Continuing the debate about measuring asthma in population studies

J K Peat, B G Toelle, G B Marks, C M Mellis

Abstract

The reasons for measuring atopy and airway hyperresponsiveness (AHR) and the methods of validating measurements of asthma in population studies continue to be debated. The debate has centred around standards against which to validate asthma measurements but the absence of a “gold standard” makes the criterion validation of measurements difficult. Questionnaires will always be useful but cannot be validated against a doctor diagnosis because of self-selection and recall biases. In practice, measurements should be selected on the merits of what they measure rather than being regarded as validated or non-validated alternatives. The measurement of AHR is invaluable because it is reliable, not influenced by variations in symptom perception or diagnostic trends, and is closely related to the underlying mechanisms of asthma. The value of AHR lies in its high specificity (rate of true negatives) and low sensitivity (rate of false positives) against asthma symptoms which gives additional information about symptomatic subjects. Atopy is also a useful test and, in quantifying its association with asthma, we should not place any currency on ecological evidence. Atopy is a strong risk factor for asthma in the presence of regionally specific allergens and ecological analyses that ignore these effects are diversionary rather than productive. For preventing asthma, we need to identify the group at greatest risk of developing it, measure the risk factors with precision, and develop interventions that are effective in changing environmental exposures and homogenous outcomes. This is the only approach that has the potential to lead to significant public health benefits. (Thorax 2001;56:406–411)

Keywords: asthma; atopy; population studies

As asthma epidemiologists reflect on the success of a new generation of exciting collaborative worldwide studies, the reasons for measuring atopy and airway hyperresponsiveness (AHR) in population studies continue to be debated. In essence, the debate has centred around the usefulness of measuring atopy and AHR, and how to validate various measurements of asthma. However, the most important issue is not which measurements to include or exclude in future studies on the basis of their perceived validity. Rather, the crucial questions should centre around how the broad range of information about respiratory symptoms, airway function, and allergic status that is available from each subject can help us to answer important questions about the prevalence, aetiology, natural history, and prevention of asthma.

The main reasons for conducting epidemiological studies are shown in table 1. When designing any study the key issue is to articulate the aims and, once these are decided, it becomes clear which measurements are needed. For each purpose it is important to use measurements of different aspects of asthma and its related syndromes, especially when we need to quantify both the presence and severity of this disease. The difficulty in measuring asthma across populations is that the disease is a complex entity with pathologists, physiologists, clinicians, patients, and epidemiologists all having different perspectives.

In all studies it is essential to use reliable and valid measurement tools to reduce the chance of error and thereby increase the potential to fulfil the study aims. In this review we re-examine the published evidence in light of the recent debate about the usefulness of measuring atopy and AHR in asthma epidemiology studies, particularly in the context of planning more purposeful studies in the future.

Methodological issues

The measurement of prevalence rates is standard practice in studying the epidemiology of any chronic disease. The accurate surveillance of illness in different populations is not only important for monitoring changes over time, but is also essential for identifying populations

<table>
<thead>
<tr>
<th>Table 1 Reasons for conducting epidemiological studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>• To measure prevalence and to quantify the burden of illness in order to generate hypotheses or allocate health resources</td>
</tr>
<tr>
<td>• To identify aetiological factors and understand mechanisms in order to develop better preventive and treatment strategies</td>
</tr>
<tr>
<td>• To identify “at risk” groups in order to administer preventive pharmacological or environmental strategies efficiently</td>
</tr>
<tr>
<td>• To evaluate responses to new treatments or interventions</td>
</tr>
</tbody>
</table>

www.thoraxjnl.com
with different rates of illness. If exposure information is also measured, such studies are powerful tools for identifying the risk factors that lead to a higher burden of illness in some communities than in others. This approach has been used to compare communities who are genetically similar but who have experienced very different living conditions.\(^{7,10}\) who are exposed to different allergens,\(^{11}\) or who were born into different time based cohorts.\(^{12}\) The inherent strength of these studies has been the use of standardised methods to minimise measurement bias and thus strengthen confidence in the comparisons.

In any prevalence studies it is important to use a wide range of measurement tools that have good repeatability and good validity and are therefore accurate markers of asthma severity. For this purpose, non-differential misclassification error is not such a major issue as has been suggested\(^{1}\) since random error will not unduly influence prevalence. In general, random error will also result in a more conservative estimate of the magnitude of risk factors, and perhaps to a greater confidence in their likely effects. On the other hand, systematic misclassification as a result of inaccurate or invalid measurements will lead to biased estimates of prevalence and associations.

In comparative prevalence studies it is important that the outcome measurement is reproducible across the settings. This approach ensures maximum flexibility to compare various measurements or combinations of measurements between populations. In the recent debate the main issue has been the selection of a measurement that is an appropriate “gold standard” against which to validate measurements of asthma. However, the absence of any gold standard for asthma means that no single measurement can reliably provide all of the information of asthma in the population. Thus, it seems sensible to assume that information that is based on a wide range of measurements, all of which are acknowledged as being relevant to asthma, will give a more comprehensive overview of the extent of the asthma problem in any community. However, the interpretation of this information relies on an understanding of what each measurement actually means in relation to the burden of asthma in an individual.

**Questionnaire measurements**

In the design of epidemiological studies, much effort has been invested in developing questionnaires for measuring the impact of asthma morbidity. Questionnaires will always be useful data collection instruments because they are cost efficient and because they collect information from the subject’s own perspective of their symptom history. However, individual questionnaires must be appropriate for the study setting. For example, for documenting the burden of illness across a population a symptom questionnaire is a suitable tool. Also, when investigating health service utilisation rates for asthma it is appropriate to select subjects who identify themselves as having asthma and to examine the rate of hospital admissions or doctor consultations that are associated with the illness. However, for research questions about risk factors or aetiological mechanisms we need to be able to identify people who have a condition that is both as recognisable and as homogenous as possible. Because of wide variations in symptom perception and diagnostic practices, questionnaires are not reliable for this purpose.

The presence of a doctor’s diagnosis obtained by questionnaire is not a robust measure of asthma for many studies or for use as a gold standard against which to validate other measurements. The identification of a doctor’s diagnosis of asthma is unreliable as an epidemiological tool because of the wide variation reported in the criteria for making a diagnosis.\(^{13}\) Also, the range of diagnostic information that is available to the general practitioner who is responsible for making the diagnosis is often more limited than the information that is available in an epidemiological setting. Furthermore, even after the diagnosis is made, reporting of the diagnosis in response to a questionnaire relies on the patient’s recall that a diagnosis was made and this may be subject to systematic or non-systematic errors. An inherent problem with the clinical labelling of asthma by a physician is that a diagnosis may have been given in the past to symptoms that have since resolved. Because few physicians give a diagnosis of “asthma in remission”, the labelling is rarely removed in a clinical context. At best, the measurement of a doctor’s diagnosis may be relevant for measuring the cumulative prevalence of asthma but is not useful for identifying current severe illness that has specific characteristics, which is particularly important for measuring clinically important risk factors.

In practice, a subject’s reporting of a doctor’s diagnosis of asthma represents a limited range of information about the presence of asthma and may not be the best gold standard against which to validate other measurements. By using questionnaires to measure symptom history, researchers are largely replicating the method that a doctor uses to reach a diagnosis of asthma during a consultation. The questions are very similar to the type of information that is sought by doctors in deciding whether to offer a diagnosis and how to treat the presenting symptoms. Thus, it is no surprise that a doctor’s diagnosis of asthma is more closely related to questionnaire information about symptoms of asthma than to the results of AHR, atopy, or lung function tests.

**Validating measurements of asthma**

It is important to recognise that epidemiological methods of measuring asthma cannot be criterion validated because no gold standard exists. Asthma is a complex multifaceted illness that can vary from an extremely mild form through to a severe life threatening illness,\(^{14}\) and from a symptomatic to a non-symptomatic abnormality.\(^{15}\) Clearly, the criterion validity of any marker of the presence or severity of asthma cannot be simply measured by the
agreement between measured diagnosis, symptom presence, symptom severity, or AHR. Because subjects with no symptoms or with mild symptoms are unlikely to seek a medical opinion, self-selection bias will have a strong influence so that a clinical diagnosis will always have a good agreement with symptom reporting. Also, subjects who classify themselves into a symptom positive group using one definition are highly likely to self-select themselves into a symptom positive group when another definition or data collection medium is used.

For these reasons, video questionnaires will invariably have good agreement with self-administered questionnaires because subjects are likely to report their symptoms, or lack of symptoms, regardless of the methods used to elicit the information. Rather than focusing on agreement between methods, it is more important to select a measurement that is highly reliable for identifying a group with a specific condition. In practice, any measurement of asthma is valuable if it has good repeatability and if it is meaningful in the context of the research study that is being conducted. Different measurements should be selected on the merits of what they measure rather than being thought of as validated or non-validated alternatives.

**Measurements of AHR**

Airway hyperresponsiveness is a fundamental characteristic of asthma that is closely related to the basic physiological abnormality that characterises this disease. The measurement of AHR is invaluable because it is reliable, it characterises this disease. The measurement of characteristic of asthma that is closely related Airway hyperresponsiveness is a fundamental Measurements of AHR validated or non-validated alternatives. meaningful in the context of the research study is valuable if it has good repeatability and if it is able for identifying a group with a specific condition. In practice, any measurement of asthma is valuable if it has good repeatability and if it is meaningful in the context of the research study that is being conducted. Different measurements should be selected on the merits of what they measure rather than being thought of as validated or non-validated alternatives.

**Measurements of AHR**

Airway hyperresponsiveness is a fundamental characteristic of asthma that is closely related to the basic physiological abnormality that characterises this disease. The measurement of AHR is invaluable because it is reliable, it characterises this disease. The measurement of characteristic of asthma that is closely related Airway hyperresponsiveness is a fundamental Measurements of AHR validated or non-validated alternatives. meaningful in the context of the research study is valuable if it has good repeatability and if it is able for identifying a group with a specific condition. In practice, any measurement of asthma is valuable if it has good repeatability and if it is meaningful in the context of the research study that is being conducted. Different measurements should be selected on the merits of what they measure rather than being thought of as validated or non-validated alternatives.

**Table 2 Calculation of diagnostic statistics using recent wheeze as the index of “disease present or absent” and airway hyperresponsiveness as the index of “test positive or negative”**

<table>
<thead>
<tr>
<th>Disease present (diagnosed asthma)</th>
<th>Disease absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test positive (AHR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(AHR) 358 (a)</td>
<td>220 (b)</td>
<td>578</td>
</tr>
<tr>
<td>Test negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>557 (c)</td>
<td>1919 (d)</td>
<td>2476</td>
</tr>
<tr>
<td>Total</td>
<td>915</td>
<td>3054</td>
</tr>
</tbody>
</table>

Sensitivity = a/(a+c) = 0.39; specificity = d/(b+d) = 0.90; positive predictive value = a/(a+b) = 0.62; negative predictive value = d/(c+d) = 0.78; positive likelihood ratio = sensitivity/ (1—specificity) = 3.9.

with other clinical markers. The high specificity indicates that subjects who are symptom negative are not likely to have AHR—that is, there is a high rate of “true negatives”. On the other hand, a low sensitivity indicates that not all subjects who are symptom positive are also AHR positive—that is, there is a high rate of “false positives”. These very qualities render AHR a perfect tool for delineating people who have symptoms associated with a physiological abnormality of the airway from those who have symptoms in the absence of an abnormality or who have an illness that is in remission. Evidence suggests that the two groups of subjects with symptomatic AHR and with symptoms in the absence of AHR differ in important clinical aspects. The former group has many of the associated aspects of asthma that has an impact on quality of life. A test that agreed perfectly with symptom information—that is, with a high sensitivity and specificity when assessed against questionnaire responses—would be of no value since it would be an expensive method with which to duplicate more efficient methods of collecting questionnaire information. It is the additional information about asymptomatic and symptomatic subjects that makes AHR a useful test.

Tables 2 and 3 show the sensitivity and specificity of AHR against a doctor diagnosis of asthma or recent symptoms of wheeze as indicated in three cross sectional studies of 8–11 year old children conducted in coastal regions of New South Wales. The values of sensitivity and specificity against the two markers are similar because only 5% of children had wheeze in the absence of a diagnosis of asthma. Thus, the children who fall into the categories of “negative symptoms” and a “negative diagnosis” are largely the same group of children. However, 11.3% of the sample

**Table 3 Calculation of diagnostic statistics using recent wheeze as the index of “disease present or absent” and airway hyperresponsiveness as the index of “test positive or negative”**

<table>
<thead>
<tr>
<th>Disease present (recent wheeze)</th>
<th>Disease absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test positive (AHR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(AHR) 337 (a)</td>
<td>246 (b)</td>
<td>583</td>
</tr>
<tr>
<td>Test negative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>409 (c)</td>
<td>2078 (d)</td>
<td>2487</td>
</tr>
<tr>
<td>Total</td>
<td>746</td>
<td>3070</td>
</tr>
</tbody>
</table>

Sensitivity = a/(a+c) = 0.45; specificity = d/(b+d) = 0.89; positive predictive value = a/(a+b) = 0.58; negative predictive value = d/(c+d) = 0.84; positive likelihood ratio = sensitivity/ (1—specificity) = 4.1.
reported a previous diagnosis of asthma but had no symptoms in the previous year. Thus, the specificity of AHR as a diagnostic tool for recent wheeze is lower than for diagnosed asthma because the proportion of children who have had symptoms in the past but not recently is removed from the “symptom positive” category. This is an important concept because we need a test that will not classify the group as having current asthma when the disease has resolved or is in remission.

**Interpreting measurements of atopy**

It has been suggested that asthma is best defined in terms of the phenomena involved—that is, airflow obstruction—and that no restrictions should be made on possible aetiological associations. We strongly agree with these arguments. For this reason it is very important that atopy, which is fundamental to the aetiology of asthma, is not incorporated as part of the definition of asthma. It is also important in quantifying the relation between atopy and asthma that we do not place too much currency on ecological or circumstantial evidence of effect. In studies in which symptoms, sensitisation, and exposure have been measured simultaneously, atopy is such a strong risk factor associated with asthma that its role should not be disregarded. In addition, numerous studies have provided robust evidence of the strong relation between atopy, exposure to allergens, and asthma and asthma.

It should not be argued that a lack of ecological evidence is evidence that atopy has no relation to asthma. There are obvious reasons why atopy to a regionally specific allergen has a real association with asthma within a community but may not have an ecological association with asthma symptoms across a range of communities. Because asthma is a complex disease, two regions may have a similar prevalence of illness but very different aetiological factors leading to the level of illness. Thus, any ecological comparisons that negate the effects of regionally specific environmental factors are likely to be diversionary rather than productive in identifying important risk factors. Rather than focusing on ecological evidence of no effect, it is much more important to measure the regionally specific associations between exposure to allergens, atopy, and asthma and to use them to work towards developing environmental preventive strategies.

**Atopy as a risk factor for asthma**

In examining the role of atopy in asthma it is important that both specific outcomes and causal pathways are investigated. It is also vital that the weak association between atopy and a heterogeneous outcome such as a doctor’s diagnosis of asthma or reported symptoms is not regarded as evidence of no effect. The proposed mechanism for the association between atopy and asthma is that the inhalation of allergen bearing particles leads to the development of the airway wall abnormalities that are associated with asthma in subjects sensitised to the specific allergen. Thus, the level of exposure to respirable particles and the prevalence of sensitisation are unlikely to be the same across communities. In addition, the dose response relation between exposure and sensitisation and between exposure and the development of disease may vary widely between communities. The nature of these relationships may be different between allergens and between regions because some allergenic particles, such as those from pollens, may be too large to enter the lower respiratory tract but may enter the nose and lead to sensitisation and rhinitis, but not to asthma.

Because these mechanisms are complex and exposures vary widely both within and between regions, the presence of any atopy is a very poor marker of a specific and clinically relevant sensitisation and exposure. Thus, the use of atopy as a surrogate for these measurements will obviously dilute the measured association between the level of sensitisation and exposure to the region’s dominant respiratory allergen particles and the prevalence and severity of disease in that region. This problem cannot be solved by treating atopy to any one specific allergen as the relevant exposure since this exposure will vary widely between sites. For example, mold allergens are important in humid regions and moulds such as *Alternaria* are important in dry, desert, or rural regions. For these reasons, the role of atopy cannot be measured by ecological associations. The true role of atopy can only be measured with any accuracy in studies in which exposure and sensitisation are specifically quantified and in which a homogenous outcome measurement is used.

Any measurement of attributable risk that does not take these factors into account will be significantly biased towards the null. The only method of accurately assessing the overall attributable risk for atopy is to take the weighted mean of the strongest attributable risk for a specific allergen at a specific location. In addition, a robust measurement of current asthma would need to be used. Only this type of approach could be used to measure the proportion of asthma cases that are attributable to sensitisation to a locally dominant respirable allergen. Any errors in these methods would lead to an underestimation of effect, but this will not be nearly as significant as the underestimation that is introduced by calculating attributable risk from atopy and diagnosed asthma alone.

**Predicting prognosis**

As with all public health interventions, effectiveness will be increased if interventions target the group at greatest risk of developing an ongoing clinically important illness. For this purpose, we need to be able to predict which children will have the most severe illness in the future. Atopy and AHR measured during childhood have a relatively high sensitivity and specificity for this outcome and, as a result, have a high positive likelihood ratio for predicting the presence of asthmatic symptoms in later life. In a longitudinal cohort study we found...
that the positive likelihood ratio for predicting active asthma in early adulthood was 2.2 for atopy, 3.1 for recent wheeze, and 3.8 for AHR measured at 8–10 years. However, data from all of the other prospective studies that have been conducted show that symptoms in early life have relatively poor positive likelihood ratios of 1.0–1.7 for predicting future illness—that is, they are barely better than tossing a coin. Clearly, atopy and AHR are much better predictive tools than are measurements of symptom history.

In recent years the inclusion criteria for enrolling “high risk” infants in primary prevention studies has been a genetic disposition—that is, the presence of symptoms in a parent or sibling. However, data from the cross sectional studies shown in tables 2 and 3 show that the presence of maternal asthma has a low sensitivity of 0.21, a specificity of 0.85, and a low positive likelihood ratio of 1.40 for predicting AHR at 8–10 years. If both parents have asthma, the sensitivity increases to 0.37 while specificity falls to 0.76 and the positive likelihood ratio is 1.54. Thus, parental asthma is a very poor predictor of children developing AHR. Although a diagnosis in the mother has a better predictive value for a diagnosis in the child with a sensitivity 0.27 and a positive likelihood ratio 2.3, this estimate is likely to be inflated as a result of a shared diagnostic awareness by the mother who usually completes the questionnaire.

We may never be able to predict perfectly which children will develop asthma or will continue to be symptomatic, or which children will grow up to be non-symptomatic as adults. Thus, we need to take a more pragmatic approach and continue the search for factors to improve the specificity of screening tests so that we can target a group who are highly likely to have an ongoing and clinically important illness. We also need to improve the effectiveness of screening tests so that we avoid imposing interventions on a group who have no need for them. In striving towards this goal, tests of atopy and AHR during childhood seem a more useful starting point for building algorithms to predict future asthma than either early symptoms or a positive family history of asthma.

Planning treatment and environmental interventions

In planning treatment and environmental interventions it is essential that we maintain an inclusive view that encompasses all of the broad epidemiological knowledge that has been accumulated in recent years. We need to design studies that build on the evidence that atopic children are now more susceptible to developing asthma than their parents were at the same age, that different allergens are important in causing asthma in different regions, and that other factors such as diet, environmental tobacco smoke, and breast feeding are regionally influential. To develop preventive strategies we need precise information on the effects of the major preventable risk factors and of the factors that predict which children will benefit most. This can only be achieved with good markers of asthma severity. The quest for preventing or treating asthma by the manipulation of allergen exposure began over a decade ago and, since then, many studies of primary, secondary and tertiary prevention have been conducted with largely encouraging outcomes. Although a recent meta-analysis suggested that there was no conclusive evidence for the effectiveness of house dust mite allergen avoidance, we should certainly not be discouraged. Few studies included in the meta-analysis actually achieved reductions in allergen levels or were specifically designed to measure effectiveness, and most enrolled adults in whom established asthma may not be reversible by any means. Thus, the benefits of allergen avoidance will not be certain until studies of a much better quality are conducted.

Future directions

It is important to continue to appraise prevention studies critically and to strive to conduct more rigorous studies in the future. Because disease prevention as a result of public health strategies has been attained in many other health areas, we need to strive to develop effective public health strategies for the prevention of asthma in the next generation of children. Any new strategies will only be successful if better information about predictive factors is collected and more rigorous interventions are developed.

Clinical trials that build on epidemiological evidence for the role of environmental factors in childhood asthma are urgently needed. Such studies will need to enrol children who are at high risk of having symptoms in later life, to ensure that the intervention itself is effective in achieving changes in environmental exposures, and to use an outcome measurement that is reliable and homogenous. Only studies that fulfil these criteria will be able to change thinking about effective methods for managing childhood asthma and for reducing the burden of respiratory symptoms in the community. In taking this approach, we need to build on our investments in asthma epidemiology in the last two decades and to work towards a more informed consideration of the purpose of conducting new population studies in the future.

We need to conduct prospective studies in which all possible risk factors are measured concurrently. This will lead to the development of much better predictive algorithms with a high sensitivity and specificity for clinically important outcomes such as severe asthma that has an important impact on quality of life. By taking this path we will gain a much richer breadth of knowledge about the causes and prevention of asthma. In the future it really does not matter whether we prevent the development of persistent wheeze, AHR, or atopy since a reduction in any of these conditions has the potential to lead to a significant public health gain.

The authors acknowledge the contribution of Professor Ann Woolcock who inspired them to write this article.


1160–7.


Continuing the debate about measuring asthma in population studies

J K Peat, B G Toelle, G B Marks and C M Mellis

Thorax 2001 56: 406-411
doi: 10.1136/thorax.56.5.406

Updated information and services can be found at:
http://thorax.bmj.com/content/56/5/406

References
This article cites 26 articles, 11 of which you can access for free at:
http://thorax.bmj.com/content/56/5/406#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Asthma (1782)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/