The diagnostic value of history and physical examination for COPD in suspected or known cases: a systematic review

Berna DL Broekhuizen, Alfred PE Sachs, Rimke Oostvogels, Arno W Hoes, Theo JM Verheij and Karel GM Moons

Background. According to current guidelines, spirometry should be performed in patients suspected of chronic obstructive pulmonary disease (COPD) by the results of history taking and physical examination. However, little is known about the diagnostic value of patient history and physical examination for COPD.

Objectives. To review the existing evidence on the diagnostic value of history taking and physical examination in recognizing COPD in patients suspected of COPD.

Methods. A systematic literature search was performed in electronic medical databases. Studies were included after using defined inclusion and exclusion criteria and judged on their methodological quality by using the Quality Assessment of Diagnostic Accuracy Studies criteria. A formal meta-analysis was not performed because all studied items of history and physical examination were investigated in only in a maximum of three studies.

Results. Six studies were included. The history items dyspnoea, wheezing, previous consultation for wheezing or cough, self-reported COPD, age and smoking and the physical examination items wheezing, forced expiratory time, laryngeal height and prolonged expiration were found to have diagnostic value for COPD. These items were studied in maximally three studies and study population studies were heterogenic. The reference test for COPD in five of the six studies concerned obstructive lung disease in general and not COPD.

Conclusion. There is insufficient evidence to assess the value of history taking and physical examination for diagnosing COPD.

Keywords. COPD, diagnosis, patient history, physical examination.

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the most important chronic diseases in terms of frequency, impact on quality of life and mortality. Since 2001, COPD is internationally classified according to the WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines on COPD, in which it is defined as airflow limitation that is not fully reversible. Diagnosing COPD in an early stage is relevant because early and appropriate management, especially smoking cessation interventions but also vaccination against influenza and medication, reduces the number and severity of exacerbations of COPD and improves quality of life of patients.

International guidelines recommend that COPD should be suspected in patients with recurrent and persistent lower respiratory tract symptoms, such as persistent cough. Despite the high prevalence of these symptoms, especially in primary care, recognizing COPD remains difficult. The current diagnostic pathway in patients suspected of COPD consists of history taking, physical examination and, ultimately, spirometry. Spirometry is a time-consuming test that must be executed by trained personnel. Screening every smoker for COPD with spirometry, however,
irrespective of signs and symptoms, has been shown to be neither efficient nor feasible.\(^7,8\)

There is still little knowledge about which items from history taking and physical examination have independent diagnostic value for COPD and to what extent. It is well known that there are no pathognomonic signs for COPD. Spirometry alone, however, is also not sufficient; without the knowledge of symptoms and signs, the results of spirometry are often not straightforward.\(^1\) Since history and physical examination are always performed in patients with complaints concerning the respiratory tract in medical practice, it is important to know which items are useful. Two earlier overviews on this topic were published >8 years ago, thus, before the GOLD guidelines were introduced.\(^9,10\) Therefore, we performed an updated systematic review to quantitatively summarize the diagnostic accuracy of symptoms and signs for COPD in patients suspected of having COPD.

The results of this review may help clinicians decide which patients should undergo spirometry. Moreover, the results are useful for researchers in future diagnostic studies on COPD to determine which items of history and physical examination should be included.

### Materials and methods

#### Searching

A systematic search for articles on the diagnostic value of items of history and/or clinical examination was performed by RO and BDLB in September 2008 in the following databases: Pubmed, Embase, Cochrane, CINAHL, Clinical Evidence and the Medion Database without restriction in date of publication. The search strategy included [(COPD OR chronic obstructive pulmonary disease OR obstructive lung disease OR obstructive airway disease OR obstructive airways disease OR airflow obstruction) AND (diagnosis OR diagnosing OR diagnostic test OR detecting OR detection OR screening OR case-finding OR differentiation OR identifying OR sensitivity OR specificity OR prediction)]. The search was limited to humans, adults and English language.

#### Selection

Inclusion criteria of relevant papers were

1. study population consisted at least partially of patients suspected of COPD (thus studies on screening of asymptomatic patients or on patients with known COPD were excluded);
2. original reports: we excluded letters, editorials, case-reports, commentaries and reviews;
3. the index tests on which the diagnostic accuracy was assessed were items from history taking and/or physical examination;
4. spirometry was used as the reference test;
5. cut-off points of spirometry parameters, i.e. the Forced Expiratory Volume in one second (FEV1), the forced vital capacity (FVC) or ratio between FEV1 and FVC, were clearly defined.

#### Validity assessment

The methodological quality of the selected original diagnostic accuracy studies was graded independently by two observers (RO and BDLB), using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool, an evidence-based tool for the quality assessment of studies on diagnostic accuracy.\(^11\) In case of doubt, a third reviewer (KGMM) was consulted. The items of the QUADAS tool enabled us to find potential sources of bias in the studies or problems regarding generalizability of the study results (Appendix 1).

#### Data extraction

The two observers extracted the following elements from the included studies: study objectives; recruitment setting, such as general practice or hospital; population characteristics such as age, gender and smoking behaviour; duration of symptoms or signs; details of the index tests, i.e. description of the physical examination and history items under study; whether the study was cross-sectional or longitudinal; details of the reference test, such as cut-off points used for the spirometry parameters and whether spirometry was executed after administration of bronchodilators (i.e. reversibility testing to acquire post-bronchodilator values); blinding of the observer of the reference test for the index tests results; whether the observer of the index tests was blinded for the reference test results (if applicable); time interval between the index tests and the reference test; number of patients enrolled; number of patients who underwent the index test and reference test and missing data.

Additionally, the estimated parameters of diagnostic accuracy of the index test—either from univariable or multivariable analysis—were extracted.

Discrepancies were resolved by discussion between the two reviewers or, if agreement could not be reached, by consultation of a third and fourth reviewer (APES and KGMM).

#### Quantitative data synthesis

Positive and negative predictive value, sensitivity, specificity and likelihood ratios of each of the index tests in relation to the outcome (i.e. diagnosis of COPD by spirometry) were retrieved or otherwise calculated if possible.\(^12\) If only multivariable diagnostic parameters were described in an article and not the univariable data or vice versa, the authors were contacted to obtain the missing information. A meta-analysis could not be performed because each index test was only studied in a maximum of three of the six included studies.
Results

Search results
Our search yielded 4396 potentially eligible articles. After applying the inclusion and exclusion criteria to the abstracts, the number of papers was reduced to 45. Of these, the full texts were retrieved and independently judged by three readers (BDLB, RO and APES). Reference lists of these 45 articles were scanned to find additional studies. Applying our inclusion and exclusion criteria resulted in six relevant articles to address our objective (Figure 1).

Study characteristics
In all six studies, the majority of QUADAS criteria were accounted for but none of the studies fulfilled all 14 criteria (Table 1). One criterion, the reporting of uninterpretable or intermediate results, was not fulfilled in five of the six studies.

The study population size varied across the six studies from 161 to 703 patients and the age of patients varied from 18 to 85 years old (Table 2). The study population of Straus et al.16 in 2002 formed part of the study population of Straus et al. in 2000.14 Both studies were nevertheless included because the index tests (forced expiratory time in 2002 and laryngeal measurements in 2000) and cut-off points for the reference tests were different in both studies. Age of the study participants, recruitment setting and the definition of ‘suspected of COPD’ differed considerably between the six studies.

In total, 10 patient history items and 9 physical examination items were evaluated in the six studies (Table 3). Most often studied items were age, gender and the physical examination item wheezing.

The definition of the outcome used in the studies varied, including obstructive Airways disease,15,16,18; obstructive lung disease13; COPD or asthma17 and COPD14. In only one of the six studies, spirometry results after administration of bronchodilators were included (Table 3).17

Quantitative data synthesis
In five of the six studies, results of multivariable analysis were provided, reflecting the independent diagnostic value of the index tests (Table 4).13,14,16–18 Three studies reported multivariable likelihood ratios and two reported $\beta$-coefficients. In four studies, both univariable and multivariable results were specified. Despite contacting the authors of the included articles for more information, it was not possible to report one single diagnostic accuracy measurement in all six studies.

Of the 10 studied patient history tests, eight were found to have independent diagnostic value for COPD: age ≥45 years, female sex, dyspnoea, wheezing, current smoking, >40 pack years of smoking, previous consultation for wheezing or cough, self-reported history of COPD and symptoms provoked by allergens. Dyspnoea, wheezing and smoking were found to have independent diagnostic value in two studies, the other five tests in only one study. The strongest tests were symptoms provoked by allergens (odds ratio = 4.5), wheezing (odds ratio = 4.4), >40 pack years of smoking (positive likelihood ratio = 11.6) and self-reported history of COPD (positive likelihood ratio = 4.4).

Of the nine studied physical examination items, five were found to have independent diagnostic value for COPD: wheezing, forced expiratory time 9 seconds14...
Table 1  Summary of the methodological quality of the six retrieved studies on the diagnostic accuracy of history and physical examination items for the diagnosis of COPD, in patients suspected of COPD, reflected by the 14 QUADAS criteria (+, Yes, ?, unclear reported, -, No, n.a., not applicable)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Representative patient sample</th>
<th>Selection criteria clearly described</th>
<th>Adequate reference standard</th>
<th>Adequate cross-sectional design</th>
<th>Complete verification of diagnosis</th>
<th>No differential verification</th>
<th>No incorporation bias</th>
<th>Adequate index-test measurement description</th>
<th>Adequate reference-test description</th>
<th>Blinding for reference test results</th>
<th>Blinding for index-test results</th>
<th>Clinical data available as in practice</th>
<th>Uninterpretable/Intermediate results reported</th>
<th>Withdrawals explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffels et al.13</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>+</td>
<td>e</td>
<td>n.a. d</td>
<td>e</td>
<td>+</td>
</tr>
<tr>
<td>Straus et al.14</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e</td>
<td>+</td>
<td>e</td>
<td>+</td>
</tr>
<tr>
<td>Garcia-Pachon</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?b</td>
<td>+</td>
<td>+</td>
<td>e</td>
<td>+</td>
<td>e</td>
<td>+</td>
</tr>
<tr>
<td>Straus et al.16</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e</td>
<td>+</td>
<td>e</td>
<td>+</td>
</tr>
<tr>
<td>Thiadens et al.17</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e</td>
<td>+</td>
<td>e</td>
<td>+</td>
</tr>
<tr>
<td>Schapira et al.18</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>e</td>
<td>+</td>
<td>e</td>
<td>+</td>
</tr>
</tbody>
</table>

Full text of items are listed in Appendix 1

*The study population existed of patients referred to a pulmonary function test laboratory at a tertiary care hospital for several reasons; it was unclear whether these patients (all) had symptoms suggestive of COPD.

*Definition of the physical signs: wheezing, rhonchi and reduced breath sounds was not provided.

*The same investigators performed index tests and reference test.

*Index tests were documented by a questionnaire, filled in by the patient and not by a physician.

*Uninterpretable results of index test or reference test were not reported.

Table 2  Characteristics of study population of the six papers included in the systematic review

<table>
<thead>
<tr>
<th>Authors</th>
<th>N</th>
<th>Definition of ‘suspected of having COPD’</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Male [N (%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffels et al.13</td>
<td>703a</td>
<td>Cough for &gt;2 weeks, breathing difficulties, wheezing, nasal allergy or hay fever.</td>
<td>Primary care</td>
<td>35–70</td>
<td>308 (44)</td>
</tr>
<tr>
<td>Straus et al.14</td>
<td>161b</td>
<td>Known COPD, suspected COPD or neither known nor suspected COPD by treating physician.</td>
<td>Primary care and secondary care</td>
<td>&gt;50</td>
<td>99 (61)</td>
</tr>
<tr>
<td>Garcia-Pachon15</td>
<td>172</td>
<td>Smoking &gt;20 years; self-reported COPD; bronchodilator treatment &gt; 6 months or dyspnoea.</td>
<td>Secondary care</td>
<td>&gt;40</td>
<td>117 (68)</td>
</tr>
<tr>
<td>Straus et al.16</td>
<td>233c</td>
<td>Known COPD or suspected COPD</td>
<td>Primary care and secondary care</td>
<td>&gt;18</td>
<td>124 (53)</td>
</tr>
<tr>
<td>Thiadens et al.17</td>
<td>192</td>
<td>Cough for &gt;2 weeks</td>
<td>Primary care</td>
<td>18–75</td>
<td>72 (38)</td>
</tr>
<tr>
<td>Schapira et al.18</td>
<td>400</td>
<td>Referred to pulmonary function laboratory (no further definition given)</td>
<td>Tertiary care</td>
<td>24–85</td>
<td>377 (98)</td>
</tr>
</tbody>
</table>

N, total number of patients.

*Of the total study population of this paper, only patients with complaints suggestive of COPD (i.e. 703 of 3158) were included in the review.

*These 161 participants form part of the 309 participants of Straus 2000.

*Of the total study population of this paper, only patients suspected of COPD were included in this review (i.e. 233 of 309), as this was the aim of our review.
### Table 3: Overview of the index tests and reference tests (outcome definition) studied in the six studies included in our review

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Prevalence COPD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffels et al.</td>
<td>2004</td>
<td>18</td>
</tr>
<tr>
<td>Schapira et al.</td>
<td>1993</td>
<td>53</td>
</tr>
<tr>
<td>Thaidens et al.</td>
<td>1998</td>
<td>6</td>
</tr>
<tr>
<td>Garcia-Pachón et al.</td>
<td>2002</td>
<td>56</td>
</tr>
<tr>
<td>Strauss et al.</td>
<td>2000</td>
<td>74</td>
</tr>
<tr>
<td>Strauss et al.</td>
<td>1998</td>
<td>56</td>
</tr>
<tr>
<td>Straus et al.</td>
<td>2002</td>
<td>56</td>
</tr>
<tr>
<td>Straus et al.</td>
<td>2000</td>
<td>74</td>
</tr>
<tr>
<td>Thiadens et al.</td>
<td>1998</td>
<td>6</td>
</tr>
<tr>
<td>Schapira et al.</td>
<td>1993</td>
<td>53</td>
</tr>
</tbody>
</table>

- +, index test which diagnostic accuracy was estimated. Hoover sign = paradoxical inspiratory indrawing of the left side of the chest.
- <50% = below the fifth percentile according to the American Thoracic Society.
- Number of patients with proven COPD among those patients suspected of COPD was not specified.
- FEV1/FVC <70% = below the fifth percentile according to the American Thoracic Society.

### Patient History
- Age
- Gender
- Smoking history
- Self-report COPD
- Family history asthma
- Cough
- Dyspnoea
- Wheezing
- Allergy
- Previous consultation for wheezing or cough

### Physical Examination
- Forced expiratory time
- Prolonged expiration
- Laryngeal height
- Laryngeal descent
- Hoover sign
- Clinical impression

### Parameter of Spirometry
- FEV1/FVC
- FEV1
- FEV1/FVC or FEV1 and ratio
- Both <70% (women)
- Both <70% (men)
- FEV1 <70% at baseline and <75% after 8 weeks

### Cut-off points
- Post-bronchodilator test

### Reference test (outcome)
- FEV1 <50%
and 6 seconds\textsuperscript{18}, maximum laryngeal height and prolonged expiration. The strongest diagnostic tests were prolonged expiration (odds ratio = 3.7), forced expiratory time >9 seconds (positive likelihood ratio = 4.6) and maximum laryngeal height (the distance between the top of the thyroid cartilage and the suprasternal notch)\textsuperscript{16} of \( < 4 \) cm (positive likelihood ratio = 3.6).

The study with only univariable results showed that reduced breath sounds, Hoover sign (paradoxical inspiratory indrawing of the lateral rib margin) and clinical impression on presence versus absence of obstructive airways disease had a significantly increased positive likelihood ratio and decreased negative likelihood ratio.\textsuperscript{15}

### Discussion

In this review, we summarized all available evidence on the diagnostic accuracy of history and physical examination in patients suspected of COPD. We found only six studies that matched our inclusion and exclusion criteria. In these studies, history items that were found to have independent diagnostic value for ruling in or out COPD were >45 years, female sex, dyspnoea, wheezing, current smoking and extensive smoking (>40 pack years), previous consult for wheezing or cough, self-reported history of COPD and allergen-induced symptoms. Physical examination items that were found to have independent diagnostic value were wheezing, forced expiratory time, laryngeal height and

### Table 4 Accuracy parameters of items that were found to have diagnostic value in the included studies

<table>
<thead>
<tr>
<th>Authors</th>
<th>Item (h) or (pe)</th>
<th>Univariable results</th>
<th>Multivariable results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sens. Spec. LR+ (95% CI)</td>
<td>LR– (95% CI)</td>
<td>PPV</td>
</tr>
<tr>
<td>Buffels et al.\textsuperscript{13}</td>
<td>Dyspnoea (h)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Wheezing (h)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Previous consultation for wheezing or cough (h)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td>Strauss et al.\textsuperscript{14}</td>
<td>Self-reported history of COPD (h)</td>
<td>n.s. n.s. 5.6</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Wheezing (pe)</td>
<td>35% 0.89% 4.0 (1.6–9.9)</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &gt; 9 seconds (pe)</td>
<td>n.s. n.s. 6.7 (2.1–1.1)</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td>Garcia-Pachon\textsuperscript{15,16}</td>
<td>Reduced breath sounds (pe)</td>
<td>59% 82% 3.38 (2.14–5.32) 0.49 (0.36–0.68) 67% 77%</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Hoover sign (pe)</td>
<td>58% 86% 4.16 (2.49–6.96) 0.49 (0.36–0.67) 71% 77%</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Clinical impression on presence/absence of COPD (pe)</td>
<td>83% 81% 4.26 (2.86–6.35) 0.21 (0.12–0.37) 72% 89%</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td>Strauss et al.\textsuperscript{16}</td>
<td>Age &gt; 45 years (h)</td>
<td>n.s. n.s. 1.5 (1.1–2.2)</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Pack years of smoking: &gt;40 (h)</td>
<td>n.s. n.s. 11.7 (2.7–50.0)</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td>Thiadens et al.\textsuperscript{17}</td>
<td>Max laryngeal height &lt;4 cm (pe)</td>
<td>n.s. n.s. 4.2 (2.3–7.9)</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Female sex (h)</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Pack years of smoking (h)</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Dyspnoea (h)</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Wheezing (h)</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>Symptoms provoked by allergens (h)</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Schapira et al.\textsuperscript{18}</td>
<td>Prolonged expiration (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>Age &lt; 60</td>
<td>n.s. n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &lt; 2 seconds (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &gt; 2 to &lt; 4 seconds (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &gt; 4 to &lt; 6 seconds (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &gt; 6 to &lt; 8 seconds (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
<tr>
<td></td>
<td>FET &gt; 8 seconds (pe)</td>
<td>n.s. n.s. n.s.</td>
<td>n.s. n.s.</td>
</tr>
</tbody>
</table>

\textsuperscript{h, patient history; pe, physical examination; FET, forced expiratory time; n.s., not specified in article; b-coeff, \( b \)- or regression coefficient; LR+, likelihood ratio of a positive index test; LR–, likelihood ratio of a negative index test; Sens., sensitivity; Spec., specificity; PPV, positive predictive value; NPV, negative predictive value; OR, odds ratio, LR, likelihood ratio.}

\textsuperscript{a}Accuracy parameters for the items were obtained by pulmonologists and by residents in this study. In this review, we included the accuracy parameters obtained by pulmonologists.

\textsuperscript{b}ORs were calculated from the \( b \)-coefficients that were stated in the articles.
prolonged expiration. Most items that were found to have diagnostic value were identified as such in no more than one study.

The number of studies that matched our inclusion criteria was surprisingly low. In contrast, many more adequate diagnostic studies have been published on the diagnostic value of history and physical examination items for other important chronic diseases, such as coronary artery disease or peripheral artery disease. A possible explanation for the relative lack of diagnostic studies on COPD might be the changing definition of COPD that has occurred over time.

One of the reasons for this review was the development of the GOLD criteria in 2001 to define the presence or absence of COPD, based on spirometry with reversibility testing as reference test. According to these guidelines, COPD is present, when after bronchodilatation the FEV1:FVC ratio is <70%. Hence, our review included only those diagnostic studies with spirometry as reference test. The cut-off values for the spirometry results to define COPD in the six studies varied. None used the GOLD criteria; instead the guidelines of the European Respiratory Society or the American Thoracic Society, both universally accepted prior to the GOLD guidelines, were used. Moreover, in five of the six studies, the spirometric diagnosis was made without bronchodilatation, which made optimal differentiation between asthma or other reversible airflow obstruction and COPD difficult. Spirometry without bronchodilatation was adequate for the end point of four of the six included studies, namely obstructive airways disease and obstructive lung disease, but not for the current definition of COPD.

According to the QUADAS criteria, we found that the overall design, methodology and reporting of the six included studies was adequate. The studies were all cross-sectional and verification of the diagnosis by a reference test was achieved in all included patients. The reporting of uninterpretable or intermediate results, however, was absent in five of the six studies. Suboptimal reporting is not unusual in diagnostic research. Guidelines for the conduct and reporting of diagnostic research have been proposed by the Standards for Reporting of Diagnostic Accuracy steering committee to improve the quality of diagnostic studies, but these guidelines were proposed after the conduct of the six included studies.

A possible weakness of this review is the fact that study populations of the six included studies differed considerably in setting and age (Table 2). The difference in setting, e.g. primary, secondary and tertiary care, is usually reflected in disease severity. Concerning age, in three studies, patients <40 years old were also included. These are not all necessarily representative for the domain of this review, namely patients suspected of COPD because COPD is rare under the age of 40. Unfortunately, it was not possible to discriminate the younger patients from the others in these papers. Finally, the definitions of ‘patients suspected of COPD’ varied (Table 2). Five of the six study populations consisted of patients with and without known COPD. Preferably, one would rather have included only patients without yet known COPD. But again, it was not possible from the publications to discriminate these patients from the other patients. We did not exclude these studies because at least part of the study populations consisted of our study domain, and we did not want to miss this information. The heterogeneity across the study populations, as well as the diversity in studied items from patient history and physical examination compromise general recommendations for medical practice.

Two overviews, though not systematic reviews, on diagnostic value of history taking and physical examination for COPD have been published previously. Holleman et al. reviewed 19 articles on the end point ‘airflow obstruction’ and found that no single item of history and physical examination sufficiently ruled in or ruled out airflow obstruction. Combinations of items, like smoking, reduced breath sounds and a peak expiratory flow <350 l/minute, were concluded to be more useful. This is in accordance with our results that also indicate that various items are independent diagnostic determinants of COPD. McAllister et al. wrote an overview about the accuracy of physical examination items for COPD but did not study history items. They showed that the accuracy of the well-known signs for COPD, such as reduced breath sounds, wheezes or hyper resonance, varies greatly between studies and that no sign is sufficiently accurate on its own. Both reviews included more studies than our review because they also included articles with another diagnostic test than spirometry as reference test, articles that formally were not a diagnostic accuracy study comparing an index tests with a reference test, articles evaluating end points other than obstructive pulmonary disease, e.g. emphysema, and text books. Furthermore, both reviews included studies on screening of COPD, meaning testing all people in a population irrespective of signs and symptoms. In regard to this screening, two original diagnostic studies have been performed by Van Schayck et al. and Geijer et al., both on the detection of chronic obstructive lung disease in smokers. In these studies, cough and age were related to chronic obstructive lung disease.

Conclusions

The available evidence for the diagnostic accuracy of patient history and physical examination for COPD is
very limited, let alone the independent diagnostic value of these items. Moreover, existing studies notably studied obstructive lung disease in general, rather than COPD according to current guidelines. Hence, the available evidence cannot yet determine with sufficient confidence which patient history and physical examination items can be used by physicians to select those patients, suspected of COPD, who require spirometry. However, this systematic review does show that dyspnea, wheezing, cough, self-reported history of COPD, age, smoking, forced expiratory time, laryngeal height and prolonged expiration are the strongest diagnostic parameters from patient history and physical examination among those that have been studied. Future diagnostic studies on COPD are necessary to estimate the true diagnostic accuracy of these items in patients suspected of COPD.

Declaration

Funding: The Netherlands Organisation for Health Research and Development, ZonMW/NWO research (945-04-015).

Ethical approval: none.

Conflict of interest: none.

References


7. Geijer RMM. Detection of COPD in smokers. Thesis. Utrecht: Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 2006.


Appendix 1. The QUADAS tool

All items to be scored as yes, no or unclear

1. Was the spectrum of patients representative of the patients receiving the test in practice?
2. Were patient selection criteria clearly described?
3. Is the reference standard likely to correctly classify the target condition?
4 Is the time period between index test and reference standard short enough to be reasonably sure that the target condition did not change between the two?
5 Did the whole sample or a random selection of the sample receive verification by the reference standard?
6 Did patients receive the same reference standard regardless of the index test results?
7 Was the reference standard independent of the index tests (i.e. the index tests did not form part of the reference standard)?
8 Was the execution of the index tests described in sufficient detail to permit replication of the tests?
9 Was the execution of the reference standard described in sufficient detail to permit its replication?
10 Were the index test results interpreted without knowledge of the results of the reference standard?
11 Were the reference standard results interpreted without knowledge of the results of the index tests?
12 Were the same clinical data available when test results were interpreted as would be available when the tests are used in practice?
13 Were uninterpretable or intermediate test results reported?
14 Were patient withdrawals from the study explained?