Defining asthma in epidemiological studies

J. Pekkanen***, N. Pearce**


ABSTRACT: It has been suggested that, in epidemiological studies, asthma should be defined as symptomatic bronchial hyperresponsiveness (BHR). This paper critically examines the validity of this and alternative methods of defining asthma by reviewing population-based studies validating BHR and symptom questionnaires against asthma defined on the basis of a clinical assessment. It is emphasized that a single definition of asthma will not be applicable to all studies.

When the aim of a study is to compare differences in prevalence of asthma between populations, Youden’s Index (sensitivity + specificity - 1) is the best single measure of validity. BHR has similar or better specificity, but much worse sensitivity, and therefore a worse Youden’s Index, than symptom questionnaires. When the aim is to estimate relative risks, the validity of the definition of asthma depends more on its positive predictive value. Therefore, more specific methods of detecting asthmatics, such as severe symptoms, diagnoses of asthma, or symptomatic BHR may be most useful in cohort and case-control studies. In contrast, conversely, the method of choice for the first phase of prevalence comparisons is standardized written or video symptom questionnaires.

In order to explore reasons for the differences in asthma prevalence, and to estimate possible differential symptom reporting, questionnaires can be supplemented with bronchial hyperresponsiveness and other testing in subsamples of the symptomatic and nonsymptomatic subjects. However, symptoms and bronchial hyperresponsiveness should usually be analysed separately rather than combined due to the poor agreement between bronchial hyperresponsiveness and clinical asthma.


Most epidemiological studies have used symptom questionnaires to distinguish between asthmatics and nonasthmatics because of their advantages in terms of cost, convenience, and the resulting optimization of sample sizes and response rates. Symptom questionnaires have, however, potential problems arising from subjective symptom recognition and recall. In search of more "objective" markers of asthma, it has been suggested that, in epidemiological studies, asthma should be defined based on the presence of asthma symptoms together with bronchial hyperresponsiveness (BHR) [1, 2]. However, the assumption that "objective" measurements, such as BHR, are more valid than a symptom questionnaire is not necessarily true [3] and needs to be tested in validation studies.

Validation of survey instruments is usually done by comparing the results from the instrument to the "gold standard" test. However, asthma has many phenotypes [4] and currently there is no gold standard for defining asthma. Most old and new definitions [2, 5] of asthma highlight variable airflow obstruction and this definition is still followed in clinical practice. The most recent definition emphasizes inflammation [2]. These definitions are, however, more descriptions of the characteristics of asthma and do not allow clear guidelines for separating asthmatics from nonasthmatics. It has also been argued that asthma has always been a clinical diagnosis and that as yet there is insufficient information available regarding the pathogenesis, prognosis and natural history of asthma to justify a major change in the criteria of asthma [6]. Although clinical diagnosis of asthma is difficult, especially among children, agreement between clinicians appears to be fairly good [7]. Thus, although clinical assessment can not be considered to be a true gold standard of asthma, it currently represents the most appropriate standard for use in validating instruments for epidemiological studies.

The validity of an instrument depends not only on its agreement with the gold standard but also on its intended use. In contrast to the clinical situation, epidemiological studies focus on comparisons between populations or groups of population rather than on individuals. Thus, in epidemiological research, the most valid instrument is the one that introduces the least bias to this comparison. The first requirement for obtaining valid results from epidemiological studies is to use the same well-standardized methods in all the populations to be compared. In addition, the choice of the survey instrument depends on the aims of the particular study and the measure of effect that will be used together with issues such as cost, convenience, sample size and response rate.

In this paper, the existing population-based studies comparing the validity of symptom questionnaires and BHR testing to a clinical examination by a physician or a...
previous diagnosis of asthma are reviewed. Before reviewing these studies, the main issues regarding the validation of epidemiological survey instruments are briefly discussed.

Validity

Measures of validity

When validating survey instruments against a "gold standard", the results are usually expressed as sensitivity, specificity, and positive and negative predictive value [8]. Sensitivity is the proportion of subjects with "true" asthma (according to the "gold standard") that the survey instrument classifies correctly and specificity is the proportion of subjects without asthma that the survey instrument classifies correctly. The positive predictive value is the proportion of "true" asthmatics among all those who test positive according to the survey instrument. The sensitivity and specificity do not depend on the underlying "true" prevalence of the disease. Conversely, the positive and negative predictive value depend strongly on the underlying prevalence of the disease, and are therefore not generalizable across populations.

As noted above, the most valid method of detecting asthma in epidemiological studies is the one that introduces the least bias to the measure of effect. In prevalence comparisons, the focus is usually on the absolute difference in prevalence between populations. In this case Youden's index provides an appropriate measure of the validity of a particular question or technique [8]. Youden's Index is the sensitivity plus the specificity minus 1.0. When the sensitivity plus the specificity are equal to 1.0, i.e. the technique used is no better than random, Youden's Index is 0 and the expected observed risk difference between any two populations is 0.

In studies of risk factors for asthma, the risk difference is sometimes used as the measure of effect, but more commonly the focus is on the relative risk, which is the rate ratio in a cohort study and the odds ratio in a case-control study [9]. It has been argued that, in this situation specificity is often the most important validity measure [10–13]. However, sensitivity is also important and the bias in relative risk is actually dependent on the positive predictive value of the test [14], which in turn depends on the "true" prevalence of asthma, as well as the specificity and sensitivity of the test. Therefore, estimates of positive predictive value cannot be extrapolated between populations unless it can be assumed that the population prevalences are similar.

Hypothetical example

These issues are illustrated by the following hypothetical example. The example assumes that a researcher has administered questionnaires and done BHR testing in two groups, A and B (table 1). The study can be thought of as either a prevalence comparison between two areas or a cross-sectional analysis of 2,000 subjects, 1,000 of whom are exposed (B) and 1,000 nonexposed (A). The example further assumes that the "true" prevalence of asthma is 5% in group A and 20% in group B. Therefore, the true prevalence difference is 15% and the true prevalence ratio 4.0. As the true prevalence of asthma is, however, not known to the researcher, two different definitions of asthma are used to compare groups A and B: 1) a question on current wheezing; and 2) current wheezing together with BHR. Assuming the sensitivities and specificities of the instruments are as reported for children by Jenkins et al. [15], the question on wheezing clearly overestimates the prevalence of asthma due to its low specificity; however, it gives a better, and more statistically significant, estimate of the true prevalence difference between the populations than does using wheezing combined with BHR (9.9 and 6.2% respectively compared with the true difference of 15%). In contrast, wheezing with BHR gives a slightly better estimate of the true prevalence ratio between the populations than does the question on wheezing alone (1.44 and 1.76 respectively compared with the true prevalence ratio of 4.0) (table 1).

Thus, even if the specificity of an instrument is fairly high, if the prevalence of the disease in the population is low, the positive predictive value of the test is low. Thus, although the bias in the prevalence ratio is mainly dependent on specificity [11], sensitivity also has a role. In fact, the positive predictive values and prevalence ratios obtained using the two instruments are quite similar.

Other considerations

The above example (table 1) assumes that the survey instruments classify asthmatics similarly in the two populations, i.e. that misclassification is nondifferential. However, this is often not a valid assumption. A person's response to a given question can depend on a wide variety of psychological, social and cultural characteristics, including healthcare practices, and also on the translation of the questionnaire. Furthermore, standardizing the performance of BHR testing is a major problem, especially in international studies [16], and lung function testing requires good cooperation. Comparisons among children are especially difficult, and it has been concluded that BHR results can not be compared between children of different ages and sizes [17]. These factors can, at least

Table 1. – Hypothetical example, comparing two instruments, of surveys done in populations A and B, which have "true" prevalences of asthma of 5% and 20%, respectively

<table>
<thead>
<tr>
<th>Instrument used for estimate</th>
<th>True value &quot;Wheezing&quot;</th>
<th>&quot;Wheezing&quot; with BHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prev. in Pop A %</td>
<td>5</td>
<td>22.3</td>
</tr>
<tr>
<td>Prev. in Pop B %</td>
<td>20</td>
<td>32.2</td>
</tr>
<tr>
<td>PPV in Pop A %</td>
<td>19</td>
<td>29</td>
</tr>
<tr>
<td>PPV in Pop B %</td>
<td>53</td>
<td>66</td>
</tr>
<tr>
<td>Prevalence difference</td>
<td>15</td>
<td>9.9</td>
</tr>
<tr>
<td>Chi-squared test</td>
<td></td>
<td>24.7</td>
</tr>
<tr>
<td>Prevalence ratio</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>(95% CI)</td>
<td></td>
<td>(1.25–1.44)</td>
</tr>
</tbody>
</table>

*: sensitivity 85%, specificity 81%; †: sensitivity 47%, specificity 94%, as reported for children by JenKINS et al. [15]. BHR: bronchial hyperresponsiveness; Prev.: prevalence, Pop: population; PPV: posit.
studies [19].

subjects, have been recommended for asthma prevalence
only those subjects with doctor-diagnosed asthma as
used to define the case group. One possibility is to define
BHR, then the association will not be identified if BHR is
risk of asthma through mechanisms that do not involve
asthmatics without BHR. Also, if a factor increases the
may affect the generalizability of the study findings to
ample, restricting asthma cases to asthmatics with BHR
affect the generalizability of study findings [4]. For ex-

Table 2. – Population based studies comparing both symptom questionnaires and bronchial hyperresponsiveness test results with asthma defined on the basis of a clinical assessment by a physician

<table>
<thead>
<tr>
<th>First author [Ref.]</th>
<th>Challenge symptom</th>
<th>Subjects n</th>
<th>Age yrs</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>Youden’s Index %</th>
<th>PPV %</th>
<th>Prevalence of asthma %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symptoms*</td>
<td>80</td>
<td>97</td>
<td>76</td>
<td>89</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>37</td>
<td>99</td>
<td>36</td>
<td>94</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypertonic saline</td>
<td>168</td>
<td>13-14</td>
<td>54</td>
<td>89</td>
<td>43</td>
<td>64</td>
<td>25§</td>
</tr>
<tr>
<td></td>
<td>Symptoms</td>
<td>85</td>
<td>81</td>
<td>66</td>
<td>61</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Both</td>
<td>47</td>
<td>94</td>
<td>41</td>
<td>74</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE MARCO [7]</td>
<td>Symptoms*</td>
<td>811</td>
<td>20-44</td>
<td>83</td>
<td>87</td>
<td>70</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Symptoms* and methacholine</td>
<td>49</td>
<td>99</td>
<td>47</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*: attacks of asthma or wheezing in last 12 months; #: wheeze, shortness of breath, attack of asthma or medicines for asthma in last 12 months; §: estimated from published values. PPV: positive predictive value; NA: not available.

partially, be controlled by developing standardized questionnaires [18, 19], guidelines for translating such questionnaires [20] and video questionnaires regarding asthma [21, 22], and by better standardization of BHR testing.

Missclassification is, however, not the only validity issue in epidemiological studies. In particular, participation in an asthma study may well be associated with having asthma, which would introduce selection bias into the study results. This possible bias is best avoided by achieving a high response rate, which is easier with simple questionnaires than in studies involving detailed tests. The use of histamine and methacholine to test hyperresponsiveness among children has also raised ethical concerns in some countries, resulting in very low response rates in some studies [23, 24]. Questionnaire surveys are also easier and cheaper to perform, enabling larger studies to be performed, and thus reducing random error. This is especially important in asthma epidemiology, as sample sizes of ≥1,000 subjects, but preferably of ≥3,000 subjects, have been recommended for asthma prevalence studies [19].

The definition of asthma may itself introduce bias and/or affect the generalizability of study findings [4]. For example, restricting asthma cases to asthmatics with BHR may affect the generalizability of the study findings to asthmatics without BHR. Also, if a factor increases the risk of asthma through mechanisms that do not involve BHR, then the association will not be identified if BHR is used to define the case group. One possibility is to define only those subjects with doctor-diagnosed asthma as asthmatics. This definition preferentially selects asthmatics that have more contact with health services. If the exposure of interest is also related to the use of health services, as in immunization or antibiotic use, the study results will be biased.

Thus, in summary, assuming that the misclassification of asthma is nondifferential: 1) the instrument with the highest Youden’s Index provides the most valid estimate of the prevalence difference; and 2) the instrument with the highest positive predictive value provides the most valid estimate of the prevalence ratio. Furthermore, more restrictive definitions of "asthma" (e.g. "a recent hospital admission for asthma") may have a high positive predictive value, but may not be representative of all cases of asthma. As it cannot always be assumed that the misclassification of asthma is nondifferential, it is important to use several methods and to achieve high response rates.

Comparative validity of symptom questionnaires and bronchial hyperresponsiveness testing

In this section, the available evidence from population-based studies comparing sensitivity, specificity, Youden’s Index, and the positive predictive value of symptom questionnaires and/or BHR testing against a clinical examination by a physician or a self-report of doctor-diagnosed asthma are reviewed. A Medline search from 1980 onwards was conducted for English language publications of population-based studies containing the keywords "sensitivity and specificity" together with "asthma", "bronchial hyperreactivity" or "bronchial provocation test". In cases

Table 3. – Population-based studies comparing either bronchial hyperresponsiveness test results or symptom questionnaires with asthma defined on the basis of a clinical assessment by a physician

<table>
<thead>
<tr>
<th>First author [Ref.]</th>
<th>Challenge symptom</th>
<th>Subjects n</th>
<th>Age yrs</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>Youden’s Index %</th>
<th>PPV %</th>
<th>Prevalence of asthma %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEARS [25]*</td>
<td>Methacholine</td>
<td>791</td>
<td>9</td>
<td>50</td>
<td>84</td>
<td>34</td>
<td>53</td>
<td>25</td>
</tr>
<tr>
<td>CERVERI [32]*,+</td>
<td>Symptoms*</td>
<td>115</td>
<td>15-65</td>
<td>56</td>
<td>97</td>
<td>53</td>
<td>45</td>
<td>3</td>
</tr>
<tr>
<td>RIEDLER [33]</td>
<td>Hypertonic saline</td>
<td>174</td>
<td>13-15</td>
<td>51</td>
<td>92</td>
<td>43</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Exercise test</td>
<td>174</td>
<td>13-15</td>
<td>57</td>
<td>90</td>
<td>47</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>REMES [29]*</td>
<td>Symptoms*</td>
<td>247</td>
<td>7-12</td>
<td>88</td>
<td>97</td>
<td>85</td>
<td>53</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Asthma*</td>
<td>247</td>
<td>7-12</td>
<td>82</td>
<td>99</td>
<td>81</td>
<td>76</td>
<td>4</td>
</tr>
<tr>
<td>STEEN-JOHNSEN [34]*</td>
<td>Symptoms*</td>
<td>96</td>
<td>7-13</td>
<td>63</td>
<td>99</td>
<td>62</td>
<td>92</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*: figures calculated from published values; +: figures differ from those in the original publication, which does not adjust for the sampling method; §: “Attacks of breathlessness with wheeze” ever; †: Doctor diagnosis of asthma ever or “attacks of wheezing” or “breathlessness” in the last 12 months; ‡: Doctor diagnosis of asthma ever; §: “ever asthma” or asthma symptoms after exposure to extrinsic factors. PPV: positive predictive value; NA: not available.
Comparison with clinical assessment by a physician

There were only two population-based studies that compared, in the same population, results from BHR and symptom questionnaires to a careful clinical assessment by a physician. 

Prevalence of diagnosed asthma %

<table>
<thead>
<tr>
<th>First author</th>
<th>Challenge</th>
<th>Subjects</th>
<th>Age yrs</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>Youden’s Index %</th>
<th>PPV %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SALOME [30]*</td>
<td>Histamine 4 mg·mL⁻¹</td>
<td>2363</td>
<td>8–11</td>
<td>53</td>
<td>87</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>PATTMORE [35]</td>
<td>Histamine 4 mg·mL⁻¹</td>
<td>2053</td>
<td>7–10</td>
<td>52</td>
<td>90</td>
<td>42</td>
<td>47</td>
</tr>
<tr>
<td>BACKER [36]*</td>
<td>Distilled water</td>
<td>495</td>
<td>7–16</td>
<td>100</td>
<td>74</td>
<td>74</td>
<td>19</td>
</tr>
<tr>
<td>FORASTIERE [37]</td>
<td>Distilled water</td>
<td>1777</td>
<td>7–11</td>
<td>72</td>
<td>52</td>
<td>24</td>
<td>NA</td>
</tr>
<tr>
<td>NICOLAI [23]</td>
<td>Cold air</td>
<td>5697</td>
<td>9–11</td>
<td>43</td>
<td>87</td>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>HABY [27]</td>
<td>Histamine 64 mg·mL⁻¹</td>
<td>94</td>
<td>7–12</td>
<td>22</td>
<td>96</td>
<td>18</td>
<td>NA</td>
</tr>
<tr>
<td>BURG [28]*</td>
<td>New-Zealand</td>
<td>868</td>
<td>12</td>
<td>35</td>
<td>92</td>
<td>27</td>
<td>48</td>
</tr>
<tr>
<td>COCKCROFT [26]*</td>
<td>New-Zealand</td>
<td>484</td>
<td>20–29</td>
<td>52</td>
<td>91</td>
<td>43</td>
<td>28</td>
</tr>
</tbody>
</table>

*: figures calculated from published values; #: self-report of diagnosis of asthma or, in a few subjects, symptoms; "Asthma ever". PPV: positive predictive value; PC6, PC12 and PC20: provocative concentration of histamine causing, respectively, a 6%, 12% and 20% fall in forced expiratory volume in one second; NA: not available.

of several reports from one fieldwork, only one report was selected. From the lists of references of the selected articles and other literature, several additional articles, of which seven are included in tables 2–4 [25–31] were found.

Several population-based studies [7, 32, 34, 41–43] did not adjust their analyses for the sampling method used; thus, the reported sensitivities and specificities may not apply to the original source population [44]. Therefore, those studies in which results adjusted for sampling could not be calculated from the published figures were excluded. However, given the small number of studies comparing both symptoms and BHR with clinical diagnoses of asthma, the study of DE MARCO et al. [7] was included, although the analyses were not adjusted for the sampling method used.

The cut-off point chosen to define BHR strongly affects the sensitivity and specificity of the BHR test. Except for two studies [36, 37], the results shown in the tables for studies using methacholine or histamine challenge are for BHR defined as a 20% fall in forced expiratory volume in one second (FEV1) at the maximum dose or concentration used. This maximum dose was, however, often different in different studies. In contrast, several of the studies using other challenges [23, 33, 38, 40] attempted to find the optimal definition of BHR in terms of sensitivity and specificity, and this definition is reported in the tables.

Comparison with clinical assessment by a physician

There were only two population-based studies that compared, in the same population, results from BHR and symptom questionnaires to a careful clinical assessment by a physician. JENKINS et al. [15], in population samples of adults aged 28–44 yrs and of children aged 13–14 yrs, compared results from a symptom questionnaire and from a hypertonic saline challenge with diagnoses of current asthma based on a blinded history taken by a trained physician (table 2). Self-reported symptoms had a higher Youden’s Index than did BHR both among children and young adults, mainly due to the better sensitivity of symptom questionnaires. Combining symptoms with BHR increased specificity, especially among children, but caused a strong decline in sensitivity, thereby decreasing Youden’s Index to a lower level than that found using either symptoms or BHR alone (table 2). The differences in positive predictive value were not large, except for the lower positive predictive value of BHR among adults. The generally high positive predictive values in this study are partly explained by the high prevalence of asthma in these populations.

In a population-based sample of young adults [7], three experienced clinicians made independent assessments regarding the presence of asthma on the basis of answers to a detailed standardized interview, and results from lung function, methacholine challenge, immunoglobulin E, and skin prick tests. Agreement in the assessments between clinicians was fairly good (Cohen κ 0.71). The results (table 2) are consistent with those of JENKINS et al. [15]. However, the reported results are not adjusted for the oversampling of those subjects with respiratory symptoms.

BURNEY et al. [45] studied 20 selected adult asthma cases and 20 controls in four centres in Europe both with
a histamine challenge and a symptom questionnaire. The selection of the asthma cases and controls was, however, not population-based and varied between centres. In this study, the question on “asthma” during the last year had a sensitivity of 55% and a specificity of 96%, and BHR with symptoms had a sensitivity of 44% and a specificity of 98%.

Several population-based studies have validated either symptoms or BHR against physician’s assessment (table 3), mostly in children. The results regarding sensitivity and specificity of BHR agree closely with those reported by Jenkins et al. [15]. The studies show that symptom questionnaires can also be worded to be very specific among children, which in turn tends to lower their sensitivity.

Comparison of bronchial hyperresponsiveness with self-reported asthma diagnosis

Many other population-based studies have compared BHR tests with a self-report of doctor-diagnosed asthma ever in life (table 4). These comparisons are less reliable as they are affected by possible underdiagnosis and under-reporting of asthma. However, in general, they agree with the results of Jenkins et al. [15] with specificities of approximately ≈90%, sensitivities of 20–50% and a Youden’s index of approximately ≈40%.

All but two of these studies have been performed among children. The results among adults are supported by two other population-based studies reporting a sensitivity of 32% and a specificity of 94% for methacholine challenge in detecting those middle-aged and elderly men reporting that they had ever had “asthma” [46] and a sensitivity of 56% and a specificity of 77% for histamine challenge in detecting those men aged 14–64 yrs reporting asthmatic attacks [47]. A study among 1,392 selected workers reported a sensitivity of 61% and a specificity of 85% for methacholine challenge in detecting those men with physician-diagnosed asthma [48].

Backer et al. [36] used several different cut-off points of BHR. Using a histamine concentration of 2.4 mg·mL⁻¹ and a 20% fall in FEV₁ (provocative concentration of histamine causing a 20% fall in FEV₁ (PC₂₀)) to define asthma, the definition was again highly specific for asthma, but was of low sensitivity (57%). Making the definition of BHR less severe increased sensitivity such that the provocative concentration of histamine causing a 6% fall in FEV₁ (PC₆) had a sensitivity of 100%, but the specificity decreased to 74%, yielding, however, the best value of Youden’s Index (0.74). The positive predictive value was best with the more severe definition (PC₂₀). Forastiere et al. [39] performed similar analyses, but varied the concentration of methacholine required to produce a 20% fall in FEV₁ (table 4).

It might be hypothesized that repeated measurements of BHR would give a better classification of subjects into asthmatics and nonasthmatics. Burrows et al. [49] studied the same 573 children at ages 9, 11, 13, and 15 yrs. Using airway responsiveness at least one of the four examinations as evidence of BHR, they produced a sensitivity of 65%, specificity of 71%, and Youden’s Index of 36% for frequent wheezing in the past 2 yrs at the age of 15 yrs. Conversely, using at least moderate responsiveness at all four examinations as evidence of BHR, the sensitivity was 17%, the specificity 98% and Youden’s Index 15%.

Thus, again, a more severe definition gave a higher specificity but a lower Youden’s Index. However, neither definition of BHR provided good agreement with asthma symptoms.

Conclusions

Findings from the population-based studies reviewed here show a poor sensitivity of BHR in detecting asthma in contrast with conclusions from clinical studies which have found the sensitivity of BHR to be >90% [50]. This is partly due to the case mix in the clinical studies [44], as BHR does not fare so well in a general population survey that includes many mild or borderline asthmatics, and in which many of the nonasthmatics have atopy, a family history of asthma or respiratory diagnoses other than asthma [51]. In addition, clinical studies as well as several population-based studies [7, 32, 34, 41–43] have not adjusted for the sampling method used in their analyses and therefore the reported sensitivities and specificities may not be representative of the original source population [44].

The continuous nature of BHR in the population means that the sensitivity of BHR in detecting asthma can be increased by defining BHR as very mild hyperresponsiveness. However, this in turn decreases the specificity, and even when the cut-off point of BHR has been examined in this way [23, 33, 36–38, 40], the agreement between BHR and asthma remains poor. The same conclusion was drawn from a recent Bayesian analysis, which allows for the estimation of test properties when no gold standard test is available [52].

Given the current problems in defining what actually constitutes asthma, the focus of epidemiological research should be less on trying to estimate the "actual prevalence of asthma" [53] in a population, and more on comparing the prevalence of asthma between population groups using standardized methods. When the aim is to study differences in asthma prevalence, the available evidence from population-based studies indicates that questions on symptoms have a higher Youden’s Index, and therefore greater validity than BHR alone or BHR in combination with symptoms.

In epidemiological studies of the causes of asthma, the situation is less clear. In these studies, the relative risk is often the effect measure of interest and the validity of a survey instrument is usually more dependent on its positive predictive value. More specific definitions of "asthma" may therefore have greater validity in this context. A more specific definition can be obtained by defining asthma as severe symptoms or a combination of symptoms [22], doctor-diagnosed asthma, positive results on BHR testing or combining BHR with symptoms [1, 2]. However, less information is available on the relative validity of these approaches.

Whatever method is used, it should be validated, preferably in a subsample of the populations studied. This allows the estimation of the degree of bias in a study and even correction for it [14]. When performing a prevalence survey, a good means of combining the best qualities of the symptom questionnaires and BHR testing is to first
perform a large questionnaire survey and then do more intensive examinations on a subsample [18, 19]. However, it should be stressed that both symptomatic and nonsymptomatic subjects need to be examined. This has unfortunately not always been carried out [53, 54], which makes it impossible to estimate the extent of misclassification in the questionnaire survey.

In conclusion, no single method of detecting asthmatics will suit all epidemiological studies. In cohort and case-control studies, specific methods for detecting asthmatics, such as severe symptoms, diagnoses of asthma or symptomatic bronchial hyperresponsiveness, are most useful. In contrast, the method of choice for the first phase of prevalence comparisons is a standardized written or video symptoms questionnaire. To explore reasons for the differences in asthma prevalence, and to estimate possible differential symptom reporting, questionnaires can be supplemented with bronchial hyperresponsiveness and other testing in subsamples of the symptomatic and nonsymptomatic subjects. However, symptoms and bronchial hyperresponsiveness should usually be analysed separately rather than combined due to the poor agreement between bronchial hyperresponsiveness and clinical asthma.

References


32. Cerveri I, Bruschi C, Ricciardi M, Zoecechi L, Zoia MC, Rampulla C. Epidemiological diagnosis of asthma:


