

Measurement of specific immunoglobulin E: intermethod comparison and standardization

MARIO PLEBANI,* DANIELA BERNARDI, DANIELA BASSO, FRANCO BORGHESEAN, and
DIEGO FAGGIAN

Recently introduced “second-generation” techniques for specific IgE measurement have produced some analytical improvement, offering better clinical sensitivity than previous techniques. The aims of our study were to compare the analytical and clinical performances of four second-generation techniques for allergen-specific IgE measurement in serum and to ascertain whether the new system for reporting quantitative results contributes to greater clinical agreement between findings using the techniques considered. Allergen-specific IgE was measured using the CAP® System, CARLA®, ENEA®, and AlaSTAT®, and the findings were compared. A significant disagreement was found between CAP and ENEA for all allergens and between CAP and CARLA for D1 and G5. However, the clinical discrepancies were reduced by selecting method-specific thresholds using ROC analysis. Second-generation techniques enable us to obtain better standardization of results; however, the identification of a specific threshold appears to be a prerequisite for the appropriate clinical interpretation of the test findings.

Allergic reactions, which are becoming more widespread, are usually diagnosed on the basis of medical history and clinical symptoms. In vitro and in vivo testing, however, play a key role in confirming clinical suspicions and tailoring treatment. Originally described in 1967 by Wide et al. (1), the radio allerge sorbent test (RAST)¹ is the standard technique for the measurement of allergen-specific IgE antibodies in serum. However, its sensitivity is limited by the solid phase used (simple paper cellulose

disk), its automation is complex, and its utilization of radioisotopic reagents causes problems (2).

Several recently introduced “RAST analogs” and “second-generation” methods for in vitro allergy testing that involve a variety of techniques, tracers, allergen extracts, and antibodies provide discordant results because of differences between their assay characteristics, reagents, and threshold criteria. In different test systems there is no consistent correspondence between arbitrary class scores and ranges for IgE antibodies. A myriad of classification schemes have been invented for reporting test results, leading to further confusion among physicians, who use different vendors and clinical laboratories. In 1992 the Executive Committee of the American Academy of Allergy and Immunology recommended that “the arbitrary reference systems with myriad class-scoring schemes should be abandoned in favor of quantitative reporting methods in which test results are reported in units that are proportional to antibody content” (3).

The so-called Pharmacia CAP® Systems (4) and other second-generation techniques have introduced some analytical improvements and a reference curve calibrated against the World Health Organization Standard for IgE 75/502 (5). Results are expressed in quantitative units, kU_a/L, where one kU_a/L corresponds to 2.4 mg of IgE per liter.

The aim of our study was therefore to compare the analytical and clinical performances of four second-generation techniques for allergen-specific IgE measurement in serum and to ascertain whether the new system for the expression of results contributes to better standardization and clinical agreement between the techniques considered.

Materials and Methods

Serum samples from 89 patients (49 females and 40 males; ages, 10–61 years; mean age, 28.5 years) with suspected inhalant allergies were analyzed. Control serum samples were obtained from 20 healthy subjects (10 females and 10 males; ages, 16–58 years; mean age, 27 years) with negative in vivo tests as described below. Participants gave

Department of Laboratory Medicine, Azienda Ospedaliera di Padova, 35128 Padua, Italy.

*Address correspondence to this author at: Servizio di Medicina di Laboratorio, Azienda Ospedaliera di Padova, Via Giustiniani, 2, 35128 Padua, Italy. Fax 39-49-663240; e-mail mariopl@ux1.unipd.it.

¹Nonstandard abbreviations: RAST, radio allerge sorbent test; kU_a/L, kilounits of allergen-specific IgE per liter; AUC, area under the curve; and CI, confidence interval.

Received October 16, 1997; revision accepted June 4, 1998.

their informed consent by signing a specific form. Each sample (10 mL of serum) was divided into several aliquots and stored at -20°C until specific IgE antibody determination. The population studied is the same as that used in a study published elsewhere (6). However, in the present study we present completely novel findings focusing on the intermethod comparison and on the actual level of standardization in this field.

DIAGNOSIS OF ALLERGY

All patients had respiratory symptoms and suffered from rhinitis and/or asthma; all had been referred to our allergy outpatient clinic by their primary care physician. A careful clinical history was taken. Inclusion criteria for the study were as follows: age between 10 and 61 years, no medication that could interfere with the results of skin testing, no immunotherapy in the past year, and no autoimmune disorders, cancer, or other immunologic conditions. The final diagnosis was based on full concordance between case history, clinical data, and in vivo test results. The choice of allergens to be tested in vitro depended on the symptoms and clinical history of each patient. Because some of the data from six patients were discordant, they were not included in the ROC analysis. Enough data to allow a reliable statistical analysis were obtained for five of the allergens considered. In particular, the number of positive and negative findings were, respectively, 30 and 53 for *Dermatophagoides pteronyssinus* (D1), 16 and 65 for cat epithelium (E1), 38 and 46 for *Lolium perenne* (G5), 26 and 57 for wall pellitory (W19), and 15 and 68 for worm wood (W5).

IN VIVO TESTS

Skin testing was performed by the prick method following the instructions of the Italian Society for Allergy and Clinical Immunology (7). We used allergen extracts [in 500 g/L glycerine from dust mites (*D. pteronyssinus*) 5000 protein nitrogen units/mL], pollens (10 000 Pollen Units/mL), grasses (Gramineae), trees (*Betula verrucosa*), and weeds (*Artemisia absinthium* and *Parietaria officinalis*); all were supplied by Bayropharm. The same batches were used throughout the study, and none were near their expiration date.

The skin was pricked with prick lancets (Hollister Stier). Positive and negative controls were also placed using histamine phosphate (5 g/L) and 500 g/L glycerine alone, respectively. Test reactions were read after 30 min, and the size was recorded as a mean wheal diameters ($D+d/2$, where D is the largest of the wheal diameters and d is the largest diameter vertical to D) (8). A wheal reaction corresponding to one-fourth of the diameter produced by histamine was graded as (+), one-half the diameter as (++), equal to the diameter as (+++), and twice the diameter as (++++).

IgE MEASUREMENT

Allergen-specific IgE was measured in serum using four different methods: the CAP System, CARLA[®], ENEA[®], and AlaSTAT[®]. Results were expressed as kU_a/L . Briefly, the fluoroimmunoassay version of the Pharmacia CAP System (Pharmacia & Upjohn Diagnostics, Milan, Italy) was used following the manufacturer's instructions as described elsewhere (9). This assay has six score classes: the 0 class includes all results $<0.35 \text{ kU}_a/\text{L}$ and the sixth class includes all results $>100 \text{ kU}_a/\text{L}$. The CARLA system is a capture assay for the measurement of specific IgE that uses mouse monoclonal antibodies against human IgE coated to the wells of a microtiter plate and biotinylated allergens in solution. Reagents were supplied by Radim (Pomezia-Rome, Italy). The assay, fully automated on an instrument for enzyme immunoassay (BRIO, Radim) as described by us elsewhere (10, 11) has a threshold of 0.5 and $75 \text{ kU}_a/\text{L}$ as the discriminating value for the sixth class. The ENEA System (BioAllergy, Rome, Italy), a fully automated assay for specific IgE measurement, was used following a technique described by us elsewhere (12); in this assay, the specified five-class score had a positive threshold at $0.36 \text{ kU}_a/\text{L}$, and the last class (4th) included all values $>17.51 \text{ kU}_a/\text{L}$. Finally, the AlaSTAT system (Medical Systems Diagnostics, Genova, Italy), a fully automated immunoassay using an enzyme as a tracer and liquid biotinylated allergens, was used as described by Koji (13); this assay provided the same quantitative values and class score as the CAP System.

For reproducibility studies we used three human pooled sera at low, medium, and high specific IgE concentrations. For the antigens considered, the mean values obtained using the CAP system were as follows: for D1, 1.95, 4.6, and $15.5 \text{ kU}_a/\text{L}$; for E1, 2.50, 7.50, and $22.5 \text{ kU}_a/\text{L}$; for W19, 0.90, 2.30, and $21.01 \text{ kU}_a/\text{L}$; for W5, 2.00, 4.6, and $18.5 \text{ kU}_a/\text{L}$; and for G5, 1.61, 5.2, and $17.1 \text{ kU}_a/\text{L}$.

STATISTICAL ANALYSIS

To compare different methods, we made regression analysis as described by Passing and Bablok (14), and data were compared using the method described by Bland and Altman (15). Because no reference method is available for specific IgE measurement, we used the CAP system for the comparative method because its analytical and clinical reliability have been well demonstrated (2, 4, 9). ROC curve analysis was performed using DDU Astute Software (16) to calculate the area under the ROC curve (AUC) and the threshold with the highest clinical efficiency for each allergen. Differences between AUCs were calculated following the method described by Hanley and McNeil (17). The statistical significance of differences between the sensitivity, specificity, and efficiency of the allergen-specific thresholds was calculated using the Confidence Interval Analysis microcomputer program (18).

Results

The mean coefficient of intraassay variation of the four methods obtained in three different human pooled sera at low, medium, and high specific IgE concentrations ranged from 2.3% to 7.7%, imprecision being lower for G5 and D1 (3.3% and 5.5%) and higher for W5 (6.1%). The interassay

reproducibility of the four methods ranged from 4.5% to 11%. Table 1 shows results from linear regression analysis used to compare the methods. The table shows the correlation coefficient (r) for each comparison and its statistical significance, the slope (s) and intercept (i) values for the regression line, and standard error of estimates ($S_{y/x}$). Table 2 shows results of analysis using the bias plot as recommended by Bland and Altman (15). Between-methods agreement was summarized using the mean of the differences (bias). Precision, described by the "limits of agreement", was calculated as the mean difference $\pm$ twice the SD of the differences. Accuracy was assessed by testing whether the mean difference was zero using a 95% confidence interval (CI) for the mean difference between the two methods. Accuracy was considered "present" when the CI included zero, showing no systematic bias; the contrary applied to "absent". Overall the correlation coefficients were satisfactory or acceptable, ranging from 0.72 to 0.98. In three cases the r value was between 0.60 and 0.68. Nevertheless, the slopes frequently differed significantly from 1.0. For ENEA the slopes ranged from 0.18 to 0.51, showing a significant calibration bias. For AlaSTAT and CARLA, the slopes were closer to the line of equality; however, in some cases (W19 for AlaSTAT; D1 and G5 for CARLA) they showed a significant bias. Concerning the intercept, we found some significant deviations from zero. The Bland-Altman plots showing the differences between findings from the different methods demonstrate a significant disagreement between CAP and AlaSTAT for W5 and W19, between CAP and ENEA for all allergens except E1, and finally, between CAP and CARLA for D1, W19, and G5. No significant disagreement between assays was found for all other comparisons. When the data obtained by the two different types of comparison analysis were summarized, a substantial concordance was observed. In general, Bland-Altman analysis confirmed the presence of accuracy in

Table 1. Comparisons by regression analysis.

	ENEA	AlaSTAT	CARLA
D1-CAP (n = 83)	$r = 0.88$ $P < 0.0001$ $s = 0.18^a$ $i = 0.42$ $S_{y/x} = 2.18$	$r = 0.83$ $P < 0.0001$ $s = 0.87$ $i = 1.00$ $S_{y/x} = 13.13$	$r = 0.86$ $P < 0.0001$ $s = 0.57$ $i = 0.35$ $S_{y/x} = 7.7$
E1-CAP (n = 79)	$r = 0.72$ $P < 0.0001$ $s = 0.26$ $i = 0.95$ $S_{y/x} = 3.45$	$r = 0.98$ $P < 0.0001$ $s = 1.15$ $i = -0.20$ $S_{y/x} = 3.24$	$r = 0.96$ $P < 0.0001$ $s = 1.01$ $i = 0.81$ $S_{y/x} = 4.14$
W19-CAP (n = 83)	$r = 0.96$ $P < 0.0001$ $s = 0.20$ $i = 0.09$ $S_{y/x} = 1.20$	$r = 0.65$ $P < 0.0001$ $s = 0.21$ $i = 0.27$ $S_{y/x} = 4.92$	$r = 0.87$ $P < 0.0001$ $s = 1.08$ $i = 3.09$ $S_{y/x} = 12.24$
W5-CAP (n = 83)	$r = 0.96$ $P < 0.0001$ $s = 0.51$ $i = -0.16$ $S_{y/x} = 0.77$	$r = 0.68$ $P < 0.0001$ $s = 0.44$ $i = -0.10$ $S_{y/x} = 2.37$	$r = 0.60$ $P < 0.0001$ $s = 0.77$ $i = 0.15$ $S_{y/x} = 5.10$
G5-CAP (n = 85)	$r = 0.93$ $P < 0.0001$ $s = 0.19$ $i = 0.52$ $S_{y/x} = 2.16$	$r = 0.94$ $P < 0.0001$ $s = 0.87$ $i = -0.18$ $S_{y/x} = 8.84$	$r = 0.87$ $P < 0.0001$ $s = 0.57$ $i = 0.80$ $S_{y/x} = 9.35$

^a s, slope; i, intercept.

Table 2. Comparisons by the Bland-Altman analysis.

	ENEA	AlaSTAT	CARLA
D1-CAP	Bias = 0.88 Agreement limits = -30-44 Accuracy ^a = absent	Bias = 0.19 Agreement limits = -26-126 Accuracy ^a = present	Bias = 3.57 Agreement limits = -20.1-28 Accuracy ^a = absent
E1-CAP	Bias = 1.50 Agreement limits = -19.4-22.4 Accuracy ^a = present	Bias = -0.29 Agreement limits = -7.78-17.19 Accuracy ^a = present	Bias = -0.84 Agreement limits = -8.9-7.2 Accuracy ^a = present
W19-CAP	Bias = 6.3 Agreement limits = -24.9-37.6 Accuracy ^a = absent	Bias = 6.07 Agreement limits = -26.1-38.3 Accuracy ^a = absent	Bias = -3.7 Agreement limits = -27.8-20.3 Accuracy ^a = absent
W5-CAP	Bias = 0.8 Agreement limits = -4.2-5.9 Accuracy ^a = absent	Bias = 0.8 Agreement limits = -6.3-8.0 Accuracy ^a = absent	Bias = 0.1 Agreement limits = -10-10 Accuracy ^a = present
G5-CAP	Bias = 10.6 Agreement limits = -35-56 Accuracy ^a = absent	Bias = 1.9 Agreement limits = -16.7-20.6 Accuracy ^a = present	Bias = 5.07 Agreement limits = -25-35 Accuracy ^a = absent

^a Accuracy assessed using the 95% CI for the mean difference between the two methods.

comparisons for which a good correlation coefficient ($r \geq 0.83$) and slope (0.87–1.15) were found, whereas it showed no agreement for cases in which the slope differed significantly from 1 (0.19–0.57). Exceptions to this rule are the comparisons between CAP and ENEA for E1 and CAP and CARLA for W5 (Fig. 1, top and middle), the Bland-Altman analysis showing the presence of accuracy despite an unsatisfactory slope, intercept, and $S_{y/x}$ in regression analysis. On the other hand, on comparing CAP and CARLA for W19 (Fig. 1, bottom), a good

correlation coefficient ($r = 0.87$) and slope (1.08) were obtained; however, Bland-Altman analysis demonstrated the absence of accuracy.

To evaluate the clinical efficiency of the four techniques, we calculated and compared the AUCs obtained for each allergen, using the different assays. Table 3 shows the comparison between AUCs obtained by the different methods, using CAP as the reference system. Significant differences were found between the CAP System, CARLA, and ENEA for D1, and between CAP and ENEA for G5. Using the same ROC analysis, we selected the threshold assuring the highest possible clinical efficiency for each allergen and each method. The selected thresholds and the corresponding clinical efficiency, sensitivity, and specificity, and their 95% CIs are reported in Table 4.

At the selected thresholds, the clinical efficiencies of the various allergens measured with different methods were compared with the efficiency of the CAP System, calculating the statistical significance of the differences by the 95% CIs for proportion. A comparable diagnostic efficiency was generally found, and a statistically significant difference between methods was found between ENEA and CAP only for D1 and G5 and between CARLA and CAP for W5 and G5.

Discussion

Despite advances in several RAST analogs or second-generation techniques for specific IgE measurement, the clinical interpretation of in vitro findings for IgE antibodies is beset by difficulties.

During recent years, many reports have appeared on the improved analytical performances of the new in vitro techniques; however, little additional information on the diagnostic accuracy of individual IgE antibody tests has been acquired. However, the quantitative potential of new generation techniques of in vitro testing has made it possible to increase diagnostic accuracy and interlaboratory agreement. In fact, quantitative units have important analytical and clinical advantages.

From the analytical viewpoint, quantitative reporting enables the evaluation of the reproducibility and accuracy of specific IgE assays according to well-defined and accepted criteria (3). Moreover, it facilitates intermethod and interlaboratory comparison, thus endorsing more useful proficiency testing programs.

From the clinical viewpoint, quantitative reporting allows threshold selection and a better definition of the relationship between IgE antibody concentrations and symptoms and/or risk of disease. In particular, it allows ROC analysis to be used as an objective method for evaluating the clinical accuracy of in vitro assays, as already demonstrated by our group and by others (6, 19).

We evaluated four different quantitative techniques for specific IgE measurement in serum, focusing our attention on the clinical accuracy of and agreement between these assays. Our first finding, the satisfactory reproducibility of all assays, cannot be considered merely the expression

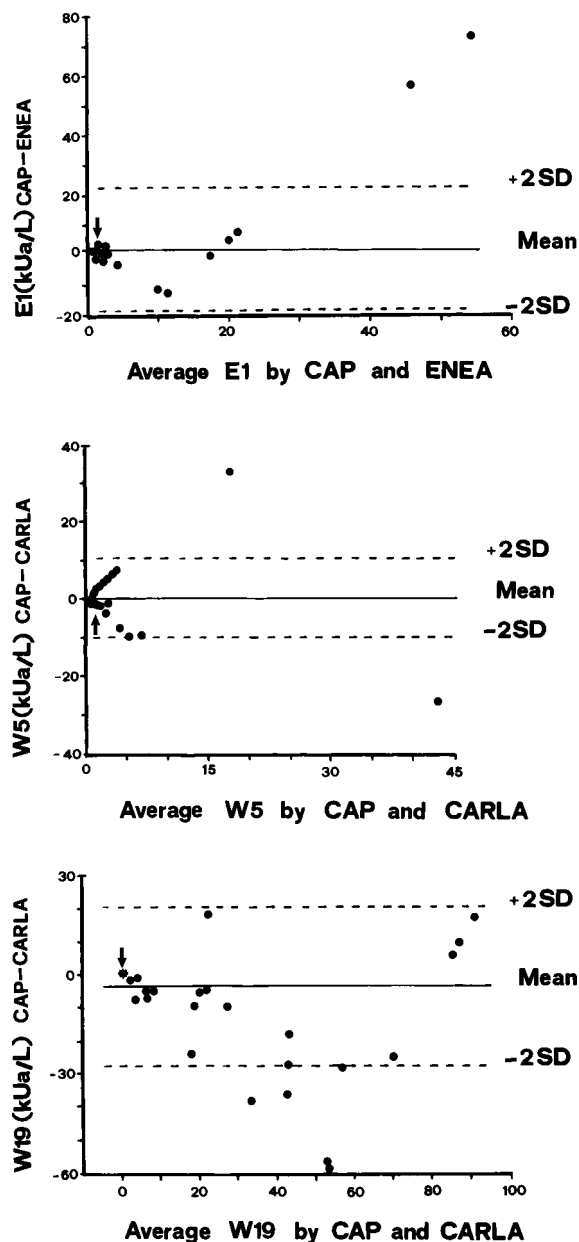


Fig. 1. Bland-Altman plot showing the comparison between (top) E1 values obtained with CAP and ENEA, (middle) W5 values obtained with CAP and CARLA, and (bottom) W19 values obtained with CAP and CARLA.

The square (top) represents 67 cases, the triangle (middle) represents 66 cases, and the asterisk represents 60 cases, indicated by arrows.

Table 3. Comparison between AUC obtained by the different methods using CAP as reference method.

	CAP	ENEA	AlaSTAT	CARLA
D1	AUC = 0.97	AUC = 0.80 z-score = 3.6	AUC = 0.96 NS ^a	AUC = 0.90 z-score = 5.1
E1	AUC = 0.96	AUC = 0.97 NS	AUC = 0.97 NS	AUC = 0.96 NS
W19	AUC = 0.98	AUC = 0.90 NS	AUC = 0.90 NS	AUC = 0.95 NS
W5	AUC = 0.81	AUC = 0.66 NS	AUC = 0.78 NS	AUC = 0.74 NS
G5	AUC = 0.99	AUC = 0.91 z-score = 2.08	AUC = 0.99 NS	AUC = 0.97 NS

^a NS, not significant.

of an analytical improvement. Reproducibility markedly affects the clinical reliability of the assay (20). Other analytic performances (e.g., linearity of the four methods) evaluated elsewhere (2, 10, 11, 13) confirm the higher validity of second-generation techniques compared with the well-known RAST. The intermethods comparison using both linear regression and Bland-Altman analysis gave substantially concordant information and demonstrated a lack of agreement between the results for most of the allergens assayed by the four techniques. No significant disagreement between assays by the two statistical techniques was found between CAP and AlaSTAT for D1, E1, and G5 and between CAP and CARLA for E1. For three assay comparisons, however, linear regression analysis

and Bland-Altman analysis gave differing results. For E1, CAP-ENEA in particular, the lack of agreement between the analytical techniques was probably because of a concentration-dependent bias, although few samples with high concentrations of the specific allergen were assayed. For W5 CAP-CARLA, the lack of a high correlation coefficient may have depended on the presence of two outliers in the high range of values. Finally, for W19 CAP-CARLA, the lack of agreement between Bland-Altman findings and regression analysis was probably because of a constant positive bias of CARLA compared with CAP; this is clearly shown by the significant intercept value (3.09). We thus found that calibration against the same standard (WHO 75/502) does not automatically

Table 4. Clinical efficiency of the different techniques at their selected thresholds.

		Threshold, kU _a /L	Efficiency	95% CI	Sensitivity	95% CI	Specificity	95% CI
D1	CAP	0.40	0.95	0.85–0.97	0.97	0.83–1.00	0.94	0.84–0.99
	ENEA	1.00	0.83	0.73–0.90	0.67	0.47–0.83	0.92	0.82–0.98
	AlaSTAT	0.41	0.95	0.88–0.99	0.90	0.73–0.98	0.98	0.90–1.00
	CARLA	0.75	0.93	0.84–0.97	0.80	0.61–0.92	1.00	0.93–1.00
E1	CAP	0.50	0.99	0.93–1.00	1.00	0.79–1.00	0.98	0.92–1.00
	ENEA	1.00	0.99	0.90–0.99	0.94	0.70–1.00	1.00	0.94–1.00
	AlaSTAT	1.00	0.99	0.93–1.00	0.94	0.70–1.00	1.00	0.94–1.00
	CARLA	0.51	0.97	0.90–0.99	1.00	0.75–0.99	0.97	0.91–1.00
W19	CAP	0.65	0.96	0.90–0.99	0.92	0.75–0.99	0.98	0.91–1.00
	ENEA	0.45	0.91	0.83–0.96	0.77	0.56–0.91	0.98	0.91–1.00
	AlaSTAT	0.70	0.93	0.85–0.97	0.77	0.56–0.91	1.00	0.94–1.00
	CARLA	0.70	0.96	0.90–0.99	0.92	0.75–0.99	0.98	0.91–1.00
W5	CAP	1.00	0.85	0.76–0.92	0.60	0.32–0.84	0.91	0.82–0.97
	ENEA	0.80	0.84	0.75–0.91	0.20	0.04–0.48	0.98	0.92–1.00
	AlaSTAT	0.55	0.90	0.82–0.96	0.53	0.27–0.79	0.98	0.92–1.00
	CARLA	0.51	0.88	0.79–0.94	0.53	0.27–0.79	0.95	0.88–0.99
G5	CAP	0.55	0.99	0.93–1.00	0.97	0.86–1.00	1.00	0.92–1.00
	ENEA	0.60	0.87	0.81–0.95	0.87	0.72–0.96	0.87	0.79–0.98
	AlaSTAT	0.60	0.98	0.92–1.00	0.95	0.82–0.99	1.00	0.92–1.00
	CARLA	0.54	0.98	0.91–1.00	0.95	0.79–1.00	1.00	0.89–1.00

guarantee intermethods agreement. Other analytical variables (characteristics of the antigen, tracer, and solid phase) are crucial in determining the final result and, therefore, the comparability between different assay methods. In effect, some of the discrepancies observed might be expected in view of the different instructions given by the different manufacturers for the interpretation of results. In fact, CAP and AlaSTAT provide a similar relationship between quantitative results and the traditional class score, whereas the ENEA system declares the same values only for the first four classes. Finally, the CARLA system provides a different class score, the lowest threshold being 0.5 and 75 kU_a/L being the discriminating value for the sixth class.

Therefore, the agreement between the evaluated four second-generation techniques for IgE measurement is not completely satisfactory, and further efforts must be made to improve standardization in this field. To overcome these limitations in the clinical setting, we selected individual thresholds for each allergen and method, choosing those that provide the highest clinical efficiency on the basis of ROC analysis. Results were reevaluated using the selected thresholds, and the clinical efficiencies of the various assays for each allergen were compared with that of CAP. An overall agreement was found for most methods and allergens, whereas a significant discrepancy was found solely between CAP and ENEA for D1 and G5. For D1, the difference is explained by the significantly lower sensitivity that the ENEA method appeared to have (0.67 vs 0.97), and for G5 by the significantly lower specificity of G5 itself (0.87 vs 1.00).

On considering the AUC, which is a measurement of overall clinical accuracy, we found no significant differences between most of the evaluated methods for each allergen, confirming results reported elsewhere in studies comparing the clinical efficiency of the four methods (6, 9).

In conclusion, second-generation techniques standardized against the 75/502 WHO preparation only partly improve the intermethod agreement between different specific IgE assays. The differences observed for some allergens in the comparative analysis did not lead to a different clinical interpretation if a specific threshold was selected on the basis of ROC analysis.

Second-generation techniques have enabled us to obtain a better standardization of results; however, the identification of a specific threshold seems to be a prerequisite for the appropriate clinical interpretation of data. This information should therefore be given in the laboratory reports.

Additional studies should be undertaken for a better understanding of intermethod discrepancies and an improved standardization of all the analytical variables that affect the final results of the assays.

We thank Pharmacia & Upjohn Diagnostic, Radim, Bio-Allergy, and Medical Systems for the reagents used to perform the evaluation.

References

1. Wide L, Bennich H, Johansson SGO. Diagnosis of allergy by an in-vitro test for allergen antibodies. *Lancet* 1967;ii:1105-7.
2. Leimgruber A, Mosimann B, Claeys M, Seppey M, Jaccard Y, Aubert V, et al. Clinical evaluation of a new in-vitro assay for specific IgE, the immuno CAP system. *Clin Exp Allergy* 1991;21:127-31.
3. Lockey R, Lichtestein L, Bloch K, Kaliner M, Zweiman B, Rocheslesky G. Position statement. The use of in vitro tests for IgE antibody in the specific diagnosis of the IgE-mediated disorders and in the formulation of allergen immunotherapy. *J Allergy Clin Immunol* 1992;90:263-7.
4. Bousquet J, Chanez P, Chanal I, Michel FB. Comparison between RAST and Pharmacia CAP system: a new automated specific IgE assay. *J Allergy Clin Immunol* 1990;85:1039-43.
5. Reed CE, Yunginger JW, Evans R. Quality assurance and standardization of allergy extracts in allergy practice. *J Allergy Clin Immunol* 1989;84:4-7.
6. Plebani M, Borghesan F, Basso D, Faggian D. Receiver-operating characteristics (ROC) curves: a fundamental tool for improving the clinical usefulness of in vitro IgE tests. *Allergy* 1996;51:407-11.
7. Società Italiana di Allergologia e Immunologia Clinica. "Memorandum" sulla diagnostica delle allergopatie. *Folia Allergol Immunol Clin* 1987;34:387-91.
8. Bernstein L. Proceedings of the Task Force on Guidelines for Standardizing Old and New Technologies Used for the Diagnosis and Treatment of Allergic Diseases. *J Allergy Clin Immunol* 1988;82:487-526.
9. Plebani M, Borghesan F, Faggian D. Clinical efficiency of in vitro and in vivo tests for allergic diseases. *Ann Allergy* 1995;74:23-8.
10. Plebani M, Borghesan F, Bernardi D, Faggian D. Clinical evaluation of a new quantitative method for specific IgE antibodies. *Eur J Clin Chem Clin Biochem* 1996;34:579-84.
11. Banfi G, Ferrara F, Bonini P, Colombo G, Plebani M, Borghesan F, et al. Multicentre evaluation of Capture Assay Radim Liquid Allergen for measurement of specific IgE antibodies. *Eur J Clin Chem Clin Biochem* 1995;33:755-9.
12. Plebani M, Faggian D, Borghesan F. Full automation in allergy testing: measurement of specific IgE by the ENEA System. *Allergy* 1995;50:229-33.
13. Ito K, Miyamoto T, Makino S, Fukuda T, Kushima A, Kobayashi S. Usefulness of AlaSTAT. *Jpn J Allergol* 1991;40:1-9.
14. Passing H, Bablok W. A new biomedical procedure for testing the equality of measurements for two different analytical methods. *J Clin Chem Clin Biochem* 1983;21:709-20.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;ii:307-10.
16. Astute. Statistics Add-in for Microsoft Excel. DDU Software Ver. 1.5. Leeds, UK: The University of Leeds, Old Medical School, 1995.
17. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology* 1982;143:29-36.
18. Gardner HJ, Gardner SB, Winter PD. Confidence interval analysis micro-computer program. London: British Medical Journal, 1989:77 pp.
19. Pastorello EA, Incorvaia C, Ortolani C, Bonini S, Canonica CW, Romagnani S, et al. Studies on the relationship between the level of specific IgE antibodies and the clinical expression of allergy. I. Definition of levels distinguishing patients with symptomatic from patients with asymptomatic allergy to common aeroallergens. *J Allergy Clin Immunol* 1995;96:580-7.
20. Fraser CG, Fogarty Y. Interpreting laboratory results. Analytical and biological variation must be taken into account. *Br Med J* 1989;298:1659-60.