Chronic obstructive lung disease “expert system”: validation of a predictive tool for assisting diagnosis

Purpose: The purposes of this study were development and validation of an expert system (ES) aimed at supporting the diagnosis of chronic obstructive lung disease (COLD).

Methods: A questionnaire and a WebFlex code were developed and validated in silico. An expert panel pilot validation on 60 cases and a clinical validation on 241 cases were performed.

Results: The developed questionnaire and code validated in silico resulted in a suitable tool to support the medical diagnosis. The clinical validation of the ES was performed in an academic setting that included six different reference centers for respiratory diseases. The results of the ES expressed as a score associated with the risk of suffering from COLD were matched and compared with the final clinical diagnoses. A set of 60 patients were evaluated by a pilot expert panel validation with the aim of calculating the sample size for the clinical validation study. The concordance analysis between these preliminary ES scores and diagnoses performed by the experts indicated that the accuracy was 94.7% when both experts and the system confirmed the COLD diagnosis and 86.3% when COLD was excluded. Based on these results, the sample size of the validation set was established in 240 patients. The clinical validation, performed on 241 patients, resulted in ES accuracy of 97.5%, with confirmed COLD diagnosis in 53.6% of the cases and excluded COLD diagnosis in 32% of the cases. In 11.2% of cases, a diagnosis of COLD was made by the experts, although the imaging results showed a potential concomitant disorder.

Conclusion: The ES presented here (COLD) is a safe and robust supporting tool for COLD diagnosis in primary care settings.

Keywords: chronic obstructive lung diseases, expert systems, diagnosis

Introduction

The umbrella term chronic obstructive lung disease (COLD) includes different pulmonary diseases whose distinctive feature is the persistent obstruction of lower airways. Irrespective of etiology or pathogenesis of the specific disease, the acronym COLD is sufficiently informative to draw attention to a worldwide health problem that must be addressed.1,2

COLD mainly involves chronic bronchitis and emphysema, although specific asthma patterns, as well as various less common lung diseases, can also be included. COPD and asthma are distinct nosology entities; however, they often present with a continuum of different patterns in which risk factors, trigger exposure, functional and biological abnormalities and symptoms interact in a complex, dynamic and heterogeneous manner.2 Owing to the variability and heterogeneity of clinical features and the limited access to lung functional tests, COLD remains often underdiagnosed, especially in primary care.2,4
Approximately 10% of the general population presents with signs of COPD and 26% of individuals suffering from chronic respiratory symptoms aged ≥45 years have indications of COLD. However, only approximately a quarter to half of these patients have received a proper diagnosis of chronic obstructive disease. In this scenario, novel case-finding strategies for the identification of hidden cases may benefit from new technology-based supporting tools estimating disease probability.

In recent years, artificial intelligence (AI) tools were used to support physicians in the diagnosis of COLD, by defining the functional respiratory defect and analyzing the results of a combination of spirometry, bronchodilatation and bronchoprovocation tests, as well as Impulse Oscillometry System (IOS). In a similar study, 323 cases of COPD were evaluated by using a clinical decision support system (CDSS), obtaining 90% specificity and 96% sensitivity. More recently, a novel tool was tested on 60 patients with COPD by using more sophisticated calculation techniques and a questionnaire based on 27 patient characteristics (including sex, dry cough, wet cough, fever, wheezing, smoking, weight loss, short of breath, chest pain, dyspnea, personal history of asthma, tuberculosis [TB], COPD, latent TB, childhood asthma patient and family history of these diseases) resulting in >90% accuracy. Although interesting results were obtained in these studies, a simple and accurate system to support physicians in their diagnostic work has not been fully developed yet. Expert systems (ESs), representing a branch of AI, are designed by combining the knowledge of human experts with inference engines suitable to answer questions on a specific topic. ESs are consulted to obtain advice, suggestions and recommendations on issues that fall within the experts’ knowledge and are widely used in many different fields of human activities (from the Internet to finance). The development of large specific data management tools is increasing the interest of the medical community.

The aim of the present study was to develop an ES to support the identification of individuals suffering from COLD in primary care settings.

To this aim, four steps were performed: questionnaire development, code development and in silico validation, expert panel pilot validation and real-life clinical validation. ES cannot be intended as a diagnostic tool. This ES quantifies the chance that an obstructive respiratory disease (eg, COPD or asthma) occurs, by means of a probabilistic value. Therefore, this ES allows addressing patients in a more appropriate and objective way to in-depth investigations leading to diagnosis. The study was approved by the Ethics Committee of Catanzaro (Italy), No 220, of November 16, 2016.

Methods

Questionnaire development

Key questions for COLD diagnosis were based on published disease recommendations. Age, sex, presence and characteristics of chronic cough, sputum and dyspnea, environmental exposure (smoke and/or known allergen sensitizations) and available diagnostic tests (chest X-ray, pre- and post-bronchodilator spirometry) were considered (Figure 1). A panel of Italian pulmonologists, with documented experience in COLD management and expertise in the field, developed, adopting a two-round Delphi method, the following rules:

1. Presence of chronic cough: even if more frequent in COPD patients (where is usually productive), cough may also be present in asthmatic patients (in whom dry cough attacks are often reported).
2. Presence of dyspnea: while persistent and progressive dyspnea is considered more common in COPD, occasional dyspnea is more prevalent in asthma. At rest, dyspnea is considered worse than exertional dyspnea.
3. Environmental and voluntary noxious exposures: the sensitization to inhalant allergens is more frequently observed in asthmatic than in COPD patients. The exposure to smoke is relevant to the diagnosis of COPD in the presence of chronic bronchial obstruction. Thus, the number of pack-years (≥15) is more frequently associated with COPD, and the risk persists also in former smokers. Smoke exposure can also occur in asthmatic patients, although to a lesser degree.
4. Age: if the subject is <45 years of age, the risk of having COPD can be considered reasonably low.
5. Asthma onset is more prevalent at younger ages, although the development of late onset of asthma in individuals >65 years of age is not uncommon.
6. Lung functional test: a post-bronchodilator persistent bronchial obstruction is typical of COPD but can also be registered in uncontrolled/severe asthma and other forms of COLD. Normal spirometry rules out the diagnosis of COPD.
7. Chest X-ray: the absence of pleural and lung thickening/infiltrates increases in the presence of respiratory symptoms and typical functional patterns increased the odds of having COLD. Vice versa, any radiologic sign (eg, bronchiectasis, pleural effusions, cysts, interstitial thickening) makes necessary primarily to investigate the occurrence of diseases other than COLD.
The expert panel assigned a specific weight to each item based on their experience in the field.

**Code development**

Starting from the weights and the rules described in the Methods section, the ES code was written using WebFlex, an advanced knowledge specification language (LPA, London, UK). The ES is based on frame rules (representing the knowledge base) driving the system itself and on forms for input and output. A user interface (UI) optimization was provided by a software house (Prospero Multilab Srl, Bologna, Italy).

The risk of COLD was calculated by adding a predefined positive or negative score to symptoms, results of the lung function test and chest X-ray findings, when available. In the presence of each required result, the ES produced a score ranging from 0 to 200 (where 200 was the highest possibility of having a COLD in the presence of all required signs and symptoms). Cutoff score, risk of suffering from COLD and score interpretation are shown in Table 1. In addition, according to the rules decided by the experts and the combination of specific symptom patterns, lung function and X-ray results, the ES also provided some specific warnings (Table 2).

### In silico validation

The in silico validation was performed to evaluate the consistency of the actual results of the ES questionnaire with the expected answers. Different information technology (IT) strategies were used (not shown) to mimic all the different possible answers to the questionnaire. The output of this validation provided the proof that the ES was performing properly and that conditions of no results were not occurring.

### Expert panel pilot validation

A pilot validation test was carried out on 60 different patients (training set). The different weights were iteratively modified (if needed) through cycles of validation, aimed at bringing the expected answers.

### Table 1: Cutoff score, risk of COLD, and interpretation implemented in the ES version used in the study (COLD<sub>es</sub>)

<table>
<thead>
<tr>
<th>Score range</th>
<th>Probability of COLD (%)</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES prediction without lung function test and X-rays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–75</td>
<td>&lt;30</td>
<td>Unlikely</td>
</tr>
<tr>
<td>75–110</td>
<td>&lt;50</td>
<td>Unlikely</td>
</tr>
<tr>
<td>110–130</td>
<td>&gt;70</td>
<td>Likely</td>
</tr>
<tr>
<td>ES prediction without lung function test and with X-rays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–50</td>
<td>&lt;20</td>
<td>Unlikely</td>
</tr>
<tr>
<td>50–100</td>
<td>&lt;40</td>
<td>Unlikely</td>
</tr>
<tr>
<td>100–150</td>
<td>&gt;70</td>
<td>Likely</td>
</tr>
<tr>
<td>ES prediction with all data (with or without bronchodilation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–60</td>
<td>&lt;10</td>
<td>Virtually excluded</td>
</tr>
<tr>
<td>60–115</td>
<td>&lt;40</td>
<td>Unlikely</td>
</tr>
<tr>
<td>115–160</td>
<td>&gt;70</td>
<td>Likely</td>
</tr>
<tr>
<td>160–200</td>
<td>&gt;85</td>
<td>Very likely</td>
</tr>
</tbody>
</table>

Abbreviations: COLD, chronic obstructive lung disease; ES, expert system.
the ES results as close as possible to experts' results. At the end of this validation step, suitable for the identification of the best scores for each sign or symptom that was relevant for the diagnosis, the definition of the final weights used in the ES tested in clinical validation was achieved. As a further result, the pilot validation phase allowed estimating the degree of concordance between the results of the ES and the results of the experts. On the basis of this concordance, it was calculated that the sample size for the clinical validation study was 260 patients for an incidence of 85%, type 1 (alpha) error of 0.01 and type 2 (power) error of 95%.

Clinical validation of the ES

For this aim, two forms of the system were built and made available in a closed network: the first was defined as the screening form where physicians uploaded the clinical and functional data registered during the visit of the patient. On this dataset, the ES ran the algorithm and calculated the probability of COLD. A second form defined as reviewer’s form was used to upload the results of the reviewer’s (the expert) final diagnosis, based not only on data uploaded during the first visit but also on all other information collected during the diagnostic process. Both forms saved data in two distinct databases in a Microsoft SQL Server. Both researchers involved in the screening procedures, and reviewers had a password to enter the system and upload the recorded information. The evaluation was blindly performed by two groups of scientists to evaluate whether the ES, loaded with real-life values, produced results in line with the experts’ opinion and experience. Starting from the results obtained in the pilot phase, the ethics committees accepted the protocol and the patient’s informed consent, and six Italian academic reference centers were engaged in the validation of the COLD<sub>ES</sub>.

The centers were requested to enroll 43 consecutive patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases. After fulfilling the informed consent, patients attending their outpatient clinics for suspected obstructive lung diseases.

Table 2 The diagnostic warnings provided by COLD<sub>ES</sub>

<table>
<thead>
<tr>
<th>Warnings</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warning 1</td>
<td>The presence of an accessional dyspnea could be associated with asthma</td>
</tr>
<tr>
<td>Warning 2</td>
<td>The presence of a sensitization to inhalant allergens is frequently associated with asthma</td>
</tr>
<tr>
<td>Warning 3</td>
<td>A younger age (&lt;45 years) is less frequently associated with COPD, but the presence of a concomitant obstructive pattern always needs further investigation (alpha-1 antitrypsin assessment)</td>
</tr>
<tr>
<td>Warning 4</td>
<td>Older age (&gt;60 years) is more frequently associated with COPD, but physiologically reduced functional lung volumes and flows could mimic COPD</td>
</tr>
<tr>
<td>Warning 5</td>
<td>The complete reversibility of bronchial obstruction is essential in the differential diagnosis between COPD and asthma. Absent or partial reversibility can be present in both asthma and COPD</td>
</tr>
<tr>
<td>Warning 6</td>
<td>Lung function tests are essential for the diagnosis of COLD</td>
</tr>
<tr>
<td>Warning 7</td>
<td>Chest X-rays are essential in excluding/confirming lung diseases other than COLD</td>
</tr>
</tbody>
</table>

Abbreviations: COLD, chronic obstructive lung disease; ES, expert system.

Figure 2 Algorithm describing the different steps in the process of validation of the tool. Abbreviations: COLD, chronic obstructive lung disease; ES, expert system.
Results
The different weights and diagnostic warnings were obtained at the end of the expert panel pilot validation by iterative cycles of improvement. The concordance analysis between the ES scores and the diagnoses performed by the experts in the pilot validation step showed an accuracy ranging between 86% and 95% for COLD and not-COLD, respectively. As specified in the “Methods” section, these findings allowed calculation of the power of the study in a total of 260 individuals.

In the clinical phase, 258 records were registered. Of these, data from 241 patients were eligible for analysis. In ten cases, symptoms alone were uploaded to the ES; in 59 patients, symptoms and absence of obstruction at baseline spirometry were recorded; in 76 patients, symptoms and obstruction at baseline spirometry were recorded without any concomitant evaluation of bronchial reversibility; in 96 patients, symptoms and bronchial obstruction at baseline spirometry with persistence of obstruction after salbutamol 400 μg were recorded in 72 patients, while in 24 subjects, the airway obstruction was fully reversible.

Among the 241 patients, chest X-ray was performed in 142 individuals. X-ray results were available in nine out of ten patients with clinical findings only. In the group of subjects with normal spirometry values, 33 had X-ray results, while 26 did not. Among the 73 patients with bronchial obstruction, 36 had X-ray results, and among those with fully reversible bronchial obstruction, nine patients had X-rays. In the 75 patients with baseline bronchial obstruction where the bronchodilator test was not performed, 55 had X-ray results.

In 28 (11.6%) subjects, X-ray allowed the ES to suspect a disease other than COLD (lung fibrosis, bronchiectasis, lymphangioleiomyomatosis [LAM], cystic fibrosis, sarcoidosis, Kartagener syndrome).

In a total of 208 out of 241 (86.3%) cases, the ES correctly identified patients with COLD and excluded individuals without COLD. In 131 (54.35%) cases, the ES identified a condition fitting with COLD in a very probabilistic manner. In this group, the expert evaluation brought to COPD diagnosis in 111 cases and severe asthma in 20 cases. In 77 (31.95% of the total) cases, both the ES and the experts excluded a COLD. In 27 (11.20%) subjects, the experts posed a final diagnosis of COLD, even if the X-ray analysis showed a pattern different from COLD (ie, post TB fibrosis). Five subjects (2.07%) had incorrect data (ie, absence of spirometry obstruction at screening and fixed obstruction shown by the expert). Therefore, the overall accuracy of the ES was 97.50%.

Chi-square analysis showed a significant association between the relative risk defined by the ES according to the weights defined during the phase of ES development and the specialist judgment (Table 3). The sample was not sufficiently large to evaluate the relative weight of each parameter of the tool (ie, spirometry, bronchodilation test) in ameliorating the accuracy of the diagnosis-supporting tool.

Discussion
Epidemiologic data constantly point out that the health resources and the physicians’ attempts to detect COLDS unable to predict the hidden prevalence of such conditions, as patients may not see the doctor until the disease is in an advanced stage. A COLD is usually diagnosed late because patients may adapt to their limiting condition or physicians may not properly detect the respiratory symptoms until lung function becomes severely impaired, sometimes below a half of normal values. As a consequence, up to 70% of the COLD population remains undiagnosed. Considering the limited resources for preventive medicine available in many countries, widespread programs of COLD detection are forced to adopt the strategy of an optimal cost-effectiveness ratio.

The ES for the diagnosis of COLD described in this work (COLD_m) was aimed at supporting the diagnosis also in primary care. The results of the validation analysis provided an ES accuracy of 86.3%. However, by removing confounding events and diagnoses that cannot be made based on only the questionnaire used in this study, the overall accuracy of the ES was 97.6%. While it cannot be excluded that a better accuracy might be obtained by adding further

### Table 3 Consolidated results in 231 unselected samples

<table>
<thead>
<tr>
<th>Score</th>
<th>Relative risk (%)</th>
<th>Number of patients</th>
<th>COLD</th>
<th>No COLD</th>
<th>% actual</th>
<th>% expected</th>
<th>Concordance</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–60</td>
<td>&lt;10</td>
<td>51</td>
<td>5</td>
<td>46</td>
<td>9.8</td>
<td>10</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>61–115</td>
<td>&lt;40</td>
<td>67</td>
<td>24</td>
<td>43</td>
<td>35.8</td>
<td>40</td>
<td>Ok</td>
<td></td>
</tr>
<tr>
<td>116–160</td>
<td>&gt;70</td>
<td>66</td>
<td>49</td>
<td>17</td>
<td>74.2</td>
<td>70</td>
<td>Ok</td>
<td>0.48</td>
</tr>
<tr>
<td>160–200</td>
<td>&gt;85</td>
<td>47</td>
<td>40</td>
<td>7</td>
<td>85.1</td>
<td>85</td>
<td>Ok</td>
<td>0.92</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>231</td>
<td>118</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: COLD, chronic obstructive lung disease.
variables, e.g., those reflecting the second-level lung function assessment, the parsimonious model proposed here proved to be very robust.

It is well known that COLD diagnosis (in particular, the distinction between COPD and severe asthma) cannot be made starting from clinical signs and symptoms, spirometry and X-ray data, as it requires more sophisticated diagnostic tools. The percentage accuracy of the system described here seems promising for a future use of COLD in different settings, including primary care. Obviously, this ES is not intended as a substitute for the clinician’s role; on the contrary, it may actually serve as an additional tool when COLD is suspected or in raising the suspicion.

Others have developed tools based on AI in the field of respiratory diseases. For example, a powerful ES was developed for the evaluation of spirometry data in the context of the clinic. In this study, the knowledge was built based on a single expert, while in the present work, a panel of experts was involved in the pilot validation study for the definition of different weights. In addition, data used to feed the inferential engine were second-level functional tests, while the aim of the present study was to suggest suspicion of COLD during the first screening visit of the patient. A further sophisticated study described the use of powerful artificial neural networks for the implementation of an ES suitable to suggest the diagnosis of COPD, TB, asthma and pneumonia. However, in this work, the validation was based on 60 cases, a number that seems too small for validating four different diseases. COLD instead was supported by an extensive clinical validation made on a collection of 241 cases, resulting in a robust and accurate tool. In addition, the ES approach has several strengths: 1) the results are more consistent and traceable compared to an individual’s assessment; 2) it allows gaining productivity and performance; 3) it is less expensive and quicker than referring to an expert, and 4) it can be made available anytime and anywhere.

One of the limitations lies in the fact that the rules were proposed by a panel of experts and need validation in larger samples. However, the system has the capability of being constantly updated, remaining efficient under the constant revision process.

To our knowledge, this is the first example of the development of an ES allowing establishment of probability of suffering from a COLD in a single individual. The current findings carry important clinical implications. The high accuracy of this tool allows moving to the implementation of the diagnosis-supporting ES in primary care settings. As already mentioned, the ES does not allow making of a definite diagnosis of COLD or to discriminate among different chronic obstructive diseases. In the hands of general practitioners (GPs), the ES can quickly contribute to identify subjects who are candidates for suffering from COLD. In other words, the high suspicion of COLD can be obtained in the office during general consultation, not requiring in this stage any supplemental test. This is of great importance in the real-life scenario, in which availability of time and specific skills may be poor. In addition, local and regional health systems suffer from the lack of standardized and organized collaborative pathways between GPs and pulmonologists, delaying lung function evaluation and consultation by the specialist.

COLD has been conceived by specialists as a support also for primary care physicians, thus fostering the collaboration between these two entities. Most importantly, this tool is designed to unveil the unexplored milieu of COLD in the general population, by supporting the GPs in raising the suspicion quickly and reliably.

Conclusion
Further investigations on larger samples are required to confirm and expand the current results, thus testing the performance of the ES in real-life primary care contexts.

Acknowledgment
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure
The authors report no conflicts of interest in this work.

References