Michiel Heron*, Marco W.J. Schreurs, Inez-Anne Haagen, Bernard China, Ana Paula Faria, Anna-Riitta Vanhanen, Marc Thelen and Cas W. Weykamp

Allergy: Evaluation of 16 years (2007–2022) results of the shared external quality assessment program in Belgium, Finland, Portugal and The Netherlands

https://doi.org/10.1515/cclm-2023-0862
Received August 7, 2023; accepted November 3, 2023;
published online November 28, 2023

Abstract

Objectives: This paper evaluates 16 year results of the Allergy EQA program shared by EQA organisers in Belgium, Finland, Portugal, and The Netherlands.
Methods: The performance of Thermo Fisher and Siemens user groups (in terms of concordance between both groups, between laboratory CV, prevalence of clinically significant errors) and suitability of samples (stability and validity of dilution of patient samples) are evaluated using data of 192 samples in the EQA programs from 2007 to 2022. Measurands covered are total IgE, screens and mixes, specific IgE extracts and allergen components.
Results: There is perfect (53 %), acceptable (40 %) and poor (6 %) concordance between both method groups. In case of poor concordance the best fit with clinical data is seen for Thermo Fisher (56 %) and Siemens (26 %) respectively. The between laboratory CV evolves from 7.8 to 6.6 % (Thermo Fisher) and 7.3 to 7.7 % (Siemens). The prevalence of blunders by individual laboratories is stable for Siemens (0.4 %) and drops from 0.4 to 0.2 % for Thermo Fisher users. For IgE, the between year CV of the mean of both user groups is 1 %, and a fifteen-fold dilution of a patient sample has an impact of 2 and 4 % on the recovery of Thermo Fisher and Siemens user groups.
Conclusions: The analytical performance of Thermo Fisher is slightly better than that of Siemens users but the clinical impact of this difference is limited. Stability of the sample and the low impact of dilution on the recovery of measurands demonstrates the suitability for purpose of the EQA program.

Keywords: allergy; between laboratory CV; external quality assessment; Siemens; sample quality; Thermo Fisher

Introduction

Diagnosis of type I hypersensitivity is based on anamnesis, provocation as well as blood- and skin testing [1, 2]. For detection of specific IgE (sIgE) antibodies in serum of patients with allergic symptoms different test methods are available. Singleplex allergen extract tests are used to measure sIgE against all the different proteins of a single allergen, while multiplex allergen extract tests make it possible to test multiple allergens at once. The development of singleplex and multiplex tests with allergen components or molecular allergens has made it possible to refine sIgE sensibilisation at protein level thereby increasing clinical specificity and prognostic value of sIgE measurements [3, 4]. Clinical laboratories that perform allergen sIgE antibody and total serum IgE measurements have to demonstrate satisfactory performance in inter-laboratory proficiency testing surveys. Proficiency testing is an external quality control check where the primary goal is to verify that all clinical laboratories accurately measure total serum IgE and correctly identify sera that contain IgE antibody of different allergen specificities.

Previously published data based on results acquired in external quality surveys showed marked differences between

***Corresponding author: Michiel Heron**, Medical Microbiology and Immunology, Diakonessenhuis, Utrecht, 3582 KE, The Netherlands, E-mail: mheron@diakhuis.nl. https://orcid.org/0000-0003-0576-8652
Marco W.J. Schreurs, Microvida Laboratory of Medical Microbiology and Immunology, Elisabeth-TweeSteden Ziekenhuis, Tilburg, The Netherlands
Inez-Anne Haagen, Department of Hematology and Clinical Chemistry Laboratory, OLVG Lab, Amsterdam, The Netherlands
Bernard China, Department of Quality of Laboratories, Sciensano, Brussel, Belgium
Ana Paula Faria, Department of Epidemiology, External Quality Assessment Unit Nacional Institute of Health, Doctor Ricardo Jorge, Lisbon, Portugal
Anna-Riitta Vanhanen, Labquality, Helsinki, Finland
Marc Thelen, Department of Laboratory Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands; and Foundation for Quality Assurance in Laboratory Medicine (SKML), Nijmegen, The Netherlands. https://orcid.org/0000-0003-1771-669X
Cas W. Weykamp, Department of Clinical Chemistry, Queen Beatrix Hospital, Winterswijk, The Netherlands

different methods or techniques for detection of sIgE against allergen extracts as well as between laboratories using identical methods. In these studies however, analysis was either based on a limited number of allergens [5–7], results were collected from a short-term period [8, 9], or data evaluation was limited to analysis of semiquantitative classes [10]. In general, no clinical information was available to interpret discrepancies between methods [6–10].

In this paper we describe 16 years results (2007–2022) of the shared programme of four external quality assessment (EQA) organizers: Sciensano (BE), Labquality Oy (FI), Programa Nacional de Avaliacao Externa da Qualidade (PNARQ; PT), and Stichting Kwaliteitsbewaking Medische Laboratoria (SKML; NL). The program covers the four categories of measurands: total IgE, mixes and screens, specific allergens, and components (Figure 1). Two aspects are evaluated: comparison of the performance of method groups and the relation between sample type and performance.

Comparison method groups

With a few exceptions, participating laboratories used the fluorescence enzyme immunoassay (FEIA) method of Thermo Fisher (previously Pharmacia and Phadia) and the chemiluminescence immunoassay (CLIA) of Siemens (previously DPC). The performance of these user groups is evaluated in terms of (a) concordance between results of user groups, (b) in case of discordance, comparison of analytical results with clinical data, (c) trend in between laboratory coefficient of variation (CV) within both user groups, and (d) trend in prevalence of clinically significant errors of individual laboratories in each of the user groups.

Sample type and performance

One of the challenges of EQA organizers is to warrant commutability, stability, and homogeneity of the samples. And, in the specific case of allergy-EQA, to obtain sufficient volume of sera to manufacture samples with relevant concentrations of relevant measurands. Commutability can be achieved by using single donations of volunteers. An option to acquire sufficient volume is to dilute sera with a (very) high concentration with negative sera, whether this is valid remains to be determined. In this paper we address these issues: the same sample was included in four consecutive years to investigate stability, and a sample diluted in different proportions was included to investigate the validity of dilution (and, as collateral information, linearity of the allergy tests) respectively.

Methods

Design EQA program

The programs are organized according to International Standardization Organization (ISO) 17,043 and samples are prepared and validated according to ISO 13485 and ISO 13528. Samples are manufactured from single donations of volunteers who gave written consent. To obtain samples with a broad range of allergens and concentrations covering the relevant range, donors without a history of allergy (for samples with low concentrations) and donors with a range of different allergies are selected. After donation the serum is dispensed in aliquots in polypropylene vials and frozen. Once a year the samples for the annual cycle are shipped in bulk to the respective EQA organizers. On arrival the EQA organizers forward the samples to their participants who freeze them at –20 °C or below until analysis. The annual cycle consists of 12 samples that have to be assayed per 3 samples at four points in time throughout the year. Guided by the deadlines for submission, laboratories thaw the samples and assay the measurands of their interest. Table 1 shows the options: total IgE, mixes (2 per sample requested), specific allergens (3 per sample); and components (0–6 per sample). Screens refer to inhalation screen (consisting of completely different inhalation allergen sources for example animal epithelia and tree pollen) or food screen (consisting of completely different food allergen sources for example peanut and milk) while mixes consist of allergens belonging to the same allergen source (for example different grasses or different moulds). The table also shows the number of participants for the respective categories of the Thermo Fisher (Thermo Fisher Scientific Inc., Waltham, Massachusetts, U.S.) and Siemens (Siemens Healthineers, Erlangen, Germany) users of the four EQA organizers, and the frequency of the respective

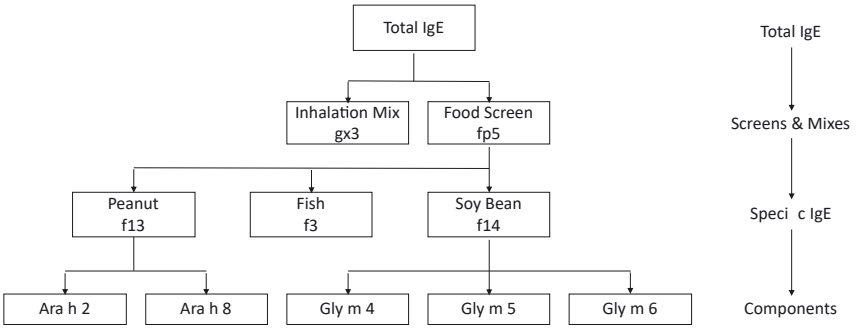

Figure 1: Allergy tests in the laboratory with increasing specificity from total IgE (unspecific) to components (most specific).

measurands in the four categories during the 16 years. Results are submitted to the website (www.allergyqc.com) and reports are on-line available immediately after the deadline. Reports show the result of the laboratory in relation to the means of both user groups in a graph with scaling in kIU/L and classes. The classes are generally interpreted as follows: class 0, <0.35 kIU/L, negative; class 1, 0.35 <> 0.7 kIU/L, equivocal; class 2, 0.7 <> 3.5 kIU/L, positive; class 3, 3.5 <> 17.5 kIU/L, positive; class 4, 17.5 <> 50 kIU/L, strongly positive.

Comparison method groups

Means in kIU/L and classes of Thermo Fisher and Siemens user groups were collected and plotted to investigate the concordance between both groups. Discordance, defined as a negative mean result (<0.35 kIU/L; class 0) in one group and a significant positive result (>0.7 kIU/L; class 2 or higher) in the other group, was compared with available clinical data (anamnesis, intradermal skin tests, skin prick tests, challenge tests, clinical conclusion of the donor) to establish the best fit between analytical result and clinical picture.

Between laboratory CVs of both user groups were collected and evaluated in terms of proportion of CVs below 10 %. To investigate trend, the median between laboratory CVs in two time intervals (2007–2014 and 2015–2022) was calculated. The percentages of clinically significant errors of individual laboratories (defined as a result 2 classes or more apart from the mean of the user group) throughout the years were collected and to estimate trend in prevalence of blunders, the mean percentages in two time intervals (2007–2014 and 2015–2022) were calculated.

Table 1: Participants[a] (A) and measurands (B).

A. Participants

EQA organizer	Total IgE		Mixes		Specific allergens		Components	
	Thermo Fisher	Siemens	Thermo Fischer	Siemens	Thermo Fisher	Siemens	Thermo Fisher	Siemens
Sciensano, BE	63	16			80	22		
Labquality Oy, FI	13	3	10	1	24	6	8	
PNAEQ, PT	4	4	5	4	5	4	1	
SKML, NL	40	15	40	16	44	16	34	
Total company	120	38	76	21	153	48	43	
Total overall	158		97		201		43	

B. Measurands

Category	Frequency	Category	Frequency
Total IgE	192	Specific allergens	
		d1 house dust mite	62
		e1 cat dander	48
		f13 peanut	47
		e5 dog dander	41
Allergen screens/mixes		g5 rye grass	40
fp5 food screen	96	t3 birch pollen	39
Inhalation screen	96	e3 horse dander	38
tx9 tree pollen mix	64	g6 timothy grass	34
gx3 grass pollen mix	63	t4 hazel pollen	32
wx3 weed pollen mix	35	f14 soy	32
mx1 mould mix	30	f4 wheat	29
		w6 mugwort	25
		f2 milk	24
		f1 egg white	20
		f3 fish	20
Components		m2 cladosporium herbarum	17
Ara h 8	18	i3 wasp	10
Ara h 2	17	m6 alternaria alternaria	9
Gal d 1	10	i1 bee	8
Gly m 4	11	f49 apple	4
Gly m 5	9	f20 almond	3
Gly m 6	9	e82 rabbit epithelium	1
Cor a 1	1	f31 carrot	1
Cor a 9	1	f40 tuna	1
Cor a 14	1	k82 latex	1

[a]Mean number of submission in 2022.

Sample type and performance

A donation with very high concentrations of relevant measurands in all four categories (Figure 1) was selected to study stability and the impact of dilutions. The serum was 4, 8, 10, 20, 40, and 60 times diluted with a serum with low (total IgE: 49 kIU/L) and negligible (other measurands) concentrations. The respective dilutions were included as sample in the EQA program to investigate the impact of dilution. One of the dilutions (10 times) was included in four consecutive EQA cycles (2019–2022) to investigate stability.

Results

Comparison method groups

Concordance between results of user groups

Figure 2 shows an overview of the relation between the Thermo Fisher and Siemens user groups for (a) total IgE (n=192), (b) food and inhalation screens (n=192), and (c) the major specific allergens (n=20–62 dots; plots for specific allergens with less than 20 dots not shown). Plots for components are missing because there were no submissions of Siemens users. Each dot represents the mean of the Thermo Fisher group vs. the mean of the Siemens group. It can be seen that for total IgE results of both groups are very close to the unity line over a wide range from less than 10 to more than 1,000 kIU/L. For the screens and the specific allergens results are more dispersed, although there is a clear concordance. Sometimes Siemens results are higher (e.g. cat dander and grass pollen) and sometimes Thermo Fisher results are higher (e.g. hazel pollen and wheat). In Figure 3 the results of specific allergens are plotted in the common sIgE classes 0 to 5. Dark blue fields represent the unity line: the mean class as measured by the Thermo Fisher users is the same as measured by the Siemens users. Example: in 58 cases during the 16 years the mean of both groups was class 3. In light blue are the fields where there is one class difference between the groups. In red there is a difference of at least 2 classes between the groups with the mean of one group being class 0 and the mean of the other group class 2 or 3. Dark blue is interpreted as “perfect concordance” between Thermo Fisher and Siemens (53 % of all cases), light blue as acceptable concordance (40 % of all cases; in 30 % Siemens is one class higher; in 10 % Thermo Fisher is one class higher), and red as discordance between the method groups (6 % of all cases; in 3 % Thermo Fisher is two classes higher; in 3 % Siemens two classes higher).

Comparison discordance analytical results with clinical data

For those 34 “red” cases of discordance the analytical results were compared with the clinical data (Table 2 and Table S1). Of 34 discordant cases, Thermo Fisher had the better clinical accordance in 19 occasions (56 %), in 9 cases (26 %) Siemens results had a better clinical correlation and in 6 cases (18 %) clinical accordance was inconclusive.

Trend in between laboratory CV

Figure 4 shows the inter-laboratory % CV of sIgE measurements against allergen extracts, allergen components and total IgE measurements by Thermo Fisher users (a) and sIgE and total IgE measurements by Siemens users (b). The majority of inter-laboratory % CV per allergen sIgE measurements was below 10 %, for Thermo Fisher and Siemens respectively 86 and 76 %. Most inter-laboratory CVs for total IgE measurements were below 10 %, respectively 99 % for the Thermo Fisher user group and 93 % for the Siemens user group. Given the wide range of measurands and the very wide range of concentrations detailed statistics are not very informative and therefore we did not generate overviews per measurand per year.

Table 3 and Figure S1 (Supplementary Material) show an overview of median between laboratory CVs in the first and second half of the investigated time interval from 2007 to 2022. Example: for the Thermo Fisher user group the mean median between laboratory CV of all specific allergens in 2007–2014 was 7.8 % with per year a range from 5.7 to 13.5 %. In 2015–2022 this dropped to a mean of 6.6 % with a range of 5.4–8.2 %.

Prevalence of clinically significant errors

A result of an individual laboratory is interpreted as a clinically significant error when the class reported by a laboratory is 2 or more classes away from the mean class of the user group. Table 3 shows the percentage of laboratories with a clinically significant error in both time intervals for the respective user groups. For Siemens the prevalence was a stable 0.4 %, for Thermo Fisher the prevalence dropped from 0.4 to 0.2 %.

Sample type and performance

Part A of Table 4 shows the results on stability. Columns show the investigated measurands in the categories specific

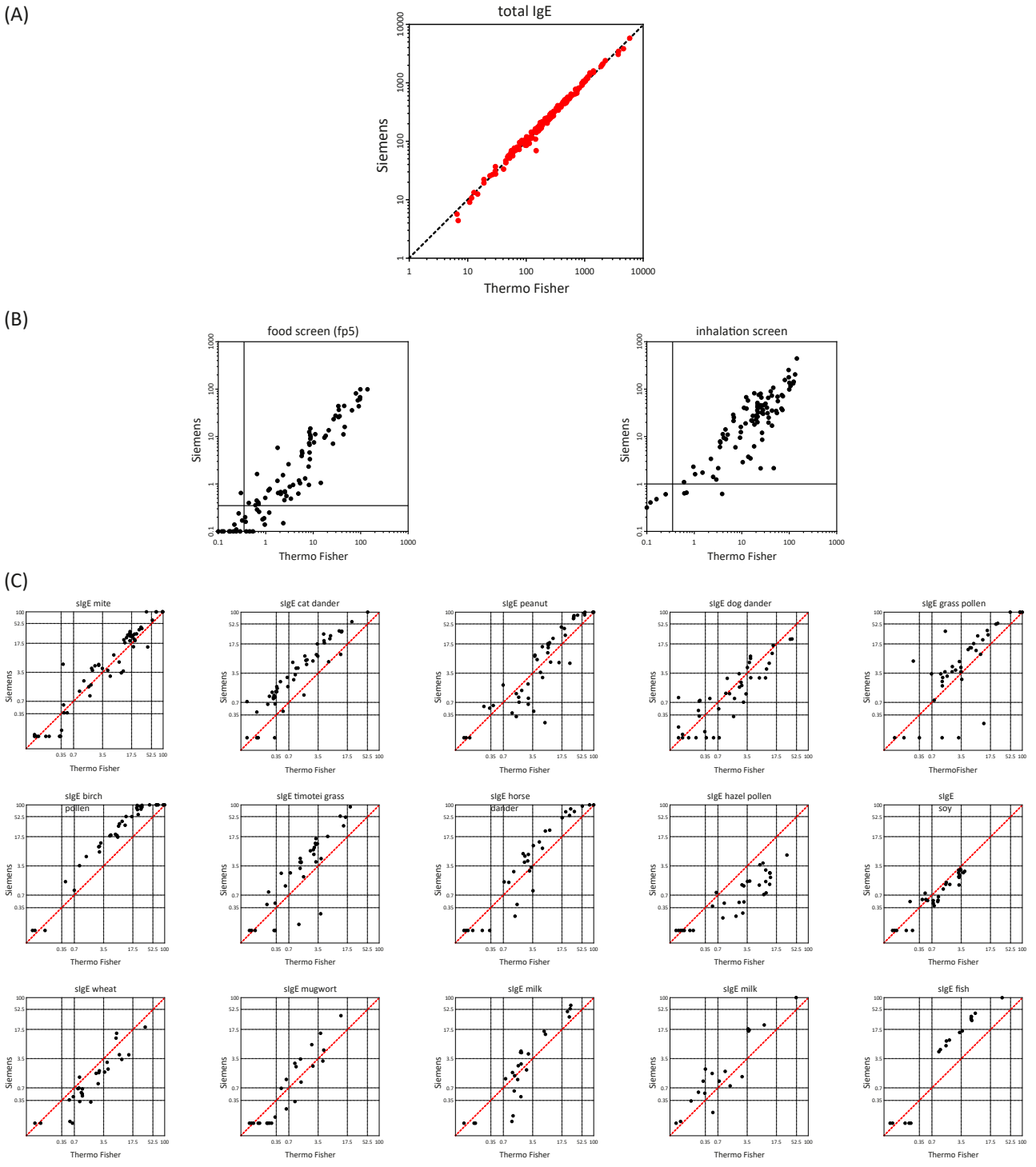

Figure 2: Scatterplots showing the mean results of Siemens and Thermo Fisher user groups of (A) total IgE, (B) food- and inhalation screens and (C) the most frequently tested (≥20 times) specific allergens during 16 years allergy EQA scheme. The black lines (B) represent threshold values for positive/negative. The grid lines (C) represent the five arbitrary semiquantitative specific IgE classes (class 1: 0–0.35 kIU/L, —, class 5: 52.5–100 kIU/L). The dotted black (A) and red (C) diagonal lines represent the line of identity.

allergens (f3, f13, f14), mixes (gx3, fp5), total IgE, and components (Ara h 2, Ara h 8, Gly m 4, Gly m 5, Gly m 6). The first column shows the years in which the serum was included as EQA sample (2019–2022) as well as the statistical parameters mean in 4 years, number of participants, between year standard deviation and between year CV. The other columns

Siemens							
	5				3	27	24
	4				52	19	
	3		1	63	58	4	
	2	14	19	83	25		
	1	12	12	22	1		
	0	116	11	18	2		
		0	1	2	3	4	5
		Thermo Fisher					

Figure 3: Means of classes 1 to 5 in 192 samples as reported by Thermo Fisher (grey; horizontal axis) and Siemens (grey, vertical axis). No difference in class (dark blue), 1 class difference (light blue), 2 classes difference with class 0 for one of the two manufacturers.

show the means of Thermo Fisher and Siemens users for the respective measurands along with the statistical parameters derived. For total IgE it can be seen that the mean of means over 4 years is exactly the same for Thermo Fisher and Siemens users (447) and that the between year CV is 1 % for both groups of users. The picture for the other measurands is variable; in general the between year CV of Thermo Fisher users is lower than of Siemens users. In the years of investigation there were no Siemens users submitting results for components.

Table 3: Trend in between laboratory CV of specific allergens (A) and prevalence clinically significant errors (B).

Manufacturer	A. Median (range) between laboratory CV		B. Prevalence clinically significant errors (1)	
	2007–2014	2015–2022	2007–2014	2015–2022
Thermo Fisher	7.8 % (5.7–13.5)	6.6 % (5.4–8.2)	0.4 %	0.2 %
Siemens	7.3 % (5.7–12.4)	7.7 % (6.9–11.9)	0.4 %	0.4 %

Table 2: Summary of discordance of specific allergen results with clinical data (1) of Thermo Fisher and Siemens user groups in case of contradicting analytical results (2) of both user groups.

Mean class		Total number cases	Accordance with clinical data		
Thermo Fisher	Siemens		Thermo Fisher	Siemens	Undecisive
2 or higher	0	20	Dog (n=1) Milk (n=1) Mugwort (n=1) Grass (n=2) Wheat (n=2) Peanut (n=2) Cla her (n=3)	Hazel (n=1) Milk (n=1) Horse (n=2)	Cladosporium herbarum (n=1) Dog (n=1) Hazel (n=2)
0	2 or higher	14	Wasp (n=1) Horse (n=1) Dog (n=1) Cat (n=4)	Egg white (n=1) Dog (n=2) Cat (n=2)	Cat (n=2)
Total		34	19	9	6

(1) Clinical data: anamnesis, intradermal skin tests, skin prick tests, challenge tests, clinical conclusions. (2) Defined as a difference of 2 classes or more between the means in classes of Thermo Fisher and Siemens user groups (see also Supplementary Table S1).

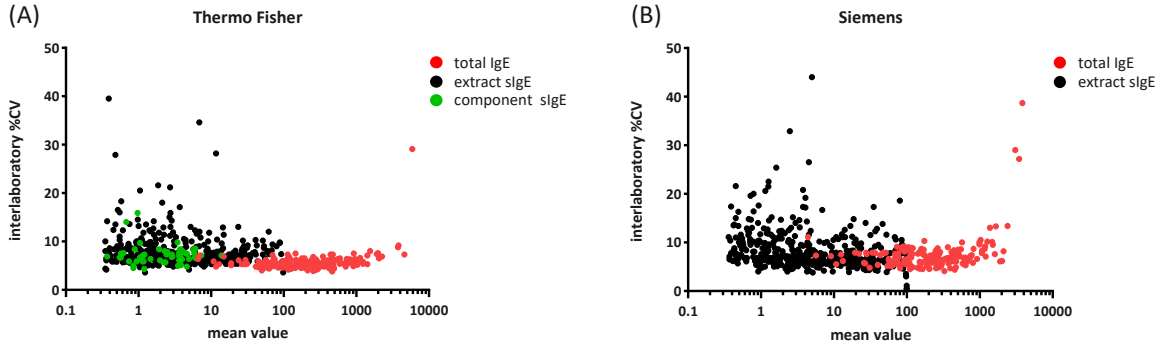

Figure 4: Thermo Fisher (A) and Siemens (B) user group inter-laboratory % CV (vertical axis) plotted against sIgE against allergen extracts, allergen components and total IgE results in kIU/L (horizontal axis) during 16 years of allergy EQA scheme.

Table 4: Stability (A), linearity and dilution (B).

A. Stability

Year	Specific allergens, kIU/L						Mixes, kIU/L				Total IgE, kIU/L		Components, kIU/L				
	f3		f13		f14		gx3		fp5				Ara h 2	Ara h 8	Gly m 4	Gly m 5	Gly m 6
	The Fi	Siem	The Fi	Siem	The Fi	Siem	The Fi	Sie	The Fi	Siem	The Fi	Siem	The Fi	The Fi	The Fi	The Fi	The Fi
2019	6.3	28.4	33.2	77.8	3.0	2.4	9.6	19.2	34.3	25.8	451	441	6.3	2.4	4.0	1.2	3.8
2020	6.1	30.3	33.9	82.8	3.1	2.7	9.7	20.7	34.8	36.7	443	444	5.9	2.3	3.7	1.1	3.6
2021	6.1	35.0	33.5	72.6	3.1	2.1	8.9	21.8	36.2	26.9	450	453	6.0	2.4	4.0	1.0	3.9
2022	6.0	35.4	33.0	70.5	3.1	1.9	9.7	24.4	35.9	38.0	444	451	5.9	2.3	3.9	0.9	3.8
Mean	6.1	32.3	33.4	75.9	3.1	2.3	9.5	21.5	35.3	31.9	447	447	6.0	2.4	3.9	1.1	3.8
n(1)	132	40	141	45	142	46	39	12	53	23	118	35	44	41	22	20	20
SD	0.1	3.5	0.4	5.5	0.1	0.4	0.4	2.2	0.9	6.4	4.1	5.7	0.2	0.1	0.1	0.1	0.1
CV	2 %	11 %	1 %	7 %	2 %	15 %	4 %	10 %	3 %	20 %	1 %	1 %	3 %	2 %	2 %	12 %	3 %

B. Linearity and dilution

Dilution	Specific allergens, kIU/L						Mixes, kIU/L				Total IgE, kIU/L		Components, kIU/L				
	f3		f13		f14		gx3		fp5				Ara h 2	Ara h 8	Gly m 4	Gly m 5	Gly m 6
	The Fi	Siem	The Fi	Siem	The Fi	Siem	The Fi	Siem	The Fi	Siem	The Fi	Siem	The Fi	The Fi	The Fi	The Fi	The Fi
4×											985	1,030					
8×	7.6	42.5	41.2	85.0	3.9	2.4	12.5	27.7	45.1	44.5	497	486	7.3	2.9	4.7	1.0	4.6
10×	6.1	30.3	33.9	82.8	3.1	2.7	9.7	20.7	34.8	36.7	393	394	5.9	2.3	3.7	1.1	3.6
20×	3.2	14.8	16.9	44.3	1.5	1.4	5.1	10.0	17.2	9.5	201	201	2.9	1.2	2.0	0.6	1.9
40×	1.5	9.2	8.9	18.6	0.8	0.5	2.4	6.2	9.0	9.6	101	103	1.5	0.6	0.9	0.2	0.8
60×	1.0	5.1	6.0	15.0	0.5	0.6	1.6	3.3	5.7	3.9	64	64	0.9	0.4	0.6	0.1	0.6
Normalised to dilution 8×=100																	
4×											99	106					
8×	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
10×	100	89	103	122	99	141	97	93	96	103	99	101	101	99	98	131	98
20×	105	87	103	130	96	146	102	90	95	53	101	101	99	103	106	143	103
40×	99	108	108	109	103	104	96	112	100	108	102	106	103	103	96	90	92
60×	99	90	109	132	96	188	96	89	95	66	97	97	92	96	99	100	90
Mean	101	95	105	119	99	136	98	97	97	86	100	100	99	100	100	113	97
n(1)	132	40	141	45	142	46	39	12	53	23	118	35	44	41	22	20	20
SD	3	9	4	14	3	36	3	10	3	25	2	4	4	3	4	23	5
CV	2 %	9 %	4 %	12 %	3 %	26 %	3 %	10 %	3 %	29 %	2 %	4 %	4 %	3 %	4 %	20 %	6 %

Part B of Table 4 shows the results of dilution. The structure of the table is identical to part A, with the exception that instead of years, the degree of dilution is in the first column. The upper part shows the mean measured concentrations in kIU/L. To facilitate interpretation in relation to the impact of dilution, the concentrations are normalized to the 8 times dilution=100. Statistics are for the normalized results. It can be seen that for both groups of users, normalized total IgE results are between 97 and 106 for the respective dilutions, indicating that the decrease of measured measurand is proportional to the degree of dilution. This can also be expressed in the between dilution CV, being 2 and 4 % for Thermo Fisher and Siemens users, respectively. The picture for the other measurands is variable with in general lower CVs for Thermo Fisher and higher CVs for Siemens, especially for f14 and fp5.

Discussion

Comparison method groups

The data described in this report were collected during 16 years (2007–2022) of the shared program of four EQA organizers. Most of the EQA participants used either FEIA method (ImmunoCAP, Thermo Fisher) or CLIA method

(Immulite, Siemens). The total IgE results of both groups were very close to the unity line over a wide range from less than 10 to more than 1,000 kIU/L. This may reflect the availability of a reference material, both CLIA and FEIA total IgE methods are calibrated against the 3rd WHO International Standard for human serum IgE, 11/234 [11].

In contrast, the results of IgE screens and specific allergens were more dispersed. World Health Organization or Joint Committee for Traceability in Laboratory Medicine [12] reference materials for sIgE are not available and preparation of allergen extracts is not standardized. The used source material and composition of the allergen extracts may vary between manufacturers. In addition, allergen binding methods, signal detection and test running time are different between both assays. As a consequence, binding of an individual IgE repertoire will vary by assay which may affect clinical decision-making. Especially when using published cut-offs, used for clinical decision making, measured with the other assay [7, 9, 13–17].

The general agreement between both methods based on semi-quantitative classes was very good as has been shown previously for EQA samples [10], however in their study no clinical data was available to interpret discordant results. In our data, in only 6 % (39 cases) of all cases the difference was two or more classes. In 34 cases the mean of Thermo Fisher results was class 2 and of the Siemens users class 0 or vice versa. Interestingly, the Thermo Fisher results matched the clinical data twice as often compared to Siemens (19 vs. 9). The most prevalent discordant allergen was cat dander; in 24 % (8 cases) the mean sIgE results against cat dander fell in class 0 for Thermo Fisher users vs. class 2 or higher for Siemens users (Table 2 and Table S1). This marked difference in detection of sIgE against cat dander between both methods has been described by Guilloux in 2004 [18], since then it seems not much has changed. Interestingly, a study by Bienboire-Frosini in 2012 showed that cat dander extract measurements may suffer from artefactual bias, as they contain more or less traceable Fel d 1 material or are composed of different Fel d 1 variants [19].

Interestingly, the market for allergy *in vitro* diagnostics in the Netherlands is changing and (new) players such as Sysmex (Hycor, NOVEOS), IDS-Immunodiagnostic Systems and BMD's BioCLIA are on the market. Indeed, when a reasonable amount of participants will make use of these new platforms, analysis of these results will be conducted in the near future.

An interesting goal of sIgE laboratory diagnostics would be the introduction of allergen-specific likelihood ratios related to the measured concentration of sIgE that permit the clinician to estimate the probability of disease [20–22]. The introduction of cross-company sIgE standards for calibration of their assays might contribute to interchangeability of results between different platforms and avoid the need to determine likelihood ratio for each assay separately.

Trend in between laboratory CV

The majority of the inter-laboratory % CV of sIgE measurements against allergen extracts was below 10 % for both user groups. For total IgE measurements the inter-laboratory CV was even better for Thermo Fisher compared to Siemens user group. The between laboratory CVs in the first and second half of the investigated time interval from 2007 to 2022 showed improvement of mean median and more narrow range for Thermo Fisher users. In contrast, no improvement was observed for median and range in between laboratory CV for Siemens users comparing both time frames.

Sample type and performance

Results on stability show a very low between-year CV for total IgE for both groups of users. For the other measurands, between year CVs are low for the group of Thermo Fisher and higher for the group of Siemens users. From the fact that CVs are low for both groups of users for total IgE, low for Thermo Fisher for the other measurands, and the fact that there is no upward or downward trend in the CVs of Siemens, it can be concluded that the samples are stable for at least 4 years. Although no definite evidence, these data also suggest homogeneity of these samples during the storage of 4 years. The fluctuating results of Siemens suggest variation in standardization throughout the years.

Results on dilution show that the between dilution CV of normalized results is low for both groups of users for total IgE, in general lower for Thermo Fisher and higher for Siemens for the other measurands. The higher CVs of the group of Siemens users might be partially caused by batch differences (note that samples were included in different years; compare variability of stability results) rather than to dilution effects. From the low between dilution CVs it can be concluded, at least for the Thermo Fisher method, that linearity is warranted for a wide range of concentrations. And, from perspective of sample preparation, that dilution of sera with a high concentration of measurands is a valid option to acquire sufficient volume of scarce material as

source to manufacture relevant samples. Although good recovery of dilution results validates the use of diluted sera in EQA, poor recovery of dilution results is not a concern in standard patient care as long as undiluted or less diluted samples yield results with correct clinical agreement.

In conclusion, the low interlaboratory CV and proven stability of samples underscores the high quality of allergy EQA, at least in the countries included in this study. Our retrospective evaluation of a large EQA data collection allows powerful assessment of performance of diagnostic laboratories and used methods. In addition to offering a broad panel of allergens including a growing list of allergen components, the SKML organizes a bi-yearly evaluation and education in the form of a symposium for (Dutch speaking) EQA subscribers. The results of a survey on the use and application of allergen components presented in the 2021 EQA evaluation (personal communication, https://www.diakonessenhuis.nl/artsen-zorgverleners/heron), showed the need for standards and education on component resolved diagnostics, in line with Saleem et al. [23] and in especially in light of the rapid developments in the field of molecular allergy [24].

Research ethics: The local Institutional Review Board deemed the study exempt from review.
Informed consent: Informed consent was obtained from all individuals included in this study.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: None declared.
Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Kowalski ML, Ansotegui I, Aberer W, Al-Ahmad M, Akdis M, Ballmer-Weber BK, et al. Risk and safety requirements for diagnostic and therapeutic procedures in allergology: world allergy organization statement. World Allergy Organ J 2016;9:1–42.
2. Ansotegui I, Melioli G, Canonica GW, Caraballo L, Villa E, Ebisawa M, et al. IgE allergy diagnostics and other relevant tests in allergy, a world allergy organization position paper. World Allergy Organ J 2020;13:100080.
3. Canonica GW, Ansotegui IJ, Pawankar R, Schmid-Grendelmeier P, van Hage M, Baena-Cagnani CE, et al. A WAO – ARIA – GA(2)LEN consensus document on molecular-based allergy diagnostics. World Allergy Organ J 2013;6:4517–51.
4. Passalacqua G, Melioli G, Bonifazi F, Bonini S, Maggi E, Senna G, et al. The additional values of microarray allergen assay in the management of polysensitized patients with respiratory allergy. Allergy 2013;68:1029–33.
5. Libeer JC, Van Hoeyveld E, Kochuyt AM, Weykamp C, Bossuyt X. In vitro determination of allergen-specific serum IgE. Comparative analysis of three methods. Clin Chem Lab Med 2007;45:413–5.
6. Wojtalewicz N, Goseberg S, Kabrodt K, Schellenberg I. Six years of INSTAND e. V. sIgE proficiency testing. An evaluation of in vitro allergy diagnostics. Allergo J Int 2017;26:43–52.
7. Wojtalewicz N, Kabrodt K, Goseberg S, Schellenberg I. Evaluation of the manufacturer-dependent differences in specific immunoglobulin E results for indoor allergens. Ann Allergy Asthma Immunol 2018;121:490–5.
8. Hamilton RG. Proficiency survey-based evaluation of clinical total and allergen-specific IgE assay performance. Arch Pathol Lab Med 2010;134:975–82.
9. Kleine-Tebbe J, Poulsen LK, Hamilton RG. Quality management in IgE-based allergy diagnostics. J Lab Med 2016;40:81–96.
10. Koch L, Aberer W. Comparability and quality of IgE-based in vitro allergy diagnosis: 25 years of external quality assessment. Wien Klin Wochenschr 2014;126:634–41.
11. Thorpe SJ, Heath A, Fox B, Patel D, William Egner W. The 3rd International Standard for serum IgE: international collaborative study to evaluate a candidate preparation. Clin Chem Lab Med 2014;52:1283–9.
12. Armbruster D, Miller RR. The Joint committee for traceability in laboratory medicine (JCTLM): a global approach to promote the standardisation of clinical laboratory test results. Clin Biochem Rev 2007;28:105–14.
13. Szecsi PB, Steender S. Comparison of immunoglobulin E measurements on IMMULITE and ImmunoCAP in samples consisting of allergen-specific mouse-human chimeric monoclonal antibodies towards allergen extracts and four recombinant allergens. Int Arch Allergy Immunol 2013;162:131–4.
14. Schuurman J, Perdok GJ, Lourens TE, Parren PWHI, Chapman MD, Aalberse RC. Production of a mouse/human chimeric IgE monoclonal antibody to the house dust mite allergen Der p 2 and its use for the absolute quantification of allergen-specific IgE. J Allergy Clin Immunol 1997;99:545–50.
15. Wang J, Godbold JH, Sampson HA. Correlation of serum allergy (IgE) tests performed by different assay systems. J Allergy Clin Immunol 2008;121:1219–24.
16. Park KH, Lee J, Sim DW, Lee SC. Comparison of singleplex specific IgE detection immunoassays: ImmunoCAP Phadia 250 and Immulite 2000 3gAllergy. Ann Lab Med 2018;38:23–31.
17. Wood RA, Segall N, Ahlstedt S, Williams PB. Accuracy of IgE antibody laboratory results. Ann Allergy Asthma Immunol 2007;99:34–41.
18. Guilloux L, Hamberger C. Évaluation du dosage des IgE spécifiques sur l'Immulite® 2000 DPC. Immuno-Anal Biol Specialisee 2004;19:71–80.
19. Bienboire-Frosini C, Lebrun R, Vervloet D, Pageat P, Ronin C. Variable content of Fel d 1 variants in house dust and cat extracts may have an impact on allergen measurement. J Investig Allergol Clin Immunol 2012;22:270–9.
20. Van Hoeyveld E, Nickmans S, Ceuppens JL, Bossuyt X. Defining thresholds of specific IgE levels to grass pollen and birch pollen

allergens improves clinical interpretation. Clin Chim Acta 2015;450: 46–50.
21. Abrams EM, Chan ES, Portnoy J. Evolving interpretation of screening and diagnostic tests in allergy. J Allergy Clin Immunol Pract 2021;9:4183–91.
22. Bossuyt X, Frans G. The added value of reporting likelihood ratios to laboratory test results in allergy and clinical immunology. J Allergy Clin Immunol Pract 2022;10:1667.
23. Saleem R, Keymer C, Patel D, Egner W, Rowbottom AW. UK NEQAS survey of allergen component testing across the United Kingdom and other European countries. Clin Exp Immunol;188:387–93. https://doi.org/10.1111/cei.12950.
24. Dramburg, Hilger C, Santos AF, de las Vecillas L, Aalberse RC, Acevedo N, et al. EAACI molecular allergology user's guide 2.0. Pediatr Allergy Immunol 2023;34:e13854.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2023-0862).