Comparison of different diagnostic products for skin prick testing

**Summary**

**Background:** Different in vivo methods are used to quantify the amount of allergens in products for skin prick testing. It is unclear how this impacts on the correct diagnosis of allergies. **Aim of the Study:** We compared the allergenic potency of three commercial extracts for skin prick testing and evaluated batch-to-batch differences within each product. **Methods:** Patients with a mono-sensitization (specific IgE level > 0.70 KU/L, ImmunoCAP®, Phadia) to *Phleum pratense* (N=21), *Parietaria judaica* (N=20) or *Dermatophagoides pteronyssinus* (N=28) were evaluated by standard skin prick testing and with the end-point dilution technique using commercial products from Stallergenes (A)(Antony, France), Lofarma Allergeni (B)(Milan, Italy) and ALK Abellò (C)(Hoersholm, Denmark). Results were expressed as mean areas of the wheal (cut-off for positive reactions: 7 mm²). **Results:** With standard prick testing, the following differences in wheal areas were found: *Phleum*, C higher than B (p=0.0454); *Parietaria*, C higher than A (p=0.094); *Dermatophagoides*, C higher than A (p=0.021). With limiting dilution testing, the following differences in dilutions yielding positive skin prick tests were found: *Phleum*, C and B higher than A (p=0.0391 and 0.0039, respectively); *Dermatophagoides*, C higher than A and B (p=0.0010 and 0.0156, respectively). In the batch-to-batch comparison, mean differences between wheal areas of compared undiluted solutions did not significantly differ in any allergen tested, although in single cases large differences were observed. At the 1 to 64 dilution, agreement was significant only with *Dermatophagoides* from Manufacturer C (p= 0.262). At the 1 to 16 dilution, agreement was significant with *Phleum* from Manufacturer C (p=0.0116) and with *Dermatophagoides* from Manufacturer B and C (p=0.0239 and 0.0001, respectively). At the 1 to 4 dilution agreement was significant with *Dermatophagoides* from the three considered Manufacturers (p=0.0189, 0.0052 and 0.0077, respectively) and with *Phleum* from Manufacturer B and C (p=0.0336 and 0.0113, respectively). **Conclusion:** There are significant differences among commercially available diagnostic products for skin prick testing.

**Introduction**

Standardization of allergen extracts for diagnostic and therapeutic products is the object of intense efforts to gain a better evaluation and treatment of allergic individuals and to allow comparison of different clinical studies (1). Presently, different in vivo methods are used by each Manufacturer to quantify the amount of protein antigen.
in diagnostic products. The same methods are mostly, though not invariably used for therapeutic extracts standardization and involve the evaluation of the skin reactivity of each batch of a given product in a panel of individuals, which are allergic to proteins of that specific allergen extract, as assessed by the determination of specific IgE and by a consistent clinical history. The validity of these inclusion criteria is widely accepted. However, relevant differences may derive from specific problems such as i) the number of individuals included in the panel, which is serving as representative of the whole allergic population; ii) the level of sensitization of these same patients, and how it was established; iii) factors affecting the level of skin test reactivity to histamine, which in turn affect also reactivity to allergens, such as age (2–4), gender, ethnic origin (5–8), environmental exposure (5), specific patterns of sensitisation to allergens (9, 10) or skin prick test site (11).

The objective of the present study was the comparison of the potency and the batch-to-batch consistency of three commercially available allergen extracts for prick testing used in a real-life situation in a group of allergic patients with known sensitizations.

**Material and methods**

**Patients**

Male and female adults and children were recruited in four allergological Centres in Italy. Patients had a known sensitization to one of the following allergens: *Phleum pratense* (21 patients), *Parietaria judaica* (20 patients), *Dermatophagoides pteronyssinus* (28 patients). Subjects with multiple sensitizations to pollen allergens were not eligible, due to the possibility of cross-reaction of IgE specific to homologous allergen components (e.g. panallergens such as profilins or calcium binding proteins). The inclusion criteria were a specific IgE level > 0.70 KU/L (i.e. the cut-off for the second ImmunoCAP® class) (Phadia, Uppsala, Sweden) and a consistent clinical history (season of occurrence and circumstances when symptoms occurred). Patient characteristics are summarized in Table 1.

Subjects were evaluated in the following months, according to their sensitization, in order to minimize the effect of natural allergen exposure:

1) *Phleum*: December
2) *Parietaria*: December
3) *Dermatophagoides*: June

<table>
<thead>
<tr>
<th>Allergen</th>
<th>N=</th>
<th>Sex (M/F)</th>
<th>Age (range)</th>
<th>R</th>
<th>RC</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grass</td>
<td>21</td>
<td>10/11</td>
<td>35 (11–55)</td>
<td>8</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Pellitory</td>
<td>20</td>
<td>8/12</td>
<td>30 (21–54)</td>
<td>6</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>House Dust Mite</td>
<td>28</td>
<td>10/18</td>
<td>22 (14–60)</td>
<td>6</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>28/41</td>
<td>26 (11–60)</td>
<td>20</td>
<td>49</td>
<td>22</td>
</tr>
</tbody>
</table>

R: rhinitis; RC: rhinoconjunctivitis; A: asthma

The skin prick test was performed twice in each patient, using two different batches of the same extract, within a 15 to 30 days time lag. Patients did not take anti-histamines drugs in the previous 30 days.

**Skin testing**

The reagents for prick testing were transferred to 1 ml glass vials anonymously identified as A, B and C and delivered to the experimenters of the four participating allergological Centres. Corresponding Manufacturers were as follows: A: Stallergenes (Antony, France), B: Lofarma Allergeni, (Milan, Italy), C: ALK Abello (Hoersholm, Denmark). Solutions for prick test were purchased simultaneously, were maintained at 4°C and had the same shelf file when they were used (within their respective expiry dates).

From publicly available information, allergenic potency for standardized extracts is expressed in reagents from Manufacturer A as “reactive index” (IR) where 100 IR is the concentration inducing an average wheal of 7 mm in thirty sensitized patients. Products for skin prick testing contain 1000 IR.

Manufacturer B standardizes extracts in DBU (“Diagnostic Biological Units”), where one DBU is one hundredth of the potency of one extract inducing a wheal equal to that induced by histamine chloride 10 mg/ml. It is unclear how many patients were used to for this purpose. Products for skin prick testing contain 100 DBU. Manufacturer C measures allergen potency as Histamine Equivalent Units (HEP), where 10 HEP is defined as the allergen concentration, which in an end-point skin prick test on 30 sensitized individuals, is eliciting a wheal of the same surface as that elicited by 10 mg/ml histamine chloride. Products for skin prick testing contain 10 HEP.

The *Lymulus Amoebocyte Lysate* reaction test (Cambrex, Walkersville, MD, USA) was used to assess the presence of bacterial endotoxins in solutions used for prick testing.
Following Manufacturer’s instructions, each sample was preliminarily screened for product inhibition, which may render the test relatively insensitive to endotoxin. Samples scored free of product inhibition according to the kit criteria (geometric mean end-point of endotoxin in sample within 0.5–2 times the labelled lysate sensitivity). Readings for endotoxin contents scored in all samples ≤ 0.06 Endotoxin Units/ml.

A 20-min skin prick test was performed in an outpatient setting within Hospitals with an Emergency Room facility. A standard skin pricker was used (Allergy pricker, Bayer DHS, Milan, Italy) and a drop of each extract was applied to the volar surface of the forearm at each of the following five dilutions, extemporarily prepared in sterile saline using a 20-200 µl micropipette (Gilson, Middleton, WI, USA): 1/1, 1/4, 1/16, 1/64, 1/256, 1/1024.

A pre-marked transparent tape was used to maintain a 2-cm distance between pricking sites. In each subject, also the undiluted extract of each of the allergen preparations not used for end-point testing was included (specificity control), as well as histamine chloride (positive control) and saline (negative control). The surface of the wheal was evaluated by the transcription of the edges of the wheal on transparent planimetric paper. A positive prick corresponded to the lowest dilution yielding a wheal whose diameter was > 3 mm (i.e., the wheal surface was > 7 sq mm) (12).

Ethical issues

Informed consent was obtained from each patient or from the children’s parents. This project was revised and approved by the Ethic Committee of each participating Centre.

Statistical analysis

The undiluted extracts of the three Manufacturers were not expected to show any difference, when read as “positive” or “negative”, considering the clear-cut inclusion criteria. Thus, to detect more subtle differences both the evaluation of the area of the wheals generated by the undiluted extract and the highest dilution of each sample giving a positive prick test (i.e., a wheal > 3 mm in diameter) in the end-point dilution testing were used. The analysis of differences of these two read-outs, when comparing products of different Manufacturers, was performed with a Wilcoxon’s test for paired data, since distribution of values was not normal. Wilcoxon’s test was also used in the comparison of the wheals generated by the undiluted extracts of different batches of each Manufacturer to evaluate whether and to what extent inter-batch differences were significant. For batch-to-batch comparison tests, we used the analysis of concordance of the positive or negative results of end-point dilution scores at each given dilution. The consistency of results obtained within each pair of prick testing solution was evaluated by matching the numbers of positive and negative prick testing, taken as categorical variables. The k value of agreement (k equals “+1” when there is complete agreement) (13) was calculated at each dilution. The p value for significance of the agreement was calculated from the z value function, obtained by dividing k by the standard error. Values of p < 0.05 were considered significant. All statistical calculations were performed with the GraphPad software (San Diego, CA, USA).

Results

Potency comparison between different prick testing products

Prick testing with undiluted, ready-to-use products from Manufacturers A, B and C yielded significant differences in wheal areas only in the following cases: Phleum, C higher than B (p=0.0454); Parietaria, C higher than A (p=0.094); Dermatophagoides, C higher than A (p=0.021). Median values and inter-quartile ranges are indicated for each allergen in figure 1, top panels.

The average mean dilutions yielding an “above the threshold” (i.e., > 7 mm²) wheal reaction were significantly different between the products from the three considered manufactures in the following instances: Phleum, C and B higher than A (p=0.0391 and 0.0039, respectively); Dermatophagoides, C higher than A and B (p=0.0010 and 0.0156, respectively). Median values are indicated for each allergen in figure 1, bottom panels.

Potency comparison between two different batches of prick testing products

For comparison of the wheal areas obtained with two batches of the same undiluted, ready-to-use product, the difference between the wheal areas observed with the compared prick tests was plotted as a function of the average of the two measurements using a Bland-Altman plot (14). This allowed to appreciate the distribution of the differences in paired measurements of prick testing.
products from Manufacturer A, B and C (Figure 2, panels from top to bottom). The mean differences of wheal areas obtained with two batches of each product, indicated as A/A1, B/B1 and C/C1, are shown for each allergen in figure 2, where they appear to the right of the dotted lines drawn at the corresponding value. These differences were not significant in any of the inter-batch comparisons with any allergen.

The consistency of results obtained within each pair of prick testing solution was further evaluated by matching the numbers of positive and negative prick testing at the 1 to 64, 1 to 16 and 1 to 4 dilutions, taken as categorical variables. The k value of agreement showed a significant symmetry at the 1 to 64 dilution only with *Dermatophagoides* from Manufacturer C (p=0.262). At the 1 to 16 dilution, agreement was significant with *Phleum* from Manufacturer C (p=0.0116) and with *Dermatophagoides* from Manufacturer B and C (p=0.0239 and 0.0001, respectively). At the 1 to 4 dilution agreement was significant with *Dermatophagoides* from the three considered Manufacturers (p=0.0189, 0.0052 and 0.0077, respectively) and with *Phleum* from Manufacturers B and C (p=0.0336 and 0.0113, respectively). *Pellitory* did not yield consistent results in the inter-batch comparison analysis at any of the considered dilutions with any of the tested Manufacturers (Table 2).

**Discussion**

Allergen extracts have been used for the diagnosis and therapy of type I allergy for about a century. They are biological products of high complexity, making them prone to significant variability. Recognition of the importance of standardization of diagnostic and therapeutic extracts has steadily gained ground over the past decades (1). The driving force behind the efforts to improve standardization is both correct diagnosis and optimal safety of immunotherapy. The system of IgE based standardization for the determination of allergen potency is usually referred to as
biological standardization, which includes both in vivo assays, such as skin prick testing and in vitro competitive assays, such as RAST inhibition (1). Different Manufacturers use non-homogeneous in vivo biological standardizations criteria, aimed to warrant a good specificity and sensitivity as well as consistency of results from batch to batch, for each allergen. We asked whether, and to what extent, three different diagnostic products of common clinical usage in Italy may differ for allergen potency and for batch-to-batch consistency of results.

Firstly, we need to state clearly that patients with known sensitization to common airborne allergens were correctly diagnosed with any of three compared products for skin prick testing. Nevertheless, significant differences were observed when more subtle parameters were measured, such as the area of the wheal elicited by each ready-to-use prick test solution or the highest dilution determining an above-the-threshold wheal area (Figure 1). Notably, differences between Manufacturers appeared distributed differently for different allergens. This is not surprising, since it is easily conceivable how a given Manufacturer-defined process of standardisation may differently affect extracts of different origins, tested in allergic subjects with different sensitization profiles.

We also considered batch-to-batch reproducibility of allergenic potency both by comparing the undiluted products and by considering the extreme conditions posed by the end point dilution technique. Although the comparison analysis directly performed on the prick areas obtained with undiluted products did not show significant differences between batches of any of the tested Manufacturers, it is interesting to observe how in the case of several patients a high variability was found (see dots more distant from the x-axis in figure 2). Moreover, as in the case of the sensitivity study, also in batch-to-batch consistency analysis when comparison of prick areas was performed with diluted products, significant differences were found depending both on the allergen and on the Manufacturer. In particular, the pellitory extract from any of the three considered Manufacturers did not yield consistent results at any of the considered dilutions. In contrast, the mite extract from the three Producers was satisfactorily concordant at
Table 2 - Results of skin prick testing at 1/64, 1/16 and 1/4 dilution, with the indicated allergen extract

<table>
<thead>
<tr>
<th>Grass</th>
<th>1 to 64</th>
<th>1 to 16</th>
<th>1 to 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
</tr>
<tr>
<td>A1 POS</td>
<td>1</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>A1 NEG</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>1 to 64 B POS</td>
<td>B POS</td>
<td>B POS</td>
<td>B POS</td>
</tr>
<tr>
<td>B1 POS</td>
<td>3</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>B1 NEG</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1 to 64 C POS</td>
<td>C POS</td>
<td>C POS</td>
<td>C POS</td>
</tr>
<tr>
<td>C1 POS</td>
<td>15</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>C1 NEG</td>
<td>3</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>1 to 16 A POS</td>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
</tr>
<tr>
<td>A1 POS</td>
<td>8</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>A1 NEG</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1 to 16 B POS</td>
<td>B POS</td>
<td>B POS</td>
<td>B POS</td>
</tr>
<tr>
<td>B1 POS</td>
<td>13</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>B1 NEG</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1 to 16 C POS</td>
<td>C POS</td>
<td>C POS</td>
<td>C POS</td>
</tr>
<tr>
<td>C1 POS</td>
<td>14</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>C1 NEG</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1 to 4 A POS</td>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
</tr>
<tr>
<td>A1 POS</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>A1 NEG</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1 to 4 B POS</td>
<td>B POS</td>
<td>B POS</td>
<td>B POS</td>
</tr>
<tr>
<td>B1 POS</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>B1 NEG</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1 to 4 C POS</td>
<td>C POS</td>
<td>C POS</td>
<td>C POS</td>
</tr>
<tr>
<td>C1 POS</td>
<td>17</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>C1 NEG</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pellitory</th>
<th>1 to 64</th>
<th>1 to 16</th>
<th>1 to 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
</tr>
<tr>
<td>A1 POS</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>A1 NEG</td>
<td>6</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>1 to 64 B POS</td>
<td>B POS</td>
<td>B POS</td>
<td>B POS</td>
</tr>
<tr>
<td>B1 POS</td>
<td>10</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>B1 NEG</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1 to 64 C POS</td>
<td>C POS</td>
<td>C POS</td>
<td>C POS</td>
</tr>
<tr>
<td>C1 POS</td>
<td>16</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>C1 NEG</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1 to 16 A POS</td>
<td>A POS</td>
<td>A POS</td>
<td>A POS</td>
</tr>
<tr>
<td>A1 POS</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>A1 NEG</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1 to 16 B POS</td>
<td>B POS</td>
<td>B POS</td>
<td>B POS</td>
</tr>
<tr>
<td>B1 POS</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>B1 NEG</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 to 16 C POS</td>
<td>C POS</td>
<td>C POS</td>
<td>C POS</td>
</tr>
<tr>
<td>C1 POS</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>C1 NEG</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

(continues)
the 1 to 4 dilutions, a condition which was lost by the extract from Producer A at the 1 to 16 dilution and by the extracts from Producer A and B at the 1 to 64 dilution. As far the grass allergen is concerned, prick test solutions from Producers A and B displayed a satisfactory inter-batch concordance at the 1 to 4 dilution, which remained to Producer C solution at the 1 to 16 dilution but was lost by all products at the 1 to 64 dilution.

Our data show that the three studied products are fully respectful of EU regulations, which declare that the maximal

<table>
<thead>
<tr>
<th>Table 2 – (continues)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 4 A POS</td>
</tr>
<tr>
<td>A1 POS</td>
</tr>
<tr>
<td>A1 NEG</td>
</tr>
<tr>
<td>1 to 4 B POS</td>
</tr>
<tr>
<td>B1 POS</td>
</tr>
<tr>
<td>B1 NEG</td>
</tr>
<tr>
<td>1 to 4 C POS</td>
</tr>
<tr>
<td>C1 POS</td>
</tr>
<tr>
<td>C1 NEG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>House dust mite</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 64 A POS</td>
</tr>
<tr>
<td>A1 POS</td>
</tr>
<tr>
<td>A1 NEG</td>
</tr>
<tr>
<td>1 to 64 B POS</td>
</tr>
<tr>
<td>B1 POS</td>
</tr>
<tr>
<td>B1 NEG</td>
</tr>
<tr>
<td>1 to 64 C POS</td>
</tr>
<tr>
<td>C1 POS</td>
</tr>
<tr>
<td>C1 NEG</td>
</tr>
<tr>
<td>1 to 16 A POS</td>
</tr>
<tr>
<td>A1 POS</td>
</tr>
<tr>
<td>A1 NEG</td>
</tr>
<tr>
<td>1 to 16 B POS</td>
</tr>
<tr>
<td>B1 POS</td>
</tr>
<tr>
<td>B1 NEG</td>
</tr>
<tr>
<td>1 to 16 C POS</td>
</tr>
<tr>
<td>C1 POS</td>
</tr>
<tr>
<td>C1 NEG</td>
</tr>
<tr>
<td>1 to 4 A POS</td>
</tr>
<tr>
<td>A1 POS</td>
</tr>
<tr>
<td>A1 NEG</td>
</tr>
<tr>
<td>1 to 4 B POS</td>
</tr>
<tr>
<td>B1 POS</td>
</tr>
<tr>
<td>B1 NEG</td>
</tr>
<tr>
<td>1 to 4 C POS</td>
</tr>
<tr>
<td>C1 POS</td>
</tr>
<tr>
<td>C1 NEG</td>
</tr>
</tbody>
</table>

The value of agreement (k) and the results of the McNemar’s test of symmetry (p) are shown for each experimental condition.
permitted potency variation between batches of allergen products is 50-200% of the stated amount (15, 16). Moreover, patients included in this study had a specific IgE level > 0.70 KU/l, a titre which is reasonably including most allergic patients. Thus, it is relatively unlikely that in a real life setting a lower level of sensitization would have established a more challenging situation. Nevertheless, our data are in agreement with previous reports where significantly different potencies were found in extracts of nut allergens from different Manufacturers (17, 18)
The explanation of the differences we found is speculative, since we have incomplete information on the methods used by each Manufacturer for establishing allergen potency, namely the number of individuals included in the panel, which is serving as representative of the whole allergic population, their specific IgE levels and how they were measured. Moreover, it is well established that skin test reactivity to histamine in single individuals significantly affect reactivity to allergens, and is influenced by factors such as age (2-4), gender, ethnic origin (5-8), environmental exposure (5) or specific patterns of sensitisation (9, 10).

In this scenario, measuring the content in major allergens may help to improve the reliability of standardization methods. However, since major allergens are recognized by at least 50% of the population sensitized to that given allergen (19), the presence of minor allergens, which a consistent proportion of patients may be sensitized to, is not accounted for by this method. Moreover, there are still no validated assays available to unambiguously quantify each single major allergen with immuno–enzymatic assays, HPLC or mass spectrometry (1), although certified reference materials have been made available for the most relevant allergen sources (20-25).

Our results highlight the timeliness of a coordinated effort, such as the one supported by the CREATE project (26), aimed to implement a synergistic approach, where methods based on reference materials are integrated in biological standardization assays, including in vitro RAST inhibition and mediator release assays (27) and in vivo skin prick tests with purified allergen molecules.

Acknowledgements

This work was sponsored by ALK Abellò (Hoersholm, Denmark).

References

Prick test potency


