

TECHNICAL NOTE

Statistical approaches to the identification of late asthmatic reactions

S.C. Stenton, A.J. Avery, E.H. Walters, D.J. Hendrick

Statistical approaches to the identification of late asthmatic reactions. S.C. Stenton, A.J. Avery, E.H. Walters, D.J. Hendrick. ©ERS Journals Ltd 1994.

ABSTRACT: Late asthmatic reactions can be difficult to recognize because of their prolonged time course and the confounding effects of superimposed circadian rhythms of ventilatory function. Conventional methods of analysis are rather arbitrary. They depend for example on a 15 or 20% fall in forced expiratory volume in one second (FEV₁) from baseline or from time-matched control measurements. We have, therefore, investigated whether statistical approaches applied to individual subjects can assist in the identification of late asthmatic reactions.

In two separate series of aerosol inhalation tests, three symptomatic workers, three asthmatic controls and three nonasthmatic controls were challenged blindly with increasing doses of two chemical agents, and saline. One of the agents, sodium isononanoyl oxybenzene sulphonate (SINOS) was a suspected cause of occupational asthma. Prior to the challenges, FEV₁ was measured hourly on three control days. Cumulative late changes on both control and active challenge days were quantified as the area between a line extrapolated from a 10.00 h baseline and the actual measurements from 12.00–22.00 h (the 2–12 h area decrement). The area decrement measurements on control and active challenge days were compared using Student's t-tests. The sensitivity of this method for detecting late asthmatic reactions among potentially positive tests (SINOS challenge tests in the workers) was examined, as was its specificity. The latter was determined from the false positive rate among the negative tests. A second statistical method based on the pooled standard deviation of serial (hourly) FEV₁ measurements was investigated in the same way. In total, the data from 220 challenge and 30 control days were available for analysis.

Late responses associated with falls in FEV₁ of 8–16% were statistically significant when a t-test was used to compare area decrement on each active challenge day with three control days. This approach was, therefore, potentially more sensitive than conventional techniques for identifying late asthmatic reactions. The false positive rate was 4%. The serial FEV₁ method was more sensitive, identifying a further five positive tests, but was less specific, with a false positive rate of 7%.

These results suggest that when the day-to-day variability of lung function has been estimated from at least three control days, statistical tests can be applied to potential late asthmatic reactions, allowing them to be identified with greater precision than conventional clinical techniques.

Eur Respir J., 1994, 7, 806–812.

Chest Unit, Newcastle General Hospital,
University of Newcastle-upon-Tyne, UK.

Correspondence: S.C. Stenton
Chest Unit
Newcastle General Hospital
Westgate Road
Newcastle-upon-Tyne
NE4 6BE, UK

Keywords: Asthma
late asthmatic reaction

Received: March 29 1993
Accepted after revision November 14 1993

Late asthmatic reactions are important features of asthma. They are widely used as experimental models of the disease and in diagnosing occupational asthma [1, 2]. They can often be demonstrated in the absence of any immediate reaction, but their prolonged time course makes them difficult to study [2]. Serial measurements of ventilatory function must be made over several hours, and random or diurnal variations in ventilatory function during this period can either mimic or mask a true reaction. A number of techniques are used to identify late asthmatic reactions, but none are universally accepted and those which are in use are somewhat arbitrary.

Late asthmatic reactions are commonly diagnosed when a late fall in forced expiratory volume in one second (FEV₁), or peak expiratory flow rate (PEFR) exceeds a predetermined value, usually 15 or 20% of baseline [3, 4]. This takes no account of diurnal variation, so parallel hourly measurements on a control day and an active challenge day are sometimes compared, a late asthmatic reaction being diagnosed if these differ by more than 15% [5]. This takes account of diurnal variation, provided that it is relatively constant from day to day, and provided that the spontaneous variability of FEV₁ at any time never exceeds 15%. If a change exceeding

15% is observed on a challenge day, it is assumed to be due to a late asthmatic reaction. This carries the risk of a false positive diagnosis in a subject with considerable spontaneous variability of ventilatory function. Equally, there will be a false negative diagnosis if spontaneous variability is low and a true late asthmatic reaction decreases FEV₁ by less than 15%.

We recently performed a series of 220 inhalation challenge tests, investigating the detergent ingredient sodium isononanoyl oxybenzene sulphonate (SINOS), a new cause of occupational asthma [6, 7]. Late changes in ventilatory function were quantified using a summary (area under the curve) measurement, the 2–12 h area decrement (AD) [8]. By using slowly increasing doses from day to day, we were able to demonstrate that the magnitude of the late asthmatic reactions increased steadily with increasing challenge doses in a log-linear manner. Similar dose-response relationships have also been demonstrated with tetrachlorophthalic anhydride and with common aeroallergens [9–11], and it is now generally accepted that late asthmatic reactions cannot be regarded as "all or nothing phenomena" [1]. There is no prerequisite minimum magnitude which is essential for their diagnosis. Thus, if sufficiently sensitive statistical techniques were available, it should be possible to identify late asthmatic reactions more reliably, and at lesser levels of magnitude than is possible using conventional criteria.

By using regression analyses, we were able to show that the slopes of the dose-response relationships for late asthmatic reactions induced by SINOS in two of three workers were statistically significant. This was a novel approach, which helped confirm that the small changes in ventilatory function following the lowest SINOS doses were due to late asthmatic reactions. It is unusual to have sufficient data to generate such a detailed dose-response relationship, but similar statistical power can be achieved if lesser amounts of active challenge data can be compared with more control data from the same individual. This suggested an alternative approach, which could allow the results of a single inhalation challenge test to be tested for statistical significance. In our SINOS study, each subject had made hourly FEV₁ measurements on three control days before any active challenges were begun. We therefore re-examined the data, using Student's *t*-tests to compare the magnitudes of potential late asthmatic reactions (as ADs) on each subject's active challenge days with comparable ADs on his/her prechallenge control days. We also evaluated another statistical test using the individual serial (hourly) measurements of FEV₁. This paper reports the sensitivity and specificity of these methods.

Methods

Clinical

The data were obtained from two series of inhalation challenge tests involving the detergent agents SINOS and linear alkyl benzene sulphonate (LAS) and nine subjects (three workers with suspected occupational asthma, three asthmatic control subjects, and three nonasthmatic con-

Table 1. – Characteristics of the nine subjects

| | Age Yrs | Sex | Atopic Status * | Baseline FEV ₁ % pred | PD ₂₀ FEV ₁ ** | |
|-------------------------|------------|-----|-----------------------|--|---|-------|
| Worker | 1 | 40 | M | Y | 107 | 1020 |
| | 2 | 36 | M | N | 77 | >6400 |
| | 3 | 28 | M | Y | 110 | 239 |
| Asthmatic control | 1 | 19 | M | Y | 106 | 184 |
| | 2 | 25 | F | N | 100 | 490 |
| | 3 | 33 | M | N | 81 | 269 |
| Nonasthmatic control | 1 | 31 | M | N | 109 | >6400 |
| | 2 | 45 | M | Y | 108 | >6400 |
| | 3 | 36 | M | N | 120 | >6400 |

*: one or more skin prick tests with mean weal diameter greater than histamine control; **: airway responsiveness to methacholine, doubling cumulative protocol using a dosimeter technique, measurable range 3.125–6,400 µg [12]. No subject took medication throughout the study period. FEV₁: forced expiratory volume in one second; PD₂₀FEV₁: provocative dose producing a 20% fall in FEV₁; M: male; F: female; Y: yes; N: no.

rol subjects) (table 1). SINOS was the agent suspected of causing occupational asthma, and it gave rise to late asthmatic reactions in all three workers but none of the control subjects. LAS is chemically similar and was introduced into the study protocol to test the specificity of the effects of SINOS. LAS had never been suspected of causing occupational asthma and it did not give rise to any clinical suspicion of asthmatic reactions throughout the study.

On each test day, FEV₁ was monitored at 10 min intervals from 09.20 to 10.00 h to provide a baseline. Each subject then inhaled, in order, one of nine aerosolized doses of SINOS or LAS from a dosimeter. Only one dose was administered per day, with challenges being administered on consecutive days as far as possible. The doses increased with 3.2 ($\sqrt{10}$) fold increments over the range 0.01–100 µg. The active challenges (with SINOS and LAS) were administered double-blind in two separate series, with three or four blinded saline control challenges interspersed in each. With worker 1, the 100 µg SINOS challenge dose was omitted and the penultimate (32 µg) dose, which gave a clear late reaction, was repeated. The 100 µg SINOS dose was repeated with worker 2, but all other subjects received nine doses of both SINOS and LAS. Late changes were monitored by hourly measurements from 2–12 h after the challenge (*i.e.* from 12.00–22.00 h). Airway responsiveness was measured as the provocative dose producing a 20% fall in FEV₁ (PD₂₀FEV₁) before the challenge series with SINOS and the challenge series with LAS, and at least once as soon as each challenge series had been completed [12].

Immediately prior to the study, each subject made FEV₁ measurements at the same times for 3 days to establish the normal diurnal pattern. One subject recorded FEV₁ on a further 3 days between the SINOS and the LAS challenge series. In all, the data from 220 challenge and 30 control days were available for analysis. The tests were described in detail previously [6].

Statistical method based on summary measurement (AD)

Late changes in ventilatory function on both control and active challenge days were quantified using a summary measure, the 2–12 h area decrement (AD). This was calculated as the area between a line extrapolated from the daily mean baseline FEV₁ measurement and the FEV₁ measurements from 2–12 h after the time of the challenges, *i.e.* 12.00 to 22.00 h (fig. 1). An unpaired Student's t-test was used to compare AD measurements on three control days with the measurements following a subsequent active challenge [13]. A t-value greater than 2.92 was regarded as significant at the 5% level.

Statistical method based on serial hourly FEV₁ measurements

An alternative statistical method based on individual hourly FEV₁ measurements was also examined. It was assumed that the variability of FEV₁ measurements at each hour between 2 and 12 h postchallenge was similar, and that the standard deviations of the 11 hourly measurements could be combined to provide a pooled standard deviation. By using this pooled standard deviation in the Student's t-test, with suitable adjustments to allow for multiple comparisons, each hourly FEV₁ on a challenge day could be tested against the mean hourly measurement on three control days (see Appendix).

This technique is best represented graphically in the form of a "lower boundary" (similar to a lower 95% confidence interval) for FEV₁ measurements on control days.

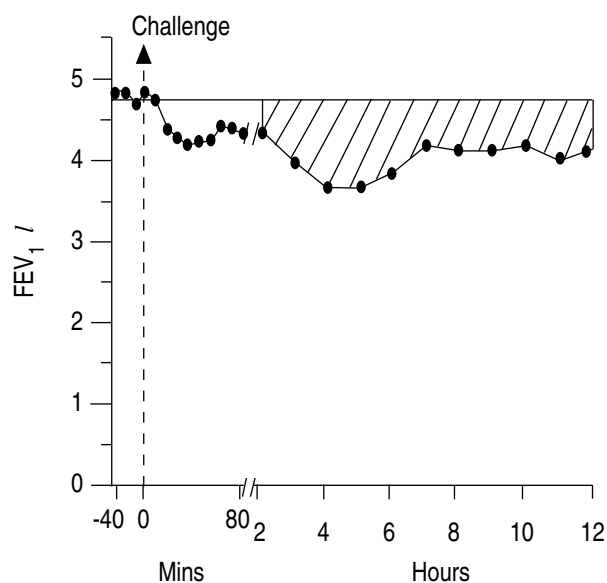


Fig. 1. – Area decrement (AD) from baseline. The response following the highest SINOS dose (100 µg) for worker 3 is shown. The AD is calculated as the area (l×h) between a line extrapolated from the mean baseline and the actual FEV₁ measurements 2–12 h postchallenge. It can be measured on control days and active challenge days. If there has been an overall increase in FEV₁ above baseline, the AD has a negative value. SINOS: sodium isononanoil oxybenzene sulphate; FEV₁: forced expiratory volume in one second.

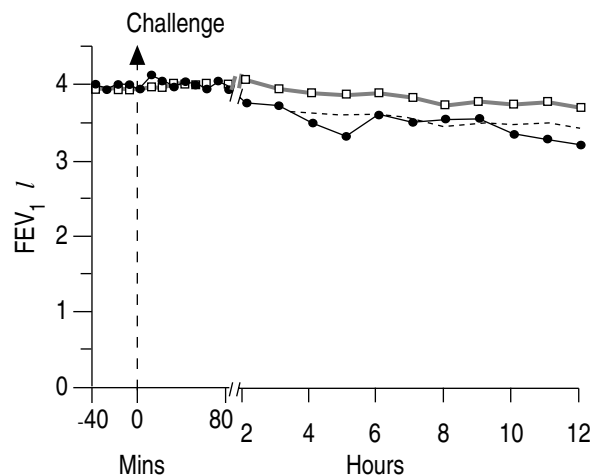


Fig. 2. – Serial FEV₁ Method. The graph illustrates the response to the 0.1 µg SINOS dose for worker 1. The excursion(s) of FEV₁ below the lower boundary represents a significant change ($p < 0.05$) on the SINOS challenge day from the control measurements. □: mean of three control days; - - -: lower boundary; -●-: SINOS 0.1 µg. For abbreviations see legend to figure 1.

The lower boundary represents the level at each hour which must be crossed for a decrement in FEV₁ on the challenge day to be significant at the 5% level. It is illustrated in figure 2, together with the mean hourly FEV₁s on the three control days and the measurements on one active challenge day. When 11 hourly measurements are available on three control days, the lower boundary can be calculated as 3.2 times the pooled standard deviation (which can be obtained from an analysis of variance) below the mean FEV₁ at each hour.

Sensitivity and specificity for detecting late asthmatic reactions

The sensitivity of these two statistical approaches for comparing active challenges with data from three control days could not be determined precisely, as there is no "gold standard" for diagnosing late asthmatic reactions. However, the number of SINOS challenge tests in the three workers (all of which were regarded as potentially positive), which each statistical method identified as significantly different from the control days, gave some guide. Each worker's SINOS responses were, therefore, compared with his three prechallenge control days using both techniques. As each worker had a number of incremental challenge doses with each agent but only a single set of control data from the three prechallenge days, multiple testing of the active challenge against the control data was necessary, and the sensitivity might, therefore, be slightly overestimated.

The problem of multiple testing against a single set of control data was avoided when determining the specificity of the statistical methods. The saline and LAS challenges in all subjects, and the SINOS challenges in the six control subjects (which had not given rise to any clinical suspicion of asthmatic reactions) were designated as

obligatorily negative, *i.e.* they were negative irrespective of whether or not either statistical method suggested a significant change. There were 183 clusters of four consecutive obligatorily negative tests among the data for the nine subjects. Within each cluster, AD and serial FEV₁ measurements on the fourth day were compared with those on the preceding 3 days using the statistical methods. Tests with saline administered to the three workers between SINOS doses were excluded from this analysis as any "responses" might have been influenced by carry-over effects from the prior exposure to SINOS.

Results

Late changes in ventilatory function during each of the 28 potentially positive challenge tests are shown in table 2. They are quantified conventionally using the maximum time-matched decrement of FEV₁ from the mean of the three prechallenge control days, and using AD. Nine of the challenges resulted in changes which might have been diagnosed as late asthmatic reactions using the conventional criterion of a 15% fall in FEV₁ from baseline, and 13 would have been diagnosed using a time-matched decrement of 15%.

Table 2. – Sensitivity of the statistical methods for detecting late asthmatic reactions

| | Worker 1 | | Worker 2 | | Worker 3 | |
|-------------------------|----------|-----------------------|----------|-----------------------|----------|-----------------------|
| | AD | FEV ₁ fall | AD | FEV ₁ fall | AD | FEV ₁ fall |
| | 1×h | % | 1×h | % | 1×h | % |
| Control 1 | 0.46 | | -0.30 | | 0.69 | |
| Control 2 | 0.24 | | 0.52 | | -1.50 | |
| Control 3 | 0.94 | | 0.34 | | -1.42 | |
| SINOS µg | | | | | | |
| 0.01 | 2.05+ | 8* | 0.72 | 3 | 0.98 | 5 |
| 0.032 | 3.04+ | 12* | 1.06 | 4 | -0.10 | 4 |
| 0.10 | 4.47+ | 14* | 1.18 | 6 | 2.83 | 10* |
| 0.32 | 4.52+ | 17* | 1.76 | 6 | 2.50 | 18* |
| 1.0 | 4.13+ | 17* | 0.20 | 2 | 3.63+ | 16* |
| 3.2 | 4.98+ | 17* | 0.44 | 7 | 1.44 | 16* |
| 10 | 6.83+ | 29* | 0.90 | 7 | 2.37 | 16* |
| 32 | 5.58+ | 30* | 0.74 | 5 | 6.45+ | 34* |
| 32 | 7.47+ | 35* | | | | |
| 100 | | | 0.98 | 16* | 7.62+ | 27* |
| 100 | | | 2.64+ | 10* | | |
| Total positive | | | | | | |
| AD method | 9/9 | | 1/10 | | 3/9 | |
| Serial FEV ₁ | | 9/9 | | 2/10 | | 7/9 |

The AD and maximum % time-matched fall in FEV₁ during the period 2–12h postchallenge are shown for each potentially positive test (*i.e.* all SINOS challenge tests in the three workers). +: statistically significant change from three control days using AD method; *: statistically significant change from three control days using serial FEV₁ method. AD: the 2–12h area decrement; FEV₁: forced expiratory volume in one second; SINOS: sodium isononyl oxybenzene sulphate.

Table 3. – Specificity of the statistical methods

| | | Obligatorily negative tests identified as positive | |
|----------------------|---|--|--------------------------------|
| | | AD method | Serial FEV ₁ method |
| Worker | 1 | 0/13 | 2/13 |
| | 2 | 1/10 | 1/10 |
| | 3 | 1/16 | 0/16 |
| Asthmatic control | 1 | 2/24 | 3/24 |
| | 2 | 1/24 | 1/24 |
| | 3 | 0/24 | 1/24 |
| Nonasthmatic control | 1 | 1/24 | 0/24 |
| | 2 | 1/24 | 3/24 |
| | 3 | 0/24 | 2/24 |
| Total | | 7/183 (4%) | 13/183 (7%) |

The clusters of obligatorily negative challenge tests (SINOS in the control plus LAS and saline in all subjects) in which the statistical methods identified a significant difference ($p < 0.05$) between the 4th test and the preceding 3 are shown. LAS: linear alkyl benzene sulphonate. For further abbreviations see legend to table 2.

Method based on summary measurement (AD)

Thirteen of the 28 potentially positive tests were significantly different ($p < 0.05$) when their ADs were tested against the three prechallenge control days. For worker 1, the method identified all challenge tests with late falls of FEV₁ of 8.2% or more as significantly different from control days. For workers 2 and 3, responses associated with late falls of FEV₁ of 9.5 and 15.7%, respectively, were identified as significant.

There were 183 sequences of four consecutive obligatorily negative tests. In 7 (4%) of these, the AD on the fourth day was significantly greater ($p < 0.05$) than those on the preceding 3 days (table 3). The false positive rate was thus 4%, *i.e.* approximately the 5% expected.

Method based on serial hourly FEV₁ measurements

This method was more sensitive than the AD method. It identified 18 of the 28 potentially positive tests in the three workers (table 2). With worker 1, the lower boundary for control FEV₁ measurements was 260 ml (7%) below the mean FEV₁ at each hour. With workers 2 and 3, the lower boundaries were, respectively, 250ml (8%) and 240ml (5%) below the mean FEV₁s. The smallest significant late changes identified for each subject by the technique are illustrated in figure 3. The specificity was less than with the AD method, the false positive rate being 7%.

A single unsustained excursion of FEV₁ below the lower boundary would not satisfy a conventional definition of a late asthmatic reaction, and so a parallel evaluation of the serial FEV₁ method was undertaken, in

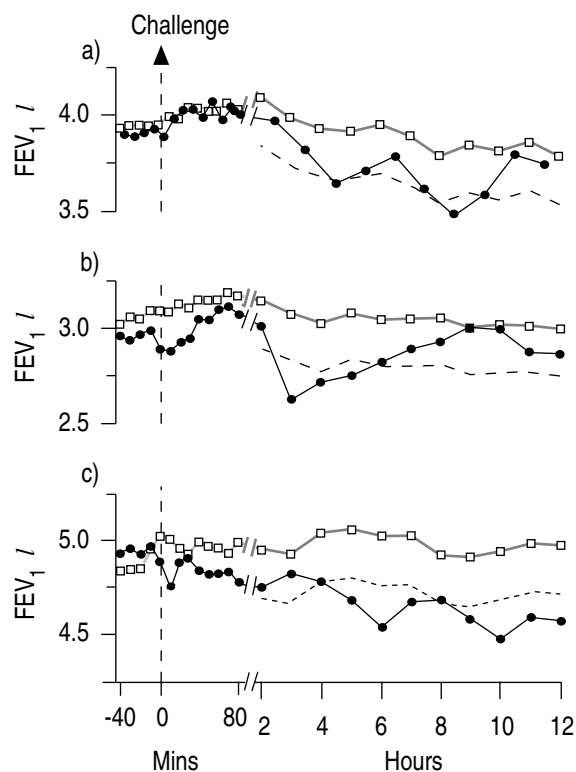


Fig. 3. — Smallest late asthmatic reactions detected. The smallest late changes detected as statistically different from control days are illustrated for each subject: a) Worker 1, SINOS 0.1 μg ; b) Worker 2, SINOS 100 μg ; c) Worker 3, SINOS 0.01 μg . For worker 1 the boundary is not breached for two consecutive hours. This did occur following all higher doses. —□—: mean of three control days; - - -: lower boundary; —●—: SINOS. For abbreviations see legend to figure 1.

which two consecutive hourly FEV_1 measurements were required to be below the lower boundary before the test was considered positive. This reduced the false positive rate to 1%, with only a slight reduction in the positive identification rate from 18 out of 28 to 17 out of 28.

Discussion

It is desirable as a general principle to support numerical results of clinical tests with an assessment of their statistical significance, and this is particularly appropriate to the detection of late asthmatic reactions following inhalation provocation tests. Numerous measurements of ventilatory function are required over a number of hours, and there are potentially confounding circadian changes which may vary appreciably from day-to-day. Only when this day-to-day variability has been quantified statistically can the investigator be reasonably confident about the meaning of any late changes on a given challenge day, unless, of course, control data show very little variability and a marked late asthmatic reaction is provoked on the active challenge day. The corollary is that statistical techniques are likely to allow less marked reactions to be recognized, thereby eliminating the need for additional, higher dose challenge tests. This, in turn, is likely to lessen any risk and discomfort to the test subject, shorten the investigation time and reduce expense.

The increased sensitivity which is possible using simple statistical tests is illustrated by the small size of some of the late changes which were identified as significant in the present investigation. The smallest decrements in FEV_1 were only 8–11% of the day's baseline and of the time-matched control measurements, considerably less than the 15 to 20% falls which are often and arbitrarily used to diagnose late asthmatic reactions. Unfortunately, there is no gold standard which can be applied in this situation to confirm that these small changes were indeed caused by late asthmatic reactions. However, in the context of the progressively increasing size of the late changes (table 2), and the clinically obvious reactions following the higher challenge doses, it seems likely that they were [6]. Furthermore, increased levels of airway responsiveness to methacholine were noted following the SINOS challenge series in the three workers but not following any of the other challenge series [6]. More importantly, the increased sensitivity of the methods was not offset by low specificity. False positive diagnoses occurred among the negative challenge tests at approximately the expected rate of 5%.

The low diagnostic thresholds of the two statistical methods in the present investigation were primarily a function of the low day-to-day variability of FEV_1 measurements in the particular test subjects. This was probably related to their mild to moderate (rather than high) levels of airway responsiveness. However, we believe that they were not atypical of the majority of subjects undergoing inhalation challenge testing, and so similar sensitivity could be generally expected. When there is greater variability, diagnostic thresholds will rise, and our wider experience suggests that in some subjects late falls in FEV_1 which are considerably in excess of 20% may not be significantly different from control days. In this situation, the use of conventional techniques to identify late asthmatic reactions could give rise to false positive diagnoses.

A Student's *t*-test can be applied to any measure of the magnitude of a late reaction, for example the lowest FEV_1 or the maximum fall from baseline, provided that data from a sufficient number of control days are available. A summary measure, such as the AD, is more valuable as it takes account of measurements throughout the period of observation and is less influenced by single outlying measurements [8]. However, it provides little information about the pattern of the response. A small fall in FEV_1 persisting throughout the 2–12 h postchallenge period can have the same AD value as a more typical late asthmatic reaction of much shorter duration. The alternative serial FEV_1 method takes this into account by allowing multiple comparisons of hourly FEV_1 measurements to be made, and it provides a useful graphical presentation.

The 7% false positive rate for the serial FEV_1 method exceeded the 5% to be expected. An analysis of simulated data suggested that this might, in part, be related to serial correlation between successive hourly FEV_1 measurements, *i.e.* the degree to which the change in FEV_1 at any hour can be predicted from change at the preceding hour. It suggested that the false positive rate might

rise to 7.7% when there is high serial correlation, and fall to 5.5% when this is low. Partly because of this, we take a conservative approach with this statistical method in clinical practice. We regard a test as positive only when two consecutive measurements are below the lower boundary. This accords better with conventional definitions of late asthmatic reactions, and in the current study it reduced the false positive rate to an acceptable 1%, whilst reducing the sensitivity only marginally.

We also find it useful in clinical practice to use the AD summary method and the serial FEV₁ method together, as each is sensitive to different aspects of a potential late asthmatic reaction. When both provide statistically significant results, we regard the challenge test as unequivocally positive, and when only one does so, the result is deemed equivocal. It is useful to monitor airway responsiveness in parallel (for example with methacholine challenge tests) because this too should increase significantly in association with a true late asthmatic reaction.

Although a positive result using these methods indicates that there has been a statistically significant late fall in ventilatory function, this does not prove it is necessarily a late asthmatic reaction to the challenge agent. Other factors, either exogenous or endogenous, could conceivably be responsible for significant late decrements in the measurements of ventilatory function (e.g. environmental factors, medications, spirometer calibration), and the investigator should endeavour to standardize the circumstances of the investigation as much as is possible, to guard against such events. A blinded technique for administering challenges is particularly useful.

These statistical approaches to the diagnosis of late asthmatic reactions require measurements on at least three control days before any active challenges are begun, in order to obtain an adequate estimate of AD or FEV₁ variance. From a statistical viewpoint, more control data would be preferable but the time, effort and expense required to gather such data would render the approach less practical. The data from the present study suggest that the use of just three control days is usually adequate.

We conclude that it is important to consider both hour-to-hour and day-to-day variability of ventilatory function when diagnosing late asthmatic reactions. Simple statistical tests can be used to take account of any changes and can provide a reliable diagnostic aid. They require increased time and effort in collecting control data, but this is offset by potentially greater sensitivity and greater confidence in interpreting the results. Such statistical methods should also allow unnecessarily high challenge doses to be avoided, thereby improving safety, and minimizing discomfort, investigation time and cost.

Appendix: derivation of the statistical methods

Area decrement (AD) method

The AD measurements from individual active challenge days were compared with the means from three control

days using a Student's t-test (one-sided, two sample). This is a standard though not commonly used application of the t-test [13]. The test statistic T was calculated as:

$$T = (AD_{\text{test}} - AD_{\text{mean control}}) / s\sqrt{(1+1/n)}$$

where AD_{test} is the AD on the test day, AD_{mean control} is the mean AD on the control days, s is the standard deviation of AD on the control days, and n is the number of control days.

T was compared with t_{n-1}(0.95) from the Student's t distribution. When T was greater than t_{n-1}(0.95) the difference was significant at the 5% level. A one-sided t-test was used, as only increases in AD were of interest.

Serial FEV₁ method

Serial FEV₁ measurements were obtained at hourly intervals between 12.00–22.00 h on three nonchallenge control days. Their pooled standard deviation (s) was then obtained from a one-way analysis of variance by time. FEV₁ measurements at matched times on a subsequent challenge day (challenge at 10.00 h) were tested by calculating a series of test statistics (T_j) as:

$$T_j = (\bar{F}_j - F_j) / s\sqrt{(1+1/n)}$$

where T_j is the test statistic for hour j, F_j is the FEV₁ at time j on the challenge day, \bar{F}_j is the mean FEV₁ at time j on the control days, n is the number of control days, and s is the pooled standard deviation.

Each T_j was compared with the 0.95^{1/k} point of a Student's t distribution with k(n-1) degrees of freedom, k being the number of hourly measurements on each day. This process ensures that if measurements at successive times are independent, and if the challenge has no effect, then the probability that none of the T_js is greater than t_{k(n-1)}(0.95)^{1/k} is 0.95.

The graphical equivalent which is illustrated in figures 2 and 3 shows the mean values for control FEV₁ measurements and the lower boundary given by:

$$\text{lower boundary} = \bar{F}_j - s\sqrt{(1+1/n)}t_{k(n-1)}(0.95)^{1/k}$$

When n=3 and k=11, as in this study, the lower boundary is 3.2 times s below the mean FEV₁ at each hour. Excursions below the lower boundary on an active challenge day suggest a statistically significant FEV₁ late fall at approximately the 5% level.

References

1. Durham SR. The significance of late responses in asthma. *Clin Exp Allergy* 1991; 21: 3–7.
2. Pepys J, Hutchcroft BJ. Bronchial provocation tests in etiologic diagnosis and analysis of asthma. *Am Rev Respir Dis* 1975; 112: 829–859.
3. Boulet LP, Roberts RS, Dolovitch J, Hargreave FE. Prediction of late asthmatic responses to inhaled allergen. *Clin Allergy* 1984; 14: 379–385.

4. O'Byrne PM, Dolovich J, Hargreave FE. Late asthmatic responses. *Am Rev Respir Dis* 1987; 136: 740-751.
5. Burge PS, Harries MG, O'Brien IM, Pepys J. Respiratory disease in workers exposed to solder flux fumes containing colophony (pine resin). *Clin Allergy* 1978; 8: 1-14.
6. Stenton S, Dennis JH, Walters EH, Hendrick DJ. The asthmagenic properties of a newly developed detergent ingredient, sodium isononanoyl oxybenzene sulphonate. *Br J Ind Med* 1990; 47: 405-410.
7. Hendrick DJ, Stenton SC, Connolly MJ, Walters EH. Occupational asthma due to isononanoyl oxybenzene sulphonate, a newly developed bleach activating agent. *Thorax* 1988; 43: 501-502.
8. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *Br Med J* 1990; 300: 230-235.
9. Venables KM, Newman Taylor AJ. Exposure-response relationships for tetrachlorophthalic anhydride asthma. *J Allergy Clin Immunol* 1990; 85: 55-58.
10. Ihre E, Axelsson GK, Zetterstrom O. Late asthmatic reactions and bronchial variability after challenge with low doses of allergen. *Clin Allergy* 1988; 18: 557-567.
11. Lai CKW, Beasley R, Holgate ST. The effect of an increase in inhaled allergen dose after terfenadine on the occurrence and magnitude of the late asthmatic reaction. *Clin Exp Allergy* 1989; 19: 209-216.
12. Beach JR, Young CL, Stenton SC *et al.* Measurement of airway responsiveness to methacholine: the relative importance of the precision of drug delivery and the method of assessing the response. *Thorax* 1993; 48: 238-243.
13. Altman DG. *Practical statistics for medical research*. London, Chapman and Hall, 1991.