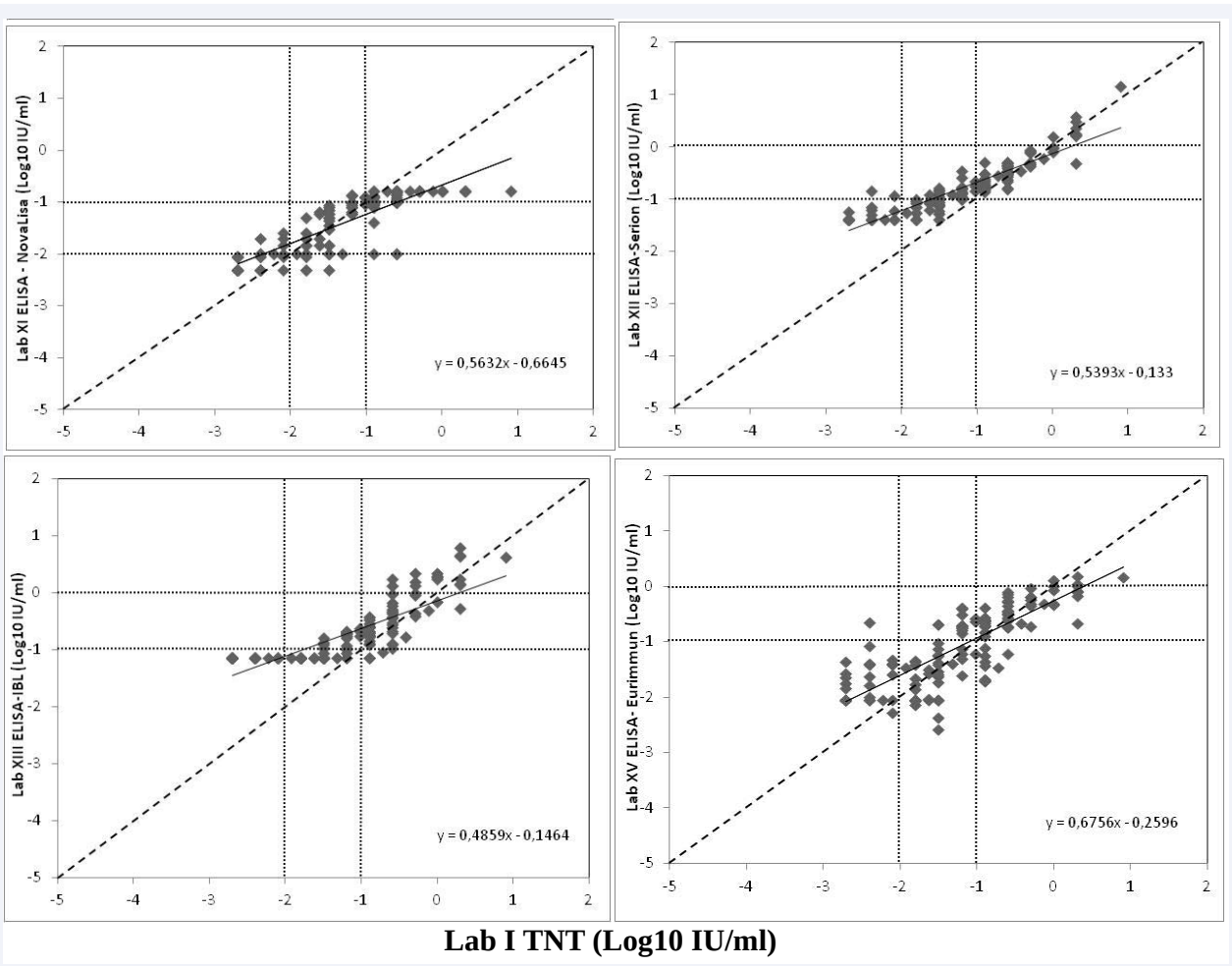


Figure 3 Interassay comparison of the diphtheria antitoxin levels (IU/ml) of the standard panel tested by TNT and eleven different ELISA tests. Equation of regression line, regression line (solid) and line of identity (dashed) are shown. Vertical and horizontal lines indicate the cutoffs used by the laboratories to determine negative, intermediate and positive sera.

of diphtheria antibody titers would be that some subjects may be wrongly assumed to require or not require immunisation. However, misidentification of antibody levels by ELISA *vs* TNT not only occurred in this EQA study, but also in the previous EQA [18]. Thus, the use of ELISA kits need to be carefully considered as there are intrinsic problems related to the nature of the assay. It remains to be seen whether the introduction and use of the recently established 1st International Standard for Diphtheria Antitoxin, Human [29] will lead to improved performance and harmonization of diphtheria serology assays in the longer term.

ACKNOWLEDGEMENT

We thank the blood donors of the UOC di Immunoematologia e Medicina Trasfusionale, Università degli Studi "La Sapienza", Rome, Italy and Prof. G. Girelli, Responsible of the UOC. We also acknowledge the collaboration of Dr A. Fakeri and M. Santonicola of the UOC for the blood collection.

We are grateful to the participants of the EDSN that took part in the EQA study.

This work was in part supported by the European Centre for Disease Prevention and Control (ECDC), Programme on European Diphtheria Surveillance Network (EDSN), ECD.2796, implementing Framework Contract No. ECDC/10/002.

REFERENCES

1. ECDC Surveillance Report. Annual epidemiological report. Reporting on 2011 surveillance data and 2012 epidemic intelligence data. 2013.
2. World Health Organization. Diphtheria reported cases. World Health Organization, Geneva, Switzerland. 2014.
3. Efstratiou A, Maple PAC. Diphtheria. Laboratory diagnosis of diphtheria. The expanded programme on immunization in the European region of WHO. 1994. WHO ICP/EPI 038.
4. Zakikhany K, Efstratiou A. Diphtheria in Europe: current problems and new challenges. Future Microbiol. 2012; 7: 595-607.
5. Dias AA, Santos LS, Sabbadini PS, Santos CS, Silva Junior FC, Napoleão F, Nagao PE. Corynebacterium ulcerans diphtheria: an emerging zoonosis in Brazil and worldwide. Rev Saude Publica. 2011; 45: 1176-1191.
6. Begg N. Diphtheria. Manual for the management and control of diphtheria in the European Region. The Expanded Programme on

Immunization in the European Region of WHO. WHO ICP/EPI 038 (B). World Health Organization, Geneva, Switzerland. 1994.

7. Sheifele DW, Ochnio JJ. The immunological basis for immunization series. Module 2: diphtheria update 2009.

8. Jensen C. Die intracutane Kaninchemethode zur Auswertung von Diphtherie-toxin und Antitoxin. Acta Pathol Microbiol Scand. 1933; 14: 1-21.

9. Miyamura K, Nishio S, Ito A, Murata R, Kono R. Micro cell culture method for determination of diphtheria toxin and antitoxin titres using VERO cells. I. Studies on factors affecting the toxin and antitoxin titration. J Biol Stand. 1974; 2:189-201.

10. Recommendations for diphtheria, tetanus, pertussis and combined vaccines, Amendments 2003. WHO Technical Report Series, No. 927, Annex 5. In: WHO Expert Committee on Biological Standardization. Fifty-forth report. Geneva, World Health Organization, 2005.

11. European Pharmacopoeia. Assay of Diphtheria Vaccine Adsorbed (2.7.6). 7th edn. Strasbourg, Council of Europe.

12. Gommer AM. Vero cell assay validation of an alternative to the Ph. Eur. diphtheria potency tests. Dev Biol Stand. 1996; 86: 217-224.

13. Winsnes R, Sesardic T, Daas A, Behr-Gross EM. Collaborative study for validation of serological methods for potency testing of diphtheria toxoid vaccines. Pharmeuropa Bio. 2003; 2: 35–68.

14. Kristiansen M, Aggerbeck H, Heron I. Improved ELISA for determination of anti-diphtheria and/or anti-tetanus antitoxin antibodies in sera. APMIS. 1997; 105: 843-853.

15. Aggerbeck H, Norgaard-Pedersen B, Heron I. Simultaneous quantitation of diphtheria and tetanus antibodies by double antigen, time-resolved fluorescence immunoassay. J Immunol Methods. 1996; 190: 171-183.

16. Hendriksen CF, van der Gun JW, Kreeftenberg JG. Combined estimation of tetanus and diphtheria antitoxin in human sera by the in vitro Toxin-Binding Inhibition (ToBI) test. J Biol Stand. 1989; 17: 191-200.

17. van Gageldonk PG, van Schaijk FG, van der Klis FR, Berbers GA. Development and validation of a multiplex immunoassay for the simultaneous determination of serum antibodies to Bordetella pertussis, diphtheria and tetanus. J Immunol Methods. 2008; 335: 79-89.

18. Di Giovine P, Pinto A, Olander RM, Sesardic D, Stickings P, Berbers G, et al. External quality assessment for the determination of diphtheria antitoxin in human serum. Clin Vaccine Immunol. 2010; 17: 1282-1290.

19. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics. 1989; 45: 255-268.

20. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20: 37-46.

21. IPSEN J. Circulating antitoxin at the onset of diphtheria in 425 patients. J Immunol. 1946; 54: 325-347.

22. Galazka A. Immunological basis for immunization-Module 2. Geneva, World Health Organization, WHO/EPI/GEN/93.12. 1996.

23. Sesardic D, Corbel MJ. Testing for neutralising potential of serum antibodies to tetanus and diphtheria toxin. Lancet. 1992; 340: 737-738.

24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33: 159-174.

25. von Hunolstein C, Aggerbeck H, Andrews N, Berbers G, Fievet-Groyne F, Maple PA, et al. European sero-epidemiology network: standardisation of the results of diphtheria antitoxin assays. Vaccine. 2000; 18: 3287-3296.

26. Lin L, Torbeck LD. Coefficient of accuracy and concordance correlation coefficient: new statistics for methods comparison. PDA J Pharm Sci Technol. 1998; 52: 55-59.

27. Melville-Smith M, Balfour A. Estimation of Corynebacterium diphtheriae antitoxin in human sera: a comparison of an enzyme-linked immunosorbent assay with the toxin neutralisation test. J Med Microbiol. 1988; 25: 279-283.

28. Walory J, Grzesiowski P, Hryniewicz W. Comparison of four serological methods for the detection of diphtheria anti-toxin antibody. J Immunol Methods. 2000; 245: 55-65.

29. Stickings P, Rigsby P, Coombes L, von Hunolstein C, Ralli L, Pinto A, Sesardic D. Calibration and commutability assessment of the 1st International Standard for Diphtheria Antitoxin Human. Biologicals. 2013; 41: 384-392.

Cite this article

von Hunolstein C, Ralli L, Pinto A, Stickings P, Efstratiou A, et al. (2014) Relevance and Criticality in an External Quality Assessment for the Determination of Diphtheria Antitoxin. J Immunol Clin Res 2(2): 1022.