

Comparability and quality of IgE-based *in vitro* allergy diagnosis: 25 years of external quality assessment

Lukas Koch · Werner Aberer

Received: 6 November 2013 / Accepted: 9 August 2014 / Published online: 6 September 2014
© Springer-Verlag Wien 2014

Summary A large variety of methods and a wide range of products from several manufacturers are available for IgE-based *in vitro* allergy testing. In the large majority of cases, the attending physician lacks information about the devices and products used in the laboratory performing the tests. To determine the comparability and reliability of different laboratories, we analyzed the findings of an international ring trial established in Austria 25 years ago.

The results were satisfactory: 97.4 % of total IgE tests and 93.9 % of antigen-specific IgE tests were correct. Further analysis revealed more false-positive than false-negative antigen-specific IgE levels (17.1 vs. 5.2 %). The investigation of different allergen groups yielded excellent results for some allergens (97.8 % correct values for mites, or 99.6 % correct values for recombinant antigens), but quite unfavorable for others (21.4 % incorrect values for drugs). In view of this paucity of comparability and reliability, it would be appropriate to establish mandatory quality control in laboratories and manufacturing units.

Keywords In-vitro allergy testing · IgE · Antigen-specific IgE · Comparability · Quality · External quality assessment

Vergleichbarkeit und Qualität IgE-basierter *in-vitro* Allergiediagnostik – 25 Jahre Ringversuch der ÖGAI

Zusammenfassung Im Rahmen der IgE-basierten *in-vitro* Allergiediagnostik stellt sich aufgrund der großen Zahl an technischen Möglichkeiten, anbietenden Firmen und durchführenden Laboratorien für den zuweisenden Arzt die Frage nach der Vergleichbarkeit und Verlässlichkeit der Resultate einzelner Allergene und Allergengruppen. In der vorliegenden Arbeit werden diese Eigenschaften mit Hilfe eines seit 25 Jahren etablierten Ringversuchs untersucht. Es wurden dabei 506 Serumproben an Laboratorien in Österreich und seinen Nachbarländern versandt. Die Resultate waren größtenteils zufriedenstellend: Insgesamt waren 97,4 % der Gesamt-IgE-Bestimmungen und 93,9 % der antigenspezifischen IgE-Bestimmungen als korrekt zu werten, mittels Aufschlüsselung der inkorrekten antigenspezifischen IgE-Werte waren mehr falsch-positive als falsch-negative Werte feststellbar (17,1 % vs. 5,2 %). Während eine Analyse der einzelnen Allergengruppen zum Teil hervorragende Resultate lieferte (z. B. Milben: 97,8 % der Bestimmungen korrekt, rekombinante Antigene 99,6 % korrekt), waren die Ergebnisse bei anderen sehr ernüchternd (z. B. Medikamente: 21,4 % inkorrekte Ergebnisse). Eine Vergleichbarkeit der Befunde zwischen verschiedenen Laboratorien war häufig, aber bei weitem nicht immer gegeben. Der Forderung einer in vielen Ländern üblichen verpflichtenden Qualitätssicherung sowohl für Firmen als auch Laboratorien wurde trotz mehrmaliger Anregung bis jetzt nicht Folge geleistet.

Schlüsselwörter *In-vitro* Allergiediagnostik · IgE · Antigen-spezifisches IgE · Vergleichbarkeit · Qualität · Ringversuch

Univ.-Prof. Dr. W. Aberer (✉) · L. Koch
Department of Dermatology, Medical University of Graz,
Auenbruggerplatz 8,
8036 Graz, Austria
e-mail: werner.aberer@medunigraz.at

Introduction

Since the discovery of immunoglobulin E (IgE) antibodies and their special significance in allergology by K. Ishizaka in 1966 and S.G.O. Johansson in 1967, IgE-based *in vitro* allergy diagnosis—in addition to the patient's medical history, skin and provocation testing—is a valuable part of the investigation of any allergy [1]. A large variety of technical detection methods are available for this purpose. Numerous companies offer suitable products. The indications are manifold: while total IgE is specifically used in allergology to detect the risk of atopy, the determination of antigen-specific IgE values is used to specify causes for allergic asthma, rhinoconjunctivitis, urticaria, drug and food allergies, and even anaphylaxis after an insect bite. Especially after developing dangerous reactions, the patients individual test result may have far-reaching consequences [2]. How reliable is the report presented to the treating physician in the individual case? The physician ordering the test has a large number of laboratories at his/her disposal. Do all of these laboratories provide the same results or does the physician's choice influence the outcome of the test? For what allergens or allergen groups is *in vitro* diagnostic investigation a reliable aid in current times and in what cases should one perform further tests? In the present investigation we try, on the basis of an external quality assessment established 25 years ago, to provide answers to these questions. In between analyses had revealed quite heterogeneous results [2].

Material and method

Five-hundred and six serum samples of allergy patients were collected from 1988 [3]. Four times per year, samples were sent to laboratories specialized in allergy diagnosis within Austria and in neighboring western countries ($n=360$, phase I). From 2006 onward, samples were also sent to laboratories in eastern and south-eastern neighboring countries ($n=146$, phase II). Every consignment included, on average, five samples to determine total IgE levels as well as antigen-specific IgE antibodies. The latter required, in most cases, the determination of three different antigens per sample. The number of participating laboratories varied over time. On average 23 laboratories tested for total IgE (minimum 11 and maximum 35) and 26 for antigen-specific IgE (minimum 12 and maximum 44).

The laboratories participating in the study were asked to treat the received serum samples as routine samples, perform the tests by the standard method used in the respective laboratory, and send the results as early as possible. Total IgE values were asked to be given in kU/l and antigen-specific values in RAST classes.

Valid samples

Of 506 dispatched serum samples for the measurement of total IgE, 480 were analyzed. The main reason for not interpreting the remaining 16 was the presence of more than fifty percent relative values (e.g. >1.000 kU/l) per sample. Of 11,386 potential values, 11,276 were available (response rate 99.0%) while 558 (5.0%) of the measurements were considered invalid because of relative values. 10,718 test outcomes were used for the evaluation.

Of the 1478 requested tests for antigen-specific antibodies, ten were invalid. The main reasons were usually irrational values or the statement of "positive" instead of a definitive RAST class. 1345 of the tests were positive sera (anticipated result with positive evidence of antibodies) while the remaining 123 were negative sera (anticipated result with no evidence of antibodies). Of 38,965 potential test outcomes, 31,008 were sent back, signifying a response rate of 79.6%. Of 30,919 outcomes, only 89 (0.3%) were invalid.

Evaluation

The data reported back were evaluated by the sender and the overall results were sent back to the participating laboratories in anonymized form. For further work-up of the data, these were summarized into one list for overall IgE and another list for antigen-specific IgE testing.

With regard to total IgE, the mean value was calculated on the basis of $\bar{X} = \sum X_i / n$ and standard deviation, using the formula $\hat{\sigma} = \sqrt{(\sum (v_i - \text{mean value})^2) / (n-1)}$. Means $\pm$ standard deviation were deemed absolutely correct. Values that deviated from means by more than three standard deviations were considered unquestionably false. A variation coefficient was determined for every consignment using the formula $\text{VarK}(X) = \text{standard deviation}(X) / \text{mean value}(X)$.

To evaluate antigen-specific values, the most frequently determined RAST class was rated definitely correct. Positive sera with a deviation >2 RAST classes or a negative value were deemed false. For negative sera, on the other hand, any positive value was deemed false. The definitely correct RAST class varied according to the heterogeneity of the samples sent for testing. Two adjacent RAST classes were equally common in 20 cases; hence both were deemed correct.

Methods and products of companies

Evaluation of the methods used and the products of companies revealed a dominance of ThermoFisher Scientific Company (previously Phadia or Pharmacia). With regard to overall IgE, 64.1% of the samples were tested with reagents of this manufacturer. At the start of the extended consignments (consignment of March 2006) the figures were 68.6% in "old" laboratories compared to 0%

in “new” ones. For antigen-specific IgE the total figures were 69.2 % in old laboratories versus 74.7 % in new ones.

Results

Total IgE

Of 10,718 valid values, 70.8 % were below one standard deviation and were therefore deemed definitely correct. A further 22.5 and 4.2 % were within two and three standard deviations, respectively, while 2.6 % of the values were beyond three standard deviations and therefore deemed false. Evaluation based on the year of dispatch revealed rather similar results over the years, although more numerous (6.6 %) false results were noted from 1992–1997 (Fig. 1). From 2006 onward, the serum samples were sent to a larger circle of laboratories. The new laboratories yielded more false values (1.7 % in the old laboratories versus 3.8 % in the new ones).

To exemplify the deviation of individual test results between the laboratories over time, independent of the IgE concentrations of the respective serum sample, a coefficient of variation was calculated for each consignment. This showed, in addition to a progressive reduction, a consistently lower fluctuation between the individual consignments. The changes were rather low especially from the turn of the century onward (Fig. 2).

Antigen-specific IgE

The overall evaluation of 30,919 antigen-specific IgE tests showed 6.1 % to be false while 74.9 % of the results were definitely correct; 17.1 and 1.9 % showed a deviation of one and two RAST classes, respectively. Compared with total IgE, much greater fluctuations were noted in the overall period of observation. In some years, about 10 % of the values were false (Fig. 3). Comparing phase I and II, again much fewer false results were noted in the “old”

laboratories (5.1 versus 13.3 %). Division of the results into positive and negative sera revealed markedly more (17.1 %) false results in negative sera compared to positive sera (5.2 %).

To analyze the entirety of tested allergens, in a first step the allergens were divided into the following groups: pollen ($n=10,298$), foodstuffs ($n=6,111$), insect venom ($n=3,944$), animal epithelia ($n=3,814$), mites ($n=3,070$), drugs ($n=1,819$), mold ($n=1,177$), environmental substances ($n=650$) and others ($n=36$). Figure 4 illustrates the percentages of false results in the respective allergen groups. Overall data as well as data for positive and negative sera are given. For comparison, we performed a total evaluation ($n=30,919$) as well as a calculation of all recombinant allergens independent of the allergen group ($n=221$). Except for environmental substances and the evaluation of all recombinant allergens, negative sera always fared much worse than positive sera. Thus, in nearly all allergen groups we registered more false-positive results than false-negative ones.

In a second step, all 141 tested antigens were analyzed individually. Owing to the diversity of antigens, the case numbers for some were rather low and the validity of the results was therefore limited. In the following, we will present a few representatives (of each allergen group) with larger case numbers. Recombinant antigens are not taken into account.

Mites

Of all allergen groups, Dermatophagoides yielded the best results in the overall assessment as well as for positive and negative sera. Nevertheless, differences in species were noted: a false rate of 1.2 % was registered for *D. farinae*, compared with 2.4 % for *D. pteronyssinus*.

Pollen

In addition to the fact that negative sera fared much worse than positive sera, the subgroup assessment revealed

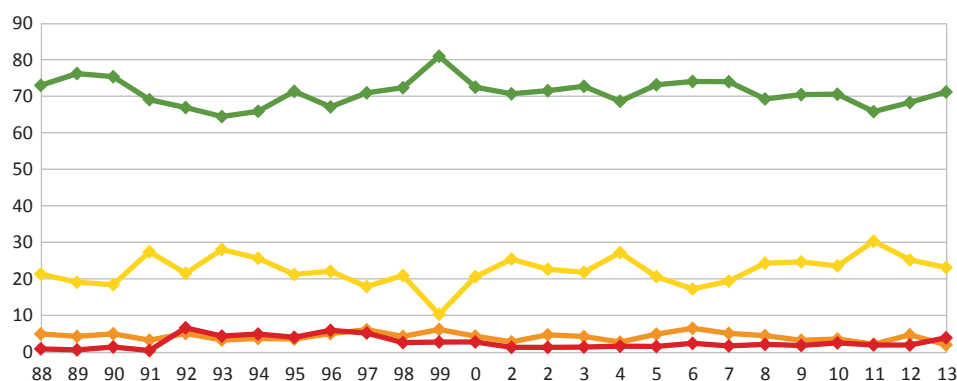


Fig. 1 Results of the evaluation of total IgE from the year 1988 to 2013. Results of the evaluation of total IgE from 1988 to 2013 are shown in percentage points. Green values were within one standard deviation from mean, yellow values between

one and two standard deviations from mean, and orange values between two and three standard deviations from mean. Red values were beyond three standard deviations and were therefore regarded as false

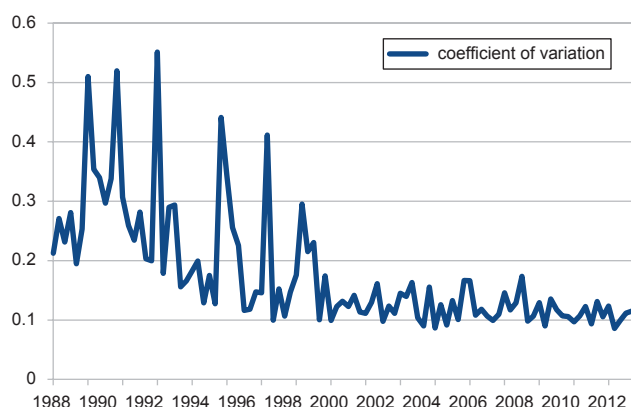


Fig. 2 Coefficient of variation of total IgE measurements from the year 1988 to 2013. The coefficient of variation of total IgE tests from 1988 to 2013 fell over the years. Fluctuations in individual values were also reduced

Fig. 3 Results of antigen-specific IgE evaluation from the year 1988 to 2013. Results of antigen-specific IgE evaluation from 1988 to 2013 are given in percentages. Green values indicate the correct RAST class while yellow and orange values deviated by one and two classes, respectively. Red values were incorrect

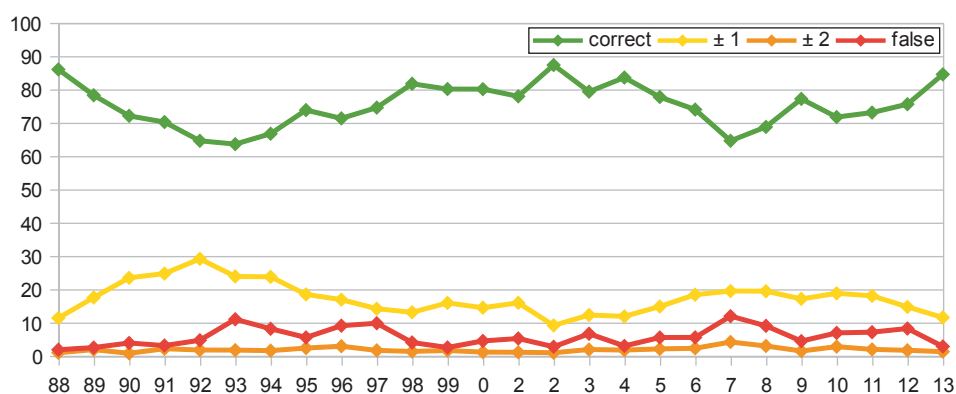
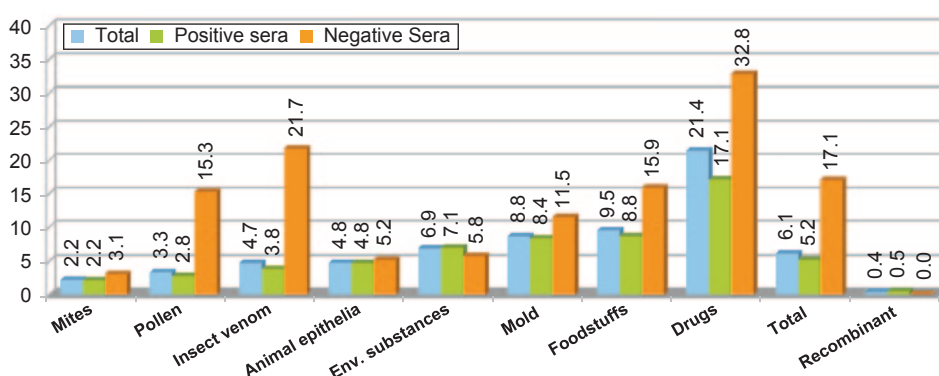


Fig. 4 Incorrect results according to antigen groups. Incorrect results for antigen-specific IgE testing, divided into allergen groups and given in percentages (y-axis). Blue columns indicate the overall evaluation; green and orange columns express the analysis of positive and negative sera



better results for grasses as opposed to tree or herb allergens (Table 1). The individual allergens of the respective groups usually yielded less than 4% false reports (e.g. 2.4% for timothy grass, 3.2% for rye, 3.2% for birch, 3.6% for mugwort, and 5.5% for ash). Thus, the test outcomes were largely correct. Only the determination of antibodies to ragweed was associated with much more numerous false results (8.4%).

Insect venom

The difference between positive and negative sera was even more pronounced for insect venoms: positive sera were associated with false-negative results in a mere 3.8% of cases (which was a favorable outcome), but at least every fifth test of negative serum was false positive (21.7%) [5]. At the level of individual allergens, very diverse results were obtained for the wasp, the bee, and the hornet (2.8, 4.9 and 10.9% false results, respectively).

Animal epithelia

Despite slight differences in cat, dog and horse dandruff, for which false results were obtained in 3.0, 4.2 and 5.5%, respectively, this group achieved average and, in respect of positive and negative sera, very homogeneous results.

Environmental substances

Environmental substances also yielded average values. The only group that fared a little better was negative sera as compared to positive sera. Notably, antibody determinations for latex yielded false results in 5% of cases.

Table 1 Specific IgE antibodies to pollen allergen groups

	Total sera	Positive sera	Negative sera
Pollen	3.33	2.77	15.30
Grasses	2.46	2.05	10.42
Trees	3.11	2.61	20.79
Herbs	4.82	3.98	17.54
False results in the determination of pollen allergens (total) and their subgroups are given in percentages. Qualitative differences were noted between grass, tree and herb pollen. Positive sera yielded much better results than negative sera			

Mold

Below-average values were obtained for molds. Slight differences in species were noted, for instance, between *Alternaria alternata*, *Aspergillus fumigatus* and *Cladosporium herbarum* (7.9, 9.0 and 11.2% false results, respectively).

Foodstuffs

Food allergens yielded somewhat poorer results [6]. Particularly negative sera yielded false-positive reports in 15.9 % of cases. Table 2 shows the data for the tested subgroups. The differences between the individual allergens were nearly as heterogeneous as the allergens themselves. Table 3 summarizes those allergens that were tested in more than 200 cases. Compared with favorable results for tomato and soybean, many false reports were obtained for substances like hazelnut or green apples.

Drugs

Drugs yielded the poorest results (incorrect values in 21.4 % of cases). Positive sera fared much better in that they were associated with a false-negative rate of 17.1 %, whereas negative sera were associated with a false-positive rate of 32.8 %. Sufficiently large case numbers were available only for penicilloyl G and V. In the overall assessment, 20.6 and 21.1 % of the tests were incorrect.

Discussion

The rather high quality of total-IgE results and the very high comparability of the individual laboratories are worthy of mention. Given a false rate of 2.6 %, only a little more than every 40th test was incorrect. Remarkably, the coefficient of variation reduced consistently over the entire time period and there was a lower fluctuation between the individual consignments. This is indicative of the stability of reagents and devices, as well as the increasing homogenization of values in the individual laboratories, especially from the turn of the century onward. The fact that the “old” laboratories fared better on comparison can be explained by the marked dominance of ThermoFisher Scientific Company (previ-

Table 2 Specific IgE antibodies to food allergen groups

	Total sera	Positive sera	Negative sera
Food	9.51	8.80	15.93
Egg/poultry	8.04	4.31	18.18
Seeds/legumes/nuts	8.98	8.61	13.62
Milk	10.17	8.81	16.08
Fruit/vegetables	10.74	10.69	13.64
Fish/crustaceans	11.66	11.66	–
False results in the analysis of food allergens (total) and in subgroups are given in percentages			

Table 3 Specific IgE antibodies to individual food allergens

	% false	n=
Food	9.51	6111
Tomato	2.70	222
Soybean	3.25	246
Rye flour	5.94	219
Peanut	5.98	669
Egg white/egg protein	8.04	858
Milk protein	9.49	643
Celery	9.51	347
Wheat flour	9.66	580
Hazelnut	14.03	677
Green apple	17.14	315
False results for the individual food allergens with case numbers > also results for the individ		

ously Phadia). This “quantity effect” has been explained elsewhere and was also true for antigen-specific IgE [3].

A summary of antigen-specific determination of IgE also reveals rather good agreement between the individual laboratories; on international comparison we found similar results [7]. At a false rate of 6.1 %, one may conclude that every 16th test result was incorrect. The difference between positive and negative sera was rather pronounced in some allergen groups: while positive sera were associated with a false-negative rate of 5.2 % (every 19th test was false negative), negative sera were associated with a false-positive rate of 17.1 % (every 6th antigen test was false positive). This is satisfactory to the extent that, in cases of positive serum, the determination of antigen-specific IgE antibodies is highly reliable for the detection of sensitization, but a positive test result is also very likely to be obtained in negative serum and may lure the inattentive interpreter in the wrong direction—with the inevitable false consequences.

Given the strong propagation of food allergies by the media and the rather lucrative screening investigation provided by manufacturers, this observation might lead to unnecessary diets and impairment of quality of life or even malnutrition—especially in the case of food allergens—in the presence of atypical or even missing clinical symptoms. In a screening investigation, given an error rate of 15.9 % registered by us, even in the absence of sen-

sensitization on the *in vitro* test every sixth investigated food allergen would test false positive. Therefore, it would be highly inadvisable to perform such screening investigations in the absence of explicit clinical suspicion [6].

The large number of false results in the investigation of insect venom or drug allergies (especially penicillin allergy) is very problematic because the current guidelines recommend first-line use of *in vitro* diagnostic testing in these cases due to the potential risk of severe reactions on skin and provocation testing [8]. Every fourth test for insect venom (21.7%) and every third test for drugs (32.8%) was false positive in negative sera. Further skin and provocation testing performed as a result thereof is a mere waste in economic terms. It endangers the tested patient to a very marginal extent, except for potential sensitization or undesirable drug effects. The alarming fact, however, is that, given a false-negative rate of 17.1%, a little more than every 6th person with drug sensitization is not identified as such and would therefore be potentially subjected to further life-threatening tests or even uncontrolled exposure to the allergen [9].

Testing children for insect venom sensitization, which is vehemently propagated by some, makes the above very clear: sensitive tests demonstrate specific IgE antibodies to insect venom in the serum of about 50% of children [10]. However, the prevalence of systemic allergic reactions is only 0.15–0.8% [11]. Thus, providing advice about the consequences of these numerous false-positive *in vitro* reports will be quite difficult [5].

Our data showed better sensitivity (favorable results for positive sera) than specificity (poorer results for negative sera) in nearly all allergen groups. A general statement as to whether enhancing the sensitivity of a test allergen by enrichment or spiking (which is done increasingly often in recent times by some manufacturers) will be counter-productive rather than meaningful cannot be made at the present time. The addition of a recombinant minor determinant such as Hev b 5 to the extract of natural latex may well enhance the sensitivity of, and therefore improve, the diagnostic investigation of latex. “Spiking” hazelnut extract with the principal hazelnut allergen Cor a 1 permits better detection of birch-pollen-associated hazelnut sensitization, but the latter is clinically harmless in most cases. Thus, the greater sensitivity appears to be of doubtful clinical value.

The reasons for the large number of false-positive reports are manifold. One reason may be the very high sensitivity of the methods. The avoidance of false-negative results is achieved at the expense of a high rate of false-positive test outcomes [7, 12]. Further causes could be the unspecific binding ability in the presence of highly elevated total IgE levels, the structural predisposition of the IgE molecule for cross-reactivity, or similar amino acid sequences in the individual allergens [13]. At least for inhalants, food and insect venom allergens, cross-reacting carbohydrate determinants also appear to be a rather significant factor in causing cross-reactivity [14]. A method for effective suppression of these antibodies was published recently [15]. Whether routine use of this

method will yield more exact and more comparable data, especially on a cross-company and cross-laboratory basis, will have to be determined in future investigations. Although the diversity of error within laboratories would generally be of rather minor significance, this aspect should also be taken into account.

The reasons for false-negative reports are also manifold. While the sensitivity of individual products provided by the companies appears to be too low, the use of methods established by empirical experience alone appears to be quite important, especially in the investigation of drug allergens. Further causes include the sensitivity of the reagents, the potential damage of samples during transport and storage and, again, the error potential within laboratories.

In summary, an outstanding degree of comparability was noted for the diagnostic investigation of mites and very good results were obtained for the determination of positive sera for pollen or insect venom allergens. While animal epithelia fared markedly better than average on comparison, environmental substances and mold ranked below these. Poor agreement was noted between the individual laboratories and companies' products for foodstuffs and, in the case of negative serum, for inhalants and insect venom antigens. The determination of specific IgE for drug allergens was least satisfactory; this was especially true of penicillin. A similar sequence of inhalants, foodstuffs and betalactam antibiotics was noted by other authors as well [7]. This sequence provides a rough idea of the reliability of individual allergen groups. Of the percentages mentioned in Fig. 4, especially those relating to the overall evaluation are relevant because, obviously, the interpreter himself/herself does not know whether the report presented to him/her concerns positive or negative serum.

Some allergens were more comparable than others in nearly all allergen groups. For mites, *D. farinae* yielded more precise values than *D. pteronyssinus*. For pollen, the subgroup of grasses was more reliable than that of trees or herbs and with the exception of ragweed, the individual allergens yielded very satisfactory results. For insect venom, the wasp fared markedly better than the bee or the hornet. With regard to animal epithelia, the sequence was cat, dog and horse. Even for mold, we noted differences in species between *Alternaria alternata*, *Aspergillus fumigatus* and *Cladosporium herbarum*. The greatest fluctuations were observed in foodstuffs. While tomato and soybean yielded comparable results, those for hazelnut or green apple were quite disillusioning.

The high agreement in recombinant allergens was remarkable [16]. This was due to the absence of the above mentioned sources of error (such as cross-reacting carbohydrate determinants) on the one hand [15], and definitely also because nearly all of the employed reagents were derived from one company and testing was performed with the same or a similar type of device. This high degree of precision is also useful for the investigation of the oral allergy syndrome. One example is a peanut-positive serum (f13): in all return consignments

f13 was tested positive and the “dangerous” recombinant allergens Ara h1–3 were negative, but Ara h8, which is an indicator of the oral allergy syndrome, was positive. The patient had the exact medical history for this condition and the test results clearly confirmed the fact. The only limiting fact was that only 13 of 26 participating laboratories performed tests for recombinant allergens. Thus, one half of the laboratories could only detect peanut sensitization, but not its clinical intensity. Despite the introduction of recombinant allergens nearly two decades ago (end of 1995/beginning of 1996), only 49.9 % of possible recombinant antigen tests were performed. This is indicative of the very meager availability of this type of testing.

The availability of diagnostic tests was very limited for some conventional allergens as well. One example is the determination of antibodies to ash, whose significance as an independent allergen has risen just in the last few years. Given a mean false rate of 5.5 % in 9 consignments, the diagnosis of ash sensitization was by no means poor, but only 14 of 34 (41.2 %) laboratories performed the test at all. This number is surprising and also regrettable because it indicates that testing for ash is either not requested due to the fact that the allergy specialist does not take it into account, or the laboratory does not offer the test because those in charge do not consider it important.

With regard to all observations and especially percentages, the strength of our data, which is the large number of contributing laboratories and the diverse methods and products used, is simultaneously its greatest weakness. By no means would it be justified to accept these data on a one-to-one basis because the sender usually serves one laboratory and the samples are usually processed by the same method of one company. Nevertheless, they serve as guidance values in that they show the probability of false results. Based on our data we cannot conclude that one company offers better products than another [17]. The comparability of the individual laboratories and the methods used is dependent on the individual allergen and is rather high in the majority of cases. Nevertheless, the question as to whether the selection of the laboratory influences the outcome must definitely be answered in the affirmative.

The long-term goal of IgE-based *in vitro* allergy diagnosis is to issue recommendations about the reliability and clinical relevance of a negative or positive report for each allergen or recombinant antigen. The ideal situation would be a percentage probability of a clinical reaction, depending on the concentration of the specific IgE. Such determinations have been attempted for some foodstuffs [18]. However, the path leading to this goal is quite long. In addition to the elimination of all of the above mentioned factors that can be eliminated and hitherto unknown factors leading to false-negative and false-positive reports, the diagnostic methods and products used for this purpose must be rendered comparable because of their sheer diversity [1, 19]. This requires the introduction of cross-company standards for each allergen or recombinant antigen being determined. All exist-

ing and future methods, especially companies offering the methods, must adhere to these standards.

Finally, it should be stated that our appeal in previous reports on the results of ring trials [3, 4] to the effect that the data obtained by various methods should be rendered more comparable and quality assurance should be mandatory for manufacturers as well as laboratories, has not borne fruit. The consequence of this measure is obvious: the referring allergy specialist would have to know the strengths and weaknesses of the method used in “his/her laboratory”. However, this is usually not the case.

A rather depressing result of ring trials will be mentioned here but not elaborated: over the years, laboratories with poor results have withdrawn from voluntary quality control based on ring trials. Obligatory participation (such as the German INSTAND e.V., the British NEQAS or the Quality Club[®] established by ThermoFisher Scientific) or even the consequences of false results are not anticipated and not prescribed for *in vitro* diagnosis in allergology. Thus, low-quality companies as well as laboratories with poor quality control remain in the market and frequently provide incorrect results for allergic patients.

Acknowledgments

We thank all laboratories involved in the study for their many years of excellent cooperation, in particular: Allergieambulatorium Burgenland (J Mikschi), Allergieambulatorium Wien-Mitte (D Blagojevic), Allergieambulatorium Wien-West (F Horak), Allergielabor dermatologische Abteilung Wilhelminenspital Wien (W Jurecka), Allergielabor Univ.-Hautklinik Innsbruck (N Reider), Allergielabor LKH Villach (B Sterz), Allergieambulatorium Hall i.Tirol (H Kofler), Gruppenpraxis Labors. at Wien (K Bertl, M Exner), HYCOR Europe (S Staten), IASON GmbH (M Resch), Immunologisches Labor Univ.-Hautklinik Salzburg (G Pohla-Gubo), Labordiagnostik mit Blutbank, Wien (A Haushofer), Nuklearmedizin AKH Linz (M Gabriel), Nuklearmedizin und Endokrinologie Elisabethinen Linz (J Dierneder), Klinikum Bad Gleichenberg (M Trinker), Klinikum Klagenfurt-Wörthersee (P Sinha), Diagnostisches Labor Innsbruck (G Wick), Landes-Frauen-/Kinderklinik Linz (Schreier-Lechner), Labordiagnostik Klagenfurt (F Holzweber), MEDIWISS Analytic GmbH Deutschland (P von Wahl), Medizinisches Labor Klagenfurt (J Perné), Phadia Austria GmbH (P Forstenlechner), Zentrallabor Kardinal Schwarzenberg'sches Krankenhaus (B Auer), Zentrallabor SMZ Baumgartnerhöhe (W Hübl).

Conflict of interest

Lukas Koch and Werner Aberer declare that there is no conflict of interest.

References

1. Aberer W, Kränke B. Measurement of IgE antibodies using liquid allergens—an inter-method and inter-laboratory quality assessment. *Wien Klin Wochenschr.* 2002;144:929–37.
2. Hamilton RG, Williams PB. Specific IgE testing task force of the AAAAI, ACAAI. Human IgE antibody serology: a primer for the practicing North American allergist/immunologist. *J Allergy Clin Immunol.* 2010;126:33–8.
3. Aberer W, Kränke B, Hager A, Wick G. In vitro allergy testing needs better standardization—test results from different laboratories lack comparability mostly due to missing effective standards. *Int Arch Allergy Immunol.* 1995;108:82–8.
4. Aberer W, Ebner H, Horak F, et al. Qualitätskontrolle im Allergielabor—4 Jahre „Ringversuch Allergie“ in Österreich. *Wien Klin Wochenschr.* 1992;104:195–9.
5. Sturm GJ, Kranzelbinder B, Schuster C, et al. Sensitization to Hymenoptera venoms is common, but systemic sting reactions are rare. *J Allergy Clin Immunol.* 2014;133:1635–43.
6. Eigenmann PA, Oh JW, Beyer K. Diagnostic testing in the evaluation of food allergy. *Pediatr Clin North Am.* 2011;58:351–62.
7. Sanz ML, Prieto I, García BE, Oehling A. Diagnostic reliability considerations of specific IgE determination. *J Investig Allergol Clin Immunol.* 1996;6:152–61.
8. Renz H, Biedermann T, Bufer A, et al. In-vitro-Allergiediagnostik—Leitlinien der Deutschen Gesellschaft für Allergologie und Immunologie. *Allergo J.* 2010;19:110–28.
9. Thong BY, Mirakian R, Castells M, et al. A World Allergy Organization international survey on diagnostic procedures and therapies in drug allergy/hypersensitivity. *World Allergy Organ J.* 2011;4:257–70.
10. Rueff F, Vos B, Przybilla B. In-vitro-Diagnostik bei Insektengiftallergie. *Allergologie.* 2013;36:43–50.
11. Bilo BM, Bonifazi F. Epidemiology of insect venom anaphylaxis. *Curr Opin Allergy Clin Immunol.* 2008;8:330–7.
12. Kleine-Tebbe J, Fuchs T, Lepp U, et al. In-Vitro-Diagnostik von Nahrungsmittelallergien. *Allergo J.* 2001;10:333–9.
13. Aalberse RC, Akkerdaas J, van Ree R. Cross-reactivity of IgE antibodies to allergens. *Allergy.* 2001;56:478–90.
14. Altmann F. The role of protein glycosylation in allergy. *Int Arch Allergy Immunol.* 2007;142:99–115.
15. Aberer W, Holzweber F, Hemmer W, et al. Inhibition kreuzreaktiver Kohlenhydratdeterminanten (CCDs) erhöht die Treffsicherheit der In-vitro-Allergiediagnostik. *Allergologie.* 2014;2:46–54.
16. Canonica GW, Ansotegui IJ, Pawankar R, et al. A WAO-ARIA-GA²LEN consensus document on molecular-based allergy diagnostics. *World Allergy Organ J.* 2013;6:17.
17. 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.
18. Sampson HA. Utility of food-specific IgE concentrations in predicting symptomatic food allergy. *J Allergy Clin Immunol.* 2001;107:891–6.
19. Contin-Bordes C, Petersen A, Chahine I, et al. Comparison of ADVIA Centaur and Pharmacia UniCAP tests in the diagnosis of food allergy in children with atopic dermatitis. *Pediatr Allergy Immunol.* 2007;18:614–20.