

Methodology of bronchial responsiveness

Susan Chinn

Recommendations for the standardisation of bronchial challenge were made in 1985,¹ and updated in 1993.² However, neither document recommended a single protocol or even criteria that would recommend the adoption of a particular method. Even the terminology lacks uniformity, bronchial responsiveness being used in this review in the hope that it offends the fewest number of readers. Distinction has been made, but not consistently, between (hyper)responsiveness as a general term,² (hyper)sensitivity as a leftward shift in the dose-response curve, and (hyper)reactivity as an increase in the dose-response slope.³ "Airway" and "bronchial" have been used interchangeably. The prefix "hyper" has connotations of excessive, implying a bimodal distribution in the population. Although this has not been substantiated, the acronyms BHR (bronchial hyperresponsiveness) and AHR (airway hyperresponsiveness) are well established, the former being adopted in this article.

A review of the methodology of BHR might be expected to cover all aspects of bronchial challenge, but one short article cannot provide this. Rather than reproduce recommendations which were comprehensive on most details,² this paper emphasises variations in protocol that lead to results from different studies being non-comparable, explains how different methods of summary have arisen, and discusses the implications of these.

Variations in protocols

Whichever term for BHR is used, challenge with a pharmacological agent or with exercise in children, rather than allergen challenge, is usually implied. Subjects are assessed for eligibility, including adequate baseline lung function. Many different provocation agents have been used¹; histamine and methacholine now predominate, although a number of others have recently become popular.² The aerosol generated by a nebuliser is inhaled during inspiration or tidal breathing in increasing doubling concentrations until the chosen measure of lung function has fallen by a predetermined amount from its value measured after inhalation of the diluent, the chosen maximum concentration is reached, or the test is stopped for some other reason.

Although specific conductance and other measures were adopted in some early work,^{1 3 4} forced expiratory volume in one second (FEV₁) is now almost always used as the measure of lung function because of its greater

reproducibility.^{2 5} Peak expiratory flow is used in exercise challenge if measurement of FEV₁ is impracticable.⁶ Exercise challenge is not considered further here. As challenge has been made with a single "dose" in general, there are more limited options for expression of the response than with challenge with a pharmacological agent.

The use of FEV₁ is virtually the only point on which testing in adults is standardised, variations in eligibility for challenge, the provoking agent, mode of delivery, starting and maximum concentrations, and expression of the response being found in every possible combination. Clinical and community studies tend to differ in eligibility criteria and maximum dose or concentration; termination of the test when a 20% fall in FEV₁ occurs is usual in community studies but greater falls and higher doses may be achieved when justified clinically and ethically.

Dose or concentration

The result of challenge is a dose or concentration response curve. With the tidal breathing method developed for clinical use results are expressed in terms of concentration of drug delivered,⁷ but it is more common with methods recommended for epidemiological studies to calculate the cumulative dose.^{8 9} Juniper *et al* suggested that methacholine had a small cumulative effect but that the effect of histamine was non-cumulative.¹⁰ However, within one study there will be a close relation between cumulative dose and either the final dose or concentration delivered, and this may explain why the issue was not discussed in the 1993 recommendations.²

CALCULATION OF DOSE

Calculation of dose requires a measure of nebuliser output. The output by weight of the nebuliser multiplied by the concentration gives a nominal dose. This will in general exceed the actual dose as aerosol output is generally less than 100% of total output.¹¹ If precise calculation of the dose delivered is thought to be essential, then nebulisers should be pre-calibrated and standardised for all aspects of operation.¹² It is probable that differences in reports of post-study calibrations in different laboratories of Mefar jet nebulisers used in the European Community Respiratory Health Survey (ECRHS)¹³ were due to differences in driving pressure of the dosimeters used in post-study calibration (E H Walters, personal communication).

Department of Public Health Medicine, United Medical and Dental Schools of Guy's and St Thomas' Hospitals, London SE1 3QD, UK
S Chinn

Correspondence to: Miss S Chinn.

Even if the nebuliser output is pre-calibrated, output by weight constantly checked, and nebulisers filled frequently with fresh solutions of histamine or methacholine, the amount of drug delivered with each nebulisation may not be constant with use. At best the calculated dose is a good approximation to the dose delivered to the upper airway, but does not necessarily represent that received by the lung.

Results from studies that use different provoking agents, methods of delivery, or maximum concentrations will not be directly comparable even if the expression of the results appears similar. This is sometimes overlooked in arguments over how the data are analysed. The rest of this review discusses how the different summary statistics have arisen and their implications for interpretation.

Shape of the dose-response curve

The term "dose" will be used from now on, but all that follows applies equally if results are expressed in terms of concentration. If FEV_1 is plotted against the log dose of histamine¹⁴ or methacholine,¹⁵ a sigmoid curve is obtained in non-asthmatic subjects and in some mild asthmatics.¹⁴ Curves may differ in the maximal response (plateau), the slope of the steep part of the sigmoid, or in position. However, it is necessary to administer high doses in order to observe the plateau of the curve in non-asthmatics. Woolcock *et al* gave a cumulative dose of up to 122 μmol histamine to laboratory staff,¹⁴ refraining from higher doses because of the side effects. The test was terminated in asthmatic subjects when FEV_1 had declined by 60%, maximal response being assumed to be 100%.

Hence, while high doses may occasionally be given to volunteers,¹⁶ it is rarely possible to measure the maximal response in clinical studies and never in epidemiological studies in which a high response rate is required. The consequences of this are that at most the mid-curve slope and position of the curve can be estimated, and the majority of subjects will have reached neither a specified fall in FEV_1 nor a plateau.

Expression of the results

SENSITIVITY

A distinction was made between bronchial sensitivity (the dose causing a specified decrease in lung function) and bronchial reactivity (the slope of the dose-response curve beyond this point),³ with the recommendation that both should be determined. Habib *et al* suggested "threshold", the dose causing a fall in lung function greater than two standard deviations below the mean of pre-histamine values, as a measure of sensitivity,¹⁷ but PC_{20} , the concentration causing a 20% fall in FEV_1 , showed greater reproducibility^{5 18} and better discrimination between asthmatic and normal subjects.¹⁸

REACTIVITY

Reactivity was generally defined as the slope of the dose-response curve beyond the threshold dose. This has been little used recently, primarily because studies showed it to have little

relation to the clinical state in asthmatic subjects,¹⁹ that it added little to the information in PC_{20} ,²⁰ and that higher doses are required to estimate reactivity than PD_{20} alone.²¹ However, it has been suggested that sensitivity and reactivity have different clinical implications.²²

In a community study, if both mid-dose slope and sensitivity are estimated from the small number of data points that will be obtained, the two measures are likely to be highly correlated for statistical reasons. Hence, for all practical purposes only, one measure of BHR is justified.

Estimation and analysis of PD_{20}

Although originally thought of as a measure of the sensitivity component of BHR, PD_{20} (or PC_{20}) has become the most used summary of BHR itself. Provided at least a 20% fall in FEV_1 has been observed, the PD_{20} can be estimated by linear interpolation between the last two points on the dose-response curve. As doubling doses are used, interpolation is usually carried out on the log dose scale, although there is some evidence that FEV_1 is linearly related to dose.^{23 24} When the FEV_1 -log dose is sigmoid, the observed part of that curve is approximately exponential.^{25 26} However, the inherent variability of FEV_1 makes the shape of the curve often difficult to observe for an individual subject, and in clinical use the practicality of linear interpolation on the log dose scale outweighs any other consideration.

It is for community studies that other methods have been recommended, when data are necessarily computerised and maximising information is the priority. Curve fitting enables all the information to be used, rather than just the final two points, and has been shown to improve repeatability slightly.²⁵ In order to increase the number of subjects with a measurable PD_{20} , extrapolation by one doubling dose has sometimes been used but is now thought inadvisable.²⁶

Inevitably, whatever maximum dose of histamine or methacholine is permitted, and whether or not extrapolation is used, most studies find that less than 50% of the population achieve a 20% fall in FEV_1 . Hence the percentage with PD_{20} less than an arbitrary cut off point is the most used summary statistic, and logistic regression is the most used method of analysis to identify risk factors for BHR. The percentage summary is simple to understand but has several drawbacks: (1) it wastes the information in the size of PD_{20} in those with an estimate; (2) it suggests that BHR is a dichotomy whereas the evidence is that the distribution is unimodal^{27 28}; and (3) the value depends on the cut off point used which is normally the maximum dose except for reasons of comparability with other studies.²⁹

Methods to overcome the first and second problems depend on assumptions. These are essentially statistical methods for censored data, an observation being "censored" when it is known only to be above a certain limit, in this case the maximum dose.³⁰ Either method designed for censored data can be used directly,³⁰ or survival analysis methods can be

exploited, in which dose is treated as “time”, reaching a 20% fall in FEV₁ being the equivalent of failure or death in conventional survival analysis.³¹ Each method requires the assumption of a distribution for PD₂₀. As most subjects have censored data, only the left hand end of the distribution can be observed, and power to detect departure from any distribution is low. Evidence in favour of a log-Normal distribution has been presented^{28, 32} and taken as a fact by some authors,³³ but a good fit to a Weibull distribution has also been reported.³¹ Whether data exist that would show any material difference in goodness of fit, whether this would be generalisable, or whether other positively skewed distributions would fit equally well is not known. The log-Normal distribution leads to estimation of the geometric mean PD₂₀ and “fold difference” or “percent change” between comparison groups.³³ The results, assuming a Weibull distribution, are expressed in relative percentiles, the ratio of the doses required for a given percentage of individuals to achieve a 20% fall in FEV₁ in different groups.³¹

Although statistical programs are now readily available for both analyses, the degree of censoring for PD₂₀ far exceeds that which is recommended for such analyses, and other assumptions of the analyses are untestable. These are more crucial than the distribution assumption for the calculation of p values and confidence intervals—namely, homogeneity of variance of the log-Normal distributions or proportionality of hazard functions in survival analysis.

Continuous measures of BHR

While some authors have tried to maximise use of the information in PD₂₀ by the above methods, others have sought alternative measures of BHR that would enable data from all tested subjects in a population study to be included in a standard statistical analysis. As the percentage decline in FEV₁ with cumulative dose is approximately linear over the range of doses permitted, the slope of this line is very highly correlated with PD₂₀ but can be measured in all subjects challenged. O'Connor *et al* proposed that the slope should be estimated simply by dividing percentage fall from post-saline FEV₁ at the highest dose given by that final dose.³⁴ Thus an estimate is possible for all subjects given at least one dose of histamine or methacholine. Abramson *et al* proposed the slope estimated using linear regression.³⁵ This least squares slope requires at least two doses to be administered for estimation to be possible, but uses all information.³²

These two measures have immediate appeal but need to be used with caution. Firstly, they require transformation in order to satisfy the statistical requirements.³² Secondly, whether a log transformation or a reciprocal transformation is used, a constant must first be added to remove negative values that can occur in subjects with low BHR. Thirdly, the fact that every subject has a value does not guarantee that the summary measure provides information extra to that in PD₂₀. The two-point slope was found to be poorly repeatable in subjects without a measurable PD₂₀ on both occasions.³²

Hence, its use is little better than analysing PD₂₀ using censored regression. The least squares slope was found to be reasonably repeatable, but no simple transformation to a Normal distribution was found.³² Its use is equivalent to estimating PD₂₀ by extrapolation beyond the maximum dose using a linear model. Verlato *et al* found this to give better agreement with observed values than extrapolation from lower doses using an exponential curve on a logarithmic scale, but with some overestimation of PD₂₀, and consequently cautioned against extrapolation.²⁶ It should be noted that, as all data points are used, these slopes do not measure “reactivity” as defined above.

PERCENTAGE DECLINE WITH LOG DOSE

A measure of decline in FEV₁ with log dose has been proposed quite separately by two research groups. Burrows *et al* suggested BRindex, defined as the log ([% decline in FEV₁/log(final methacholine concentration in mg/dl)] + 10), as a measure of BHR in children after finding no Normalising transformation for O'Connor's slope.³⁶ Chinn *et al* proposed 100/(“log slope” + 10), where “log slope” was defined as the least squares slope of % decline in FEV₁ with log₁₀ (cumulative dose in mg), to overcome a potential problem with nebuliser batch variation in the ECRHS mentioned above.²³ BRindex and ECRHS slope are not quite equivalent, even allowing for the slight difference in transformation, in the sense that O'Connor's two-point slope is an approximation for the least squares slope on the linear scale. As zero log dose is unity on the original scale, BRindex assumes zero fall in FEV₁ at a concentration of 0.01 mg/ml. BRindex does not use all information and mean values of it and the ECRHS slope may depend on the range of doses or concentrations used, as the % decline in FEV₁ curve with log dose is not linear.

Importance of variations in protocol and expression of results

Within one study the protocol should be uniform. The continuous slope measures or methods for analysis of PD₂₀ that use all the information, either by censored regression or survival analysis methods, will have greater power to detect group differences than logistic regression of the percentage achieving PD₂₀ at some arbitrary dose.²⁸ No other difference in results has been reported. This is not surprising given the very high correlation between least squares slope on the linear scale and PD₂₀. Although the ECRHS slope and PD₂₀ are less highly correlated, these also have given similar results.^{29, 37}

At the other extreme, if results are to be compared between different studies, any of the variations in protocol may matter, particularly if levels of BHR are to be compared rather than relations of BHR to risk factors. Comparisons of prevalence of BHR between populations have been hampered, in particular, by the differences in provoking agents, variety of cut off points, and age ranges of the subjects.³⁸ There is little alternative to large scale multicentre studies if truly comparative data

are to be obtained. An additional problem is that eligibility criteria necessarily entail exclusion of subjects with poor lung function, many of whom may have high BHR, and this will differ between the populations studied. Here there is an advantage of logistic regression of PD_{20} as a sensitivity analysis of the assumption that all such subjects have PD_{20} below the cut off point can be carried out easily.³⁷

Relations of BHR to risk factors can be compared qualitatively between studies employing different protocols, but results may appear different because of sample size differences or the limited power of some analyses. Increasingly, there is a desire to combine results from different studies using meta-analysis.³⁹ Although the method was developed for combining randomised controlled trials, it can also be used to provide a quantitative summary of results from several observational studies, as used in relation to passive smoking.⁴⁰ In the context of BHR, either odds ratios from logistic regression in relation to the risk factor of interest can be combined or differences in means of one of the continuous measures. If heterogeneity between studies is detected this may be due to any of the variations in protocol that occur. Conversely, differences in protocol may obscure true heterogeneity in the effect of interest. While sensitivity analyses may help to differentiate between true and spurious heterogeneity, the variations in protocol are likely to be too great for this to be convincing. An alternative is to use effect size, the difference in means divided by the within group standard deviation. This provides a dimensionless measure, although it does not guarantee comparability. Calculation of effect size requires a continuous outcome—that is, one of the slope measures—or an estimate of standard deviation of $\log(PD_{20})$ from censored regression.

Conclusion

Given the very many variations in protocol, different ethical requirements in different countries, and suitability of methods in different circumstances, it seems unlikely that researchers or clinicians could agree to standardise the measurement of BHR. The variations may be unimportant in clinical use but hamper the progress of epidemiology. Analysis of percentage PD_{20} below an arbitrary cut off point by logistic regression is misleading, given the unimodal distribution of BHR in the population, lacks power, and is unhelpful for those wishing to combine results using meta-analysis. Least squares slope or ECRHS slope should be used, or analysis of PD_{20} using censored regression, while recognising that the first and third are essentially the same analysis in different guises, and that each measure has problems which reflect the nature of the data.

I am indebted to my colleagues on the European Community Respiratory Health Survey, particularly Dr Deborah Jarvis and Professor Peter Burney, for many fruitful discussions.

1 Eiser NM, Kerrebijn KF, Quanjer PH. Guidelines for standardization of bronchial challenge with (nonspecific) bronchoconstricting agents. *Bull Eur Physiopathol Respir* 1983;19:495–514.

- 2 Sterk PJ, Fabbri LM, Quanjer PhH, *et al.* Airway responsiveness. Standardized challenge testing with pharmacological, physical and sensitising stimuli in adults. *Eur Respir J* 1993;6(Suppl 16):53–83.
- 3 Orehek J, Gayraud P, Smith AP, *et al.* Airway response to carbachol in normal and asthmatic subjects. *Am Rev Respir Dis* 1977;115:937–43.
- 4 Rubinfeld AR, Pain MCF. Relationship between bronchial reactivity, airway caliber, and severity of asthma. *Am Rev Respir Dis* 1977;115:381–7.
- 5 Dehaut P, Rachielle A, Martin RR, *et al.* Histamine dose-response curves in asthma: reproducibility and sensitivity of different indices to assess response. *Thorax* 1983;38:516–22.
- 6 Quanjer PH, Lebowitz MD, Gregg I, *et al.* Peak expiratory flow; conclusions and recommendations of the European Respiratory Society. *Eur Respir J* 1997;10(Suppl 24):2–8s.
- 7 Cockcroft DW, Killian DN, Mellon JA, *et al.* Bronchial reactivity to inhaled histamine: a method and clinical survey. *Clin Allergy* 1977;7:235–43.
- 8 Yan K, Salome C, Woolcock AJ. Rapid method for measurement of bronchial responsiveness. *Thorax* 1983;38:760–5.
- 9 Britton J, Mortagy A, Tattersfield A. Histamine challenge testing: comparison of three methods. *Thorax* 1986;41:128–32.
- 10 Juniper EF, Frith PA, Dunnett C, *et al.* Reproducibility and comparison of responses to inhaled histamine and methacholine. *Thorax* 1978;33:705–10.
- 11 Dennis JH, Stenton SC, Beach JR, *et al.* Jet and ultrasonic nebuliser output: use of a new method for direct measurement of aerosol output. *Thorax* 1990;45:728–32.
- 12 Dennis JH, Avery AJ, Walters EH, *et al.* Calibration of aerosol output from the Mefar dosimeter: implications for epidemiological studies. *Eur Respir J* 1992;5:1279–82.
- 13 Ward RJ, Reid DW, Walters EH. Variation in nebulizer output and weight output from Mefar dosimeter: implications for multicentre studies. *Eur Respir J* 1997;10:2436–7.
- 14 Woolcock AJ, Salome CM, Yan K. The shape of the dose-response curve to histamine in asthmatic and normal subjects. *Am Rev Respir Dis* 1984;130:71–5.
- 15 Sterk PJ, Naniel EE, Zamel N, *et al.* Limited bronchoconstriction to methacholine using partial flow-volume curves in nonasthmatic subjects. *Am Rev Respir Dis* 1985;132:272–7.
- 16 de Pee S, Timmers MC, Hermans J, *et al.* Comparison of maximal airway narrowing to methacholine between children and adults. *Eur Respir J* 1991;4:421–8.
- 17 Habib MP, Paré PD, Engel LA. Variability of airway responses to inhaled histamine in normal subjects. *J Appl Physiol: Respir Environ Exercise Physiol* 1979;47:51–8.
- 18 Cockcroft DW, Berscheid BA, Murdock KY. Measurement of responsiveness to inhaled histamine using FEV₁: comparison of PC₂₀ and threshold. *Thorax* 1983;38:523–6.
- 19 Beapré A, Malo JL. Histamine dose-response curves in asthma: relevance of the distinction between PC₂₀ and reactivity in characterising clinical state. *Thorax* 1981;36:731–6.
- 20 Cockcroft DW, Berscheid BA. Slope of the dose-response curve: usefulness in assessing bronchial responses to inhaled histamine. *Thorax* 1983;38:58–61.
- 21 Malo JL, Cartier A, Pineau L, *et al.* Slope of the dose-response curve to inhaled histamine and methacholine and PC in subjects with symptoms of airway hyperexcitability and in normal subjects. *Am Rev Respir Dis* 1985;132:644–7.
- 22 Sterk PJ, Bel EH. Bronchial hyperresponsiveness: the need for a distinction between hypersensitivity and excessive airway narrowing. *Eur Respir J* 1989;2:267–74.
- 23 Chinn S, Arossa WA, Jarvis D, *et al.* Variation in nebuliser aerosol output and weight output from the Mefar dosimeter: implications for multi-centre studies. *Eur Respir J* 1997;10:452–6.
- 24 Enarson DA, Schulzer M, Vedal S, *et al.* Summarizing methacholine bronchoprovocation data in epidemiological surveys. *Bull Eur Physiopathol Respir* 1987;23:387–9.
- 25 Chinn S, Britton JR, Burney P, *et al.* Estimation and repeatability of the response to inhaled histamine in a community survey. *Thorax* 1987;42:45–52.
- 26 Verlato G, Cerveri I, Villari A, *et al.* Evaluation of methacholine dose-response curves by linear and exponential mathematical models: goodness-of-fit and validity of extrapolation. *Eur Respir J* 1996;9:506–11.
- 27 Cockcroft DW, Berscheid BA, Murdock KY. Unimodal distribution of bronchial responsiveness to inhaled histamine in a random human population. *Chest* 1983;83:751–4.
- 28 Rijcken B, Schouten JP, Weiss ST, *et al.* The distribution of bronchial responsiveness to histamine in symptomatic and asymptomatic subjects. *Am Rev Respir Dis* 1989;140:615–23.
- 29 Chinn S, Burney P, Jarvis D, *et al.* Variation in bronchial responsiveness in the European Community Respiratory Health Survey (ECRHS). *Eur Respir J* 1997;10:2495–501.
- 30 Burney PGJ, Britton JR, Chinn S, *et al.* Descriptive epidemiology of bronchial reactivity in an adult population: results from a community study. *Thorax* 1987;42:38–44.
- 31 Sunyer J, Muñoz A, Spanish group of the European Asthma study. Concentrations of methacholine for bronchial responsiveness according to symptoms, smoking, and immunoglobulin E in a population-based study in Spain. *Am J Respir Crit Care Med* 1996;153:1273–9.

- 32 Chinn S, Burney PGJ, Britton JR, *et al.* Comparison of PD₂₀ with two alternative measures of response to histamine challenge in epidemiological studies. *Eur Respir J* 1993;6: 670–9.
- 33 Peat J, Unger W, Combe D. Measuring change in logarithmic data, with special reference to bronchial responsiveness. *J Clin Epidemiol* 1994;47:1099–108.
- 34 O'Connor G, Sparrow D, Taylor D, *et al.* Analysis of dose-response curves to methacholine. *Am Rev Respir Dis* 1987; 136:1412–7.
- 35 Abramson MJ, Saunders NA, Hensley MJ. Analysis of bronchial reactivity in epidemiological studies. *Thorax* 1990;45: 924–9.
- 36 Burrows B, Sears MR, Flannery EM, *et al.* Relationships of bronchial responsiveness assessed by methacholine to serum IgE, lung function, symptoms, and diagnoses in 11-year-old New Zealand children. *J Allergy Clin Immunol* 1992;90:376–85.
- 37 Chinn S, Jarvis D, Luczynska C, *et al.* Individual allergens as risk factors for bronchial responsiveness in young adults. *Thorax* 1998;53:662–7.
- 38 Neukirch F, Cartier A. Mesure de l'hyperreactivite bronchique en epidemiologie. *Rev Mal Resp* 1994;11:101–9.
- 39 Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol* 1995;48:167–71.
- 40 Cook DG, Strachan DP. Parental smoking, bronchial reactivity and peak expiratory flow variability in children. *Thorax* 1998;53:295–301.