Comparison of two epidemiological protocols for measuring airway responsiveness and allergic sensitivity in adults


ABSTRACT: In recent years, airway responsiveness has commonly been measured in epidemiological studies using one of two methods. In one method, histamine is administered via a handheld DeVilbiss nebulizer and in the other, methacholine is administered via a dosimeter. Allergic sensitivity has commonly been measured by either the allergen droplet method or by Phazet. We wanted to assess the comparability of airway responsiveness and of allergic sensitivity measured by both methods.

A total of 48 volunteers, including normal and asthmatic subjects, participated in the study. Subjects first underwent one of the two tests of airway responsiveness and allergic sensitivity, and then returned within 10 days to undergo tests using the second protocol. Commencement protocol was allocated in random order.

There was good agreement between both methods for assessing airway responsiveness and for assessing allergic sensitivity. The difference for dose response ratio (DRR) between histamine and methacholine was a 1.19 (95% CI 0.78, 1.82) fold changes, which was not statistically significant. Agreement between allergic sensitivity methods was perfect for Alternaria tenuis, good for rye-grass (kappa=0.71) and moderate for cat and Dermatophagoides pteronyssinus (kappa approximately 0.5). It is possible to compare data from epidemiological studies which use these methods.


Although a great deal of effort has been given to standardizing questionnaires used in the assessment of asthma in communities [1], a wide variety of objective methods are currently used to measure airway responsiveness and allergic sensitivity. Epidemiological protocols vary in their use of methods for measuring airway responsiveness and allergic sensitivity.

Many studies, in several countries, have used a protocol which measures airway responsiveness using histamine with the rapid method [2–5]. In adults, histamine and methacholine appear to be equipotent in field studies [6], and in asthmatics tested in the clinic [7], but not in patients with chronic obstructive lung disease [8]. However, a population study of children showed that asthmatics were slightly more sensitive to histamine than to methacholine [9], so that results using different agents may not be directly comparable. In addition to measuring airway responsiveness, allergic sensitivity has commonly been measured using the method of Pervs [10], in which droplets of selected allergens are applied to the forearm and the skin pricked through the drop with a lancet.

A major screening of populations, the European Commission (EC) Respiratory Health Survey [11] has used a protocol which measures airway responsiveness using methacholine administered by dosimeter. This method has the advantage that higher doses can be given with fewer side-effects [12]. In this protocol, allergic sensitivity has been measured using allergen coated lancets (Phazet) [13], which are thought to more accurately standardize the dose of allergen that is delivered under the skin.

In order to determine the comparability of the most widely used epidemiological protocols to measure both airway responsiveness and allergic sensitivity in epidemiological studies, we conducted a study of adults using each of two methods for each measurement within a 10 day interval. In this paper, we report results from the two different methods for assessing airway responsiveness (histamine by rapid method versus methacholine by dosimeter) and allergic sensitivity (allergen droplet versus Phazet).

Methods

Population

In September 1991, and in March 1992, both asthmatic and nonasthmatic research volunteers in Wellington,
New Zealand were invited to join a study to compare two different protocols for measuring asthma and allergy. Subjects were contacted by telephone, at which time the study design was explained and an appointment for testing arranged. Because the data were paired, it was calculated that a minimum of 33 subjects were required to detect a significant difference in dose-response ratio (DRR) between protocols at the p<0.05 level (power 80%), or 44 subjects at the p<0.01 level. A total of 48 subjects consented to be involved in the study. Subjects first underwent one of the two tests of airway responsiveness (histamine delivered by the rapid method or methacholine delivered by dosimeter) and allergic sensitivity (skin-prick test to either allergen droplets or Phazets), then returned within 10 days to have similar tests using the second protocol. Commencement protocol was allocated in random order. These studies were approved by the Wellington Area Health Board Ethics Committee.

Airway responsiveness

**Histamine airway challenge.** Lung function was measured with the subject standing using a Mijnhardt dry rolling seal spirometer connected to a lap-top computer running Scientific and Medical software (S&M Instrument Co., Doylestown, PA, USA) for immediate data acquisition. Subjects were asked to repeat forced expiratory manoeuvres until two readings of forced expiratory volume in one second (FEV1) within 100 ml of each other were obtained, of which the larger value was used in analyses. Measurements were corrected to body temperature and pressure, saturated with water vapour (BTPS). Airway responsiveness was then assessed using the rapid method [14] with histamine diphosphate, which was administered by means of a DeVilbiss No. 45 handheld nebulizer in cumulative doses ranging 0.03–3.9 µmol. Subjects were asked not to smoke in the hour before testing and to avoid taking asthma medications for 6 h before testing. The challenge was stopped if there was a fall in FEV1 of 20% or more from the saline value, or when the highest dose had been administered. Salbutamol aerosol (200 µg) was administered to aid recovery when necessary.

**Methacholine airway challenge.** Lung function was measured by means of a Vitalograph compact pneumotach (Vitalograph, Bucks, UK), with the subject sitting and wearing a noseclip. Forced expiratory manoeuvres were repeated five times and the highest technically satisfactory FEV1 was recorded. Measurements were corrected to BTPS. Methacholine was administered using a dosimeter (MEFAR, Bovezzo, Italy). The challenge was carried out with cumulative doses of methacholine from 0.01 to 20.4 µmol. Subjects were asked not to smoke in the hour before testing and to avoid taking asthma medications for 6 h before testing. The challenge was stopped if the FEV1 fell by 20% or more from saline value, or when the highest dose had been administered. Salbutamol aerosol (200 µg) was administered to aid recovery when necessary.

**Expression of results.** Airway responsiveness can be assessed in populations either by the prevalence of airway hyperresponsiveness (AHR) or by the distribution of a continuous measure of responsiveness. For both challenges, a dose-response curve was obtained by plotting the percentage change in FEV1 from the saline value against the logarithm of the dose of agonist. From the curve, the provocative dose of histamine or methacholine that caused a 20% fall in FEV1 (PD20FEV1) was read. Subjects with a PD20FEV1 were classified as having AHR. In this comparison of results, AHR was defined firstly as a fall in FEV1 of 20% or more at any dose of the agonist, which is the definition used in the EC Respiratory Survey protocol, which uses methacholine. Secondly, as a fall in FEV1 of 20% or more at a dose of 3.9 µmol or less, which is the definition commonly used in the histamine challenge protocol.

DRR was calculated for all subjects as the percentage fall in FEV1 at last dose divided by the total dose administered [15]. One subject who had an FEV1 of less than 60% predicted when presenting for the histamine challenge was given a bronchodilator test instead. This subject had an improvement in FEV1 following bronchodilator of 37.8%, and was assigned a PD20FEV1 value of 0.1 µmol and a DRR of 200% fall FEV1/µmol for inclusion in analyses.

**Allergic sensitivity**

**Allergen droplet method.** Sensitivity to eight common allergens was measured by skin-prick test reactions on the forearm using a lancet [10]. The allergens tested were produced by Hollister-Stier (Elkhart, IN). The allergens tested and the allergy units (AU) or weight/volume (W/V) as indicated by the manufacturer were: house-dust (1:10 W/V); two house-dust mites Dermatophagoides pteronyssinus (30,000 AU/ml) and D. farinae (1:50 W/V); cat dander (1:10 W/V); rye-grass (1:20 W/V); plantain (1:20 W/V); Alternaria tenuis (1:10 W/V); and cockroach (1:10 W/V). Histamine and glycerol were used as positive and negative controls. If histamine tested negative or glycerol positive, the test was repeated and, if the same result occurred, the data were excluded. Allergens and control substances were stored in a refrigerator when not in use.

**Phazet method.** Sensitivity to 11 common allergens was measured by skin reactions on the forearm using Phazets (Pharmacia Diagnostics AB, Sweden). Phazets are lancets, precoated with allergen solution standardized for total allergenic activity (100,000 biological units/ml) or with histamine dihydrochloride (10 mg/ml). The allergens tested were: cat; two moulds (Cladosporium herbarum, Alternaria tenuis); five pollens (timothy grass, birch, ragweed, parietaria, rye-grass); two dust mites (Dermatophagoides pteronyssinus, D. farinae); and olive. Histamine and an uncoated Phazet were used as the positive
and negative controls, respectively. The Phazet was applied to the forearm at a 90 degree angle, and held in place for at least one second. Phazets were stored in a refrigerator when not in use.

Expression of results. In order to compare the allergic sensitivity methods, analysis of skin test reactions was confined to the five allergens that were common to both protocols. For both methods, wheal sizes were recorded after 15 min as the long axis and its perpendicular. Mean wheal size for each allergen was used in analyses. A wheal size of 4 mm or greater was regarded as positive [16].

Statistical methods. Data were analysed using the statistical package SAS (SAS Institute Inc., Cary, NC, USA). Both DRR and PD20FEV1 values are log-normally distributed [17]; therefore, they were converted to base 10 logarithms prior to analyses and the antilogarithm (geometric) means are reported. For all analyses, p-values were standardized at 3.9

Baseline measurement of FEV1 was compared between the two methods. The mean difference was 0.35 l (95% CI 0.29, 0.41). The Vitalograph compact pneumotach spirometer measured FEV1 slightly higher than the Mijnhardt dry rolling seal spirometer. A paired t-test showed this difference was significant p<0.001. Measurements of AHR using histamine challenge by the rapid method and methacholine airway challenge by dosimeter were compared. A total of 15 subjects did not respond to either challenge, 19 responded to both challenges, and 4 responded to methacholine but not to histamine. Of the four subjects who responded to methacholine and not to histamine, two responded at the higher dose of methacholine, and the other two had a fall in FEV1 to histamine which was less than 20% and was, therefore, not sufficient to be categorized as having AHR. The mean PD20FEV1 for methacholine was 0.35 µmol (95% CI 0.17, 0.73) and the mean PD20FEV1 for histamine was 0.55 µmol (95% CI 0.33, 0.92). The proportion in agreement, average correct classification rate and kappa values for AHR are shown in table 1. A kappa value above 0.4 indicates moderate agreement and above 0.6 indicates good agreement [20]. The value for PD20FEV1 was calculated only for the 19 subjects who had a 20% fall in FEV1 during both challenges. The difference for histamine compared to methacholine for PD20FEV1 was a 0.39 (95% CI 0.27, 0.56) fold decrease. Because lower PD20FEV1 values indicate greater severity, these values show that subjects who had a 20% fall in FEV1 were more responsive to methacholine.

When the cut-off point defining AHR was standardized at 3.9 µmol, two subjects who responded to methacholine and not histamine were reclassified as nonresponders. The mean PD20FEV1 for methacholine by dosimeter was 0.26 µmol (95% CI 0.14, 0.47) and the mean PD20FEV1 for histamine by the rapid method was 0.55 µmol (95% CI 0.33, 0.92). The proportion in agreement, average correct classification rate and kappa values for AHR are shown in table 1. Thus, there was very good agreement between the two challenge methods for measuring AHR.

To compare responsiveness as a continuous variable, DRR values were calculated for all subjects irrespective of fall in FEV1. Higher DRR values indicate greater responsiveness. The mean DRR for methacholine was 6.92% fall FEV1/µmol (95% CI 2.69, 17.79) and mean DRR for histamine was 5.75% fall FEV1/µmol (95% CI 2.93, 11.29). The difference for histamine compared...
to methacholine for DRR was a 1.19 (95% CI 0.78, 1.82) fold increase (NS). Because the confidence intervals overlap, this indicates no significant difference between results from the two challenge methods. Figure 1 shows the results from the histamine and methacholine challenge, plotted with the line of identity. In subjects with a DRR to methacholine greater than 10, the severity of the response to methacholine was greater than to histamine, but in subjects with a DRR to methacholine less than 10, the severity of the response to methacholine was less than to histamine. The difference in DRR value was a 2.82 (95% CI 1.87, 4.25) fold increase for the 21 subjects with a DRR greater than 10, and a 0.41 (95% CI 0.27, 0.63) fold decrease for the 17 subjects with a DRR less than 10.

Allergic sensitivity

Table 2 shows the mean wheal size, mean difference and number of positive test results for each allergen by the two methods of determining allergic sensitivity. The mean wheal size was larger to the Phazet method than the allergen droplet method for histamine and all allergens except D. farinae. The mean difference in wheal size between methods was significantly greater for histamine by Phazet method (p<0.001), and for ryegrass (p<0.05) and D. farinae (p<0.001) by the allergen droplet method. The number of subjects that produced a wheal and the number of wheals that were categorized as positive were different between methods. The allergen droplet method produced a greater number of positive wheals for each allergen except cat dander. Also, cat dander was the only allergen to which in three subjects the Phazet method produced a positive wheal and the allergen droplet did not.

The comparability of the two tests of allergic sensitivity is shown in table 3 for the five allergens tested by both methods. Alternaria tenuis had perfect agreement and ryegrass had good agreement. Cat and D. pteronyssinus had moderate agreement, as indicated by kappa values of about 0.5. The D. farinae allergens had poor agreement, with only 60% of subjects showing agreement for a positive test and an extremely low kappa value. Only one third of the subjects who had a positive wheal to the allergen droplet method also did to the

Table 2. – Wheal size and number of positive results for each method of assessing allergic sensitivity

<table>
<thead>
<tr>
<th>Allergen</th>
<th>Droplet method</th>
<th>Phazet method</th>
<th>Mean difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean wheal size (sd)</td>
<td>Number positive</td>
<td>Mean wheal size (sd)</td>
</tr>
<tr>
<td>Histamine</td>
<td>3.8 (1.0)</td>
<td>43</td>
<td>6.1 (2.3)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternaria tenuis</td>
<td>4.0 (2.1)</td>
<td>4</td>
<td>7.0 (2.7)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ryegrass</td>
<td>5.4 (2.3)</td>
<td>19</td>
<td>5.8 (3.0)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cat</td>
<td>2.6 (0.7)</td>
<td>1</td>
<td>4.6 (2.5)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. pteronyssinus</td>
<td>5.7 (3.0)</td>
<td>24</td>
<td>6.3 (3.1)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. farinae</td>
<td>5.1 (1.9)</td>
<td>21</td>
<td>4.0 (1.4)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allergic sensitivity to any of the above allergens</td>
<td>28</td>
<td></td>
<td>24</td>
</tr>
</tbody>
</table>

Data for wheal size are presented as mean and sd in parenthesis. #: mean wheal size calculated from wheals ≥1 mm; ##: number positive only includes those wheals ≥4 mm. *: p<0.05; **: p<0.001 calculated for subjects with a wheal to either method.
Phazet method; and, of these, only half had a Phazet wheal that was classified as positive. However, overall, there was good agreement for allergic sensitivity defined as a positive wheal to any of the five allergens tested. The distribution of wheals to the rye-grass allergen for both the allergen droplet and Phazet method is shown in figure 2. Both methods produced wheals to the rye-grass allergen which were similarly distributed and this is also reflected in the good agreement between methods. The distribution of wheals to the D. pteronyssinus allergen for both the allergen droplet and Phazet method is shown in figure 3. The distribution of wheals less than 5 mm was similar between methods, however, for wheals 5 mm or greater, the distribution was different between the allergen droplet and Phazet method.

Discussion

It is important that historical data of AHR and allergic sensitivity collected in previous studies can be accurately compared with that being currently collected as part of the EC Respiratory Health Survey [11]. This study has shown that reasonable comparisons can be made between studies which have used a protocol that includes a histamine challenge using the rapid method and allergen droplet method for measuring allergic sensitivity, and the EC Respiratory Survey protocol which has used a methacholine challenge using dosimeter and Phazets for measuring allergic sensitivity. Reasonable comparisons can be made between methods of measuring airway hyperresponsiveness for calculating both DRR and PD_{20}FEV_{1}. Furthermore, when the cut-off dose determining hyperresponsiveness is standardized to 3.9 \( \mu \)mol, these methods are comparable in detecting subjects with AHR. The methods of measuring allergic sensitivity showed good overall agreement for assessing total allergic sensitivity and for comparing rye-grass, Alternaria and D. pteronyssinus allergy.

The subjects in this study included normals and asthmatics with a wide range of airway responsiveness, and atopic and nonatopic individuals. The range of DRR extended from 0.1 to 1,000 % fall FEV_{1}/\( \mu \)mol (fig. 1). The wheal sizes were reported as the geometric area which has been recommended by VOHLONEN et al. [21], and were not corrected for the size of the histamine wheal. The two airway challenge protocols have been standardized and were administered in a strict manner. Each of the protocols was administered by one of two investigators, and no investigator administered both protocols. The number of subjects was sufficient to calculate comparability with precision relative to the repeatability of the challenges.

The ability to compare results between protocols is highly dependent on the criteria used to define measurements. There has been much recent discussion about the best way to present the results of airway responsiveness, and most researchers agree that PD_{20} FEV_{1} and DRR, often called the two point slope,
provide the most information [17, 22]. Both measures have good short-term repeatability and have good validity against symptom frequency and severity. Therefore, we have presented the results in two ways, as PD$_{20}$FEV$_1$ and to avoid the censoring of data, as DRR.

There was no difference in the overall comparability of the rapid method using histamine and the dosimeter method using methacholine. To ensure that the subject who was assigned a PD$_{20}$FEV$_1$ and DRR value did not distort the comparability of methods, the analysis was also performed omitting this subject's data. The overall comparability of methods was not different, with the difference in PD$_{20}$FEV$_1$ values changing from 0.39 to 0.36 fold decreases, and the difference in DRR values changing from 1.19 to 1.21 fold increases. There was a significant difference in FEV$_1$ measured between protocols, but because the differences in FEV$_1$ were linear over the entire scale this was unlikely to affect the comparability of method. Also the fall in FEV$_1$ was standardised for baseline measurement and the order of testing was randomised to account for any change in asthmatic status. However, this suggests that lung function measurements between methods are not easily compared.

A clear difference did exist when the group with "normal" responsiveness was compared to the group with "abnormal" responsiveness [17]. The "normal" responsiveness group had a greater response to the histamine than to the methacholine protocol. This may happen because the rapid method takes only 5 min to administer up to the final dose, whilst the dosimeter method takes up to 15 min to reach this dose. BRITTEN et al. [23] compared three methods of histamine challenge testing and found that the two methods which delivered the histamine in a short time had better repeatability than the longer method, suggesting that "a greater proportion of the cumulative histamine dose would be active at the end of the protocol" [23].

The "abnormal" responsiveness group had a shorter challenge time to both protocols, because they did not reach the final dose before they had a 20% fall in FEV$_1$, and so the greater responsiveness to methacholine may be due to the more efficient delivery of the dose. The dosimeter may deliver the dose more efficiently, but in the "normal" responsiveness group this effect may be more than offset by the cumulative effect being lost over the time taken to complete the challenge. KNOX et al. [24] compared histamine challenge by dosimeter and by rapid method, and found that in responsive subjects the responsiveness was greater to the dosimeter method than to the rapid method.

A definition of AHR using methacholine by dosimeter in the EC Respiratory Health Survey [11] has yet to be finalized, so we compared the final dose of histamine with both an equal dose and the final dose of methacholine administered to measure agreement between histamine and methacholine for classifying AHR. The agreement between methods improved slightly by using an equal dose of 3.9 µmol methacholine instead of the final dose of 20.4 µmol. However, when comparing DRR for both methods, the final dose was used in calculations. Doses of methacholine to 20.4 µmol may cause a plateau, which will alter the shape of the dose-response curve and effect the comparison with histamine [25]. There was only one subject that appeared to have a plateau to methacholine, and so the shape of all but this one curve allowed us to make comparisons of DRR at any point along the dose-response curve.

The comparability of classifying atopic individuals between methods was good overall for the five allergens common to each panel. The allergens that were found to be the most comparable are those that have the most important association with asthma [26], i.e. Alternaria tenuis, D. pteronyssinus. The agreement between methods is reassuring, considering that both the method of testing and the potency of allergen extracts used was different. The correspondence between allergen units or weight/volume and allergenic units is not known. This has led many authors to stress the importance of standardizing the dose of allergen used in tests to measure allergic sensitivity. Because more than 95% of allergic subjects are sensitized to a few common allergens, it would be helpful to standardize these allergens in panels used to define allergic sensitivity, in order to compare results using different protocols.

The mean difference in wheal size between methods was significant for histamine, rye-grass and D. farinae. For histamine, both methods produced the same number of wheals but, for rye-grass and D. farinae, the Phazet method in some subjects did not produce a wheal, and so a wheal size of 0 was used, and tended to increase the mean difference. But, when mean difference was compared only for the subjects who had a wheal to both methods the difference was significant only for histamine.

The agreement between methods for the D. farinae allergen was disappointing. The D. farinae results for the Phazet method appear to be very different from the other allergens tested by this method. The Phazet method was found to have poor comparability because it produced fewer and much smaller wheals than did the allergen droplet method. As the reactions to this allergen were very different it raises questions about the potency of this batch of Phazets for D. farinae.

The allergen droplet method consistently produced more wheals than did the Phazet method and, except for the cat allergen, produced more wheals that were classified as positive. Perhaps the greater number of wheals classified as positive can be explained by a more potent allergen extract or a more effective dose being delivered under the skin to produce a wheal. However, the exact dose of allergen delivered under the skin is not known. A recent study has found the size of a wheal to be related to the degree of trauma to the skin imposed by the method [27]. The allergen droplet method lifts up the skin as it punctures the top layer of skin, whilst the Phazet method punctures a hole directly in the skin and could be considered the less traumatic method. The smaller number of subjects with a wheal to the Phazet method could be due to the Phazet not puncturing the skin. The Phazet is pushed on the skin and held in place for a second. The difficulty is in knowing whether the Phazet has actually broken the skin,
because in some subjects the skin is taut and in others the skin is loose. Also, in three subjects the Phazet method produced a wheal to the cat allergen, whilst the allergen droplet method did not. It is interesting that, of the multiple tests performed, this occurred only in the three subjects, and in all cases to the cat allergen.

In order to compare the prevalence of asthma and atopy worldwide, it is important that the screening protocols are standardized or at least comparable. This study has demonstrated that measurements of airway responsiveness in epidemiological studies using either of the protocols tested can be compared with confidence when the results are weighted with the comparability ratios presented. When using these protocols to compare the prevalence of AHR and allergic sensitivity, it is essential that the results are standardized for the dose administered and that a standardized panel of allergens is used.

Acknowledgements: The authors would like to thank the funding bodies in Australia (Allen & Hanburys, the National Health and Medical Research Council of Australia, the Asthma Foundation of NSW and the Community Health and Anti-Tuberculosis Association) and in New Zealand (the New Zealand Asthma Foundation, New Zealand Lottery Board and the Health Research Council of New Zealand) for funding this project. In addition, they thank the subjects who participated in the studies.

References