Feasibility of spirometry in primary care to screen for COPD: a pilot study

Violaine Giraud1,2
Alain Beauchet3
Thierry Gomis4
Thierry Chinet1,2

1 AP-HP, Department of Pneumology and Thoracic Oncology, Ambroise-Paré Hospital, 2 UFR Paris Ile-de-France Ouest, Versailles Saint Quentin-en-Yvelines University, 3 Public Health Department, Unité de Recherche Clinique, Hôpitaux Universitaires Paris Ile-de-France Ouest, 4 Boulogne-Billancourt, Paris, France

Background: COPD is a frequent but underdiagnosed disease whose diagnosis relies on the spirometric demonstration of bronchial obstruction. Spirometry use by general practitioners could represent the first line in COPD diagnosis.

Objective: Because duration of spirometry is retarding its development in primary care, we decided to measure the time it requires in the primary-care context in France.

Methods: Ten volunteer general practitioners were trained during two 3-hour theoretical and practical continuing education sessions. Then, from October 2013 to May 2014, they included patients without any known respiratory disease but at risk of developing COPD (age: ≥40 years, smoker: ≥20 pack-years). The duration of spirometry and its quality were evaluated according to the following acceptability criteria: 1) expiration ≥6 seconds or reaching a plateau; 2) good start with an early peak flow, curve peaked on top and not flat; 3) no artifacts; and 4) reproducibility criteria, ie, forced expiratory volume in 1 second and forced vital capacity differences between the two best spirometry curves ≤0.15 L. Quality of the spiograms was defined as optimal when all the criteria were met and acceptable when all the criteria were satisfied except the reproducibility criterion, otherwise, it was unacceptable.

Results: For the 152 patients included, the 142 assessable spirometries lasted for 15.2±5.9 minutes. Acceptability criteria 1–3, respectively, were satisfied for 90.1%, 89.4%, and 91.5% of patients and reproducibility criterion 4 for 56.3%. Quality was considered optimal for 58.5% of the curves and acceptable for 30.2%.

Conclusion: The duration of spirometry renders it poorly compatible with the current primary-care practice in France other than for dedicated consultations. Moreover, the quality of spirometry needs to be improved.

Keywords: spirometry workshop, general practitioners, COPD

Introduction

COPD is a common pathology and, according to the World Health Organization, will be the third leading cause of death by 2025.1 In France, the prevalence of COPD is in ~5%–10% of adults >45 years old.2 Its human and social impact is marked, affecting 3.5 millions of individuals, with >40,000 new cases diagnosed annually, >100,000 hospitalizations, and >16,000 deaths.3 Moreover, COPD is expensive, with direct costs in 2005 estimated at 3.5 billion euros, ie, 3.5% of health care costs.4 The underdiagnosis of COPD is high, and two-thirds of the patients are never diagnosed.5–9 The benefits of early diagnosis of COPD remain the subject of debate. However, numerous arguments support COPD screening, most notably because forced expiratory volume in 1 second (FEV1) declines rapidly at the early stage,10,11 effective treatments exist to diminish dyspnea and the frequency of exacerbations,12 and outcomes of attempts to quit smoking might improve after diagnosis.13,14 In a Finnish global screening program, early COPD management was able to lower the number of hospitalizations by ~40%.15 Furthermore,
the diagnosis of COPD at the time of hospitalization, as is
the case for ~35% of the patients admitted for the first COPD
exacerbation, evidently seems late.16

The organization of early COPD screening has not been
formalized, even though the Global Initiative for Chronic
Obstructive Lung Disease (GOLD) Group recommends
screening of all smokers >40 years old.12 The diagnosis of
COPD relies on the spirometric demonstration of airflow
obstruction (AO). General practitioners (GPs) are first in line
to screen for COPD, especially because the patient is neither
couraged nor inclined to consult a pneumologist. Notably,
in the French ZEPHIR study, >40% of the patients at risk for
COPD did not go to a pneumologist for spirometry, despite
being prescribed by their GP.17 To counter this reluctance,
spirometry in primary care was developed in some European
countries. Nonetheless, its use in this context remains very
limited in many countries, including France. The obstacles
are many: lack of time to conduct the test, absence of GP
interest, insufficient training, difficulties in choosing the
instruments, etc.

We decided to undertake a study on the feasibility
of spirometry screening for COPD in primary care in
France. The time factor was often raised as limiting
the implementation of new techniques or strategies in general
medicine, where the estimated mean consultation duration is
15 minutes.18 Thus, after two education sessions dedicated
to spirometry in the framework of COPD screening, we
examined spirometry duration.

Materials and methods
The study was approved by the Saint-Germain-en-Laye
Ethics Committee and by the Comité Consultatif sur le
Traitement de l’Information en matière de Recherche dans
le domaine de la Santé.

Recruitment and training of the
participating GPs
Ten volunteer GPs working in the suburbs west of Paris were
trained by two pneumologists during two 3-hour continuing
education sessions. During the first session, the study pro-
tocol was explained and the GPs received theoretical and
practical training in spirometry and were provided with a
portable spirometer (Spirodoc® coupled to WinspiroPRO®
software; MIR, Langlade, France). This instrument allows
direct visualization of the flow–volume curve and the expira-
tion time. A beep is emitted after 6 seconds of expiration. The
GPs were asked to come to a second session with five test
results to pursue their training. Before and after the sessions,
a 12-question questionnaire evaluating each participant’s
knowledge of spirometry, its acceptability and reproducibility
criteria, and the diagnostic criterion of AO was completed
and graded on 100 points. Moreover, after the inclusion of
the first five patients, a pneumologist instructor’s evaluation
of the quality of the first five spirometry curves was sent to
each participating GP.

Patients
From October 2013 to May 2014, each GP aimed to recruit
20 consecutive patients at risk of developing COPD. The inclu-
sion criteria were smoking >20 pack-years, age >40 years,
and mastery of French. The exclusion criteria were an acute
respiratory episode during the 4 weeks preceding inclusion,
diagnosed chronic pulmonary disease, and inability to perform
spirometry. The patient gave his consent after oral and written
information, then the GP ran the spirometry with at least three
determinations of slow vital capacity (VC) and three flow–
volume curves. The GP then had to decide whether the test
was interpretable and then about the possible presence of AO.
The quality of the flow–volume curves was assessed centrally
by a pneumologist. No reversibility test was planned.

Evaluation criteria
The principal assessment criterion was the total duration of
the spirometric examination, including explanations, mea-
sured with the computer’s clock. The secondary end points
were determination of the quality of the curves, spirometry-
interpretation concordance (interpretability of the curves and
AO diagnosis) between the GP and the pneumologist, and
questionnaire grades before and after the training sessions.
A pneumologist evaluated the curves for acceptability and
reproducibility. The acceptability criteria were 1) expira-
tion lasting at least 6 seconds or reaching a plateau on the
time–volume curve; 2) good start with an early peak flow,
curve peaked on top and not flat; and 3) absence of cough
artifacts or glottis closing. Reproducibility was defined as
FEV₁ and forced vital capacity (FVC) differences ≤0.15 L
on the two best spirometry curves.19 Curve quality was con-
sidered optimal when all the acceptability and reproducibility
criteria were met and acceptable when all the acceptability
criteria were fulfilled but not the reproducibility criterion
and the best curve was normal or indicated AO; otherwise,
it was uninterpretable.

Evolution of participants’ spirometry knowledge was
assessed based on the questionnaire grades obtained before
and after evaluating the training sessions, with a maximum
grade of 100.
Statistical analyses
Quantitative variables are expressed as mean ± standard deviation and range; qualitative variables are expressed as n (%). Two means were compared with Student’s t-test and Wilcoxon signed-rank test, whereas several means were compared with Kruskal–Wallis test. GP–pneumologist concordance on the possibility to interpret the spirometry curve and AO diagnosis was ascertained with Cohen’s kappa coefficient.

The effect of experience on spirometry duration was evaluated by comparing the first three examinations with spirometries 8, 9, and 10. A center effect on the duration was assessed by comparing each center’s spirometries 8, 9, and 10.

Statistical analyses were computed using R software version 3.0.2. A P-value <0.05 defined significance.

Results
Among the ten GPs participating in this study, two were female. The mean age of GPs was 55 years (39–64 years) and seven were private practice GPs who also trained GP residents. Two of them had been trained in spirometry at least 10 years earlier, and one had a spirometer but had never run the test in his practice before the study. The mean mark for questionnaire evaluating spirometry knowledge raised from 44/100 before to 80/100 after the training sessions (P=0.01).

During the 8-month inclusion period (October 2013–June 2014), the GPs recruited 152 patients; five refused to participate. Among the 147 spirometries performed, five were excluded from the analysis because of information missing on the duration of the examination. The characteristics of the 142 patients included are given in Table 1. A mean of 9±6 (range: 1–24) patients were recruited per center.

The mean duration of spirometry was 15.2±5.9 (5–37) minutes: 15.0±6.1 minutes for males and 15.4±5.8 minutes for females (P=0.70). Eighty-six percentage of the spirometries lasted for ≤20 minutes. No association was found between spirometry duration and age (P=0.17) or AO diagnosis (n=41, P=0.21). No center effect was found for the eight centers that performed at least ten spirometries. No experience effect was noted for the spirometry duration comparing the first three spirometries versus spirometries 8, 9, and 10 (P=0.42).

Analysis of the 142 flow–volume curves by two expert pneumologists showed that 58.5% of these curves were of optimal quality, 30.2% were acceptable, and 11.3% uninterpretable (Table 2). Reproducibility, defined as FEV₁ and FVC differences between the two best spirometry curves ≤0.15 L, was the least frequently attained quality criterion that was met on only 56.3% of the curves. GP–pneumologist concordance for the possibility to interpret the spirometry curve was moderate (κ=0.40) and good for AO diagnosis (κ=0.62). The GPs did not attempt interpretation of the test in four of the 13 spiromgrams that they did not consider interpretable, and in ten of 15 spiromgrams for which they did not assess interpretability at all.

AO, defined as prebronchodilator FEV₁/FVC <70%, was diagnosed in 41 (29%) patients, among whom FEV₁ values were >80% of predicted for 68% (GOLD classification stage I) and between 50% and 80% of predicted for 32% (GOLD classification stage II).

Discussion
The results of this study showed that, after 6 hours of continuing education devoted to spirometry, the mean duration of the test performed by participating GPs was ~15 minutes. Approximately 60% of their spirometries were of optimal quality, and 30% additional tests were considered acceptable.

Table 2 Central evaluation of the quality of the 142 spirometry curves

<table>
<thead>
<tr>
<th>Curve-quality criterion</th>
<th>N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td></td>
</tr>
<tr>
<td>Expiration ≥6 s or reached a plateau</td>
<td>129 (90.1)</td>
</tr>
<tr>
<td>Good start</td>
<td>127 (89.4)</td>
</tr>
<tr>
<td>No artifact</td>
<td>130 (91.5)</td>
</tr>
<tr>
<td>Reproducibility</td>
<td></td>
</tr>
<tr>
<td>FVC difference between two best curves ≤0.15 L</td>
<td>93 (65.5)</td>
</tr>
<tr>
<td>FEV₁ difference between two best curves ≤0.15 L</td>
<td>108 (76.1)</td>
</tr>
<tr>
<td>FEV₁ and FVC differences between two best curves ≤0.15 L</td>
<td>80 (56.3)</td>
</tr>
<tr>
<td>Interpretation of quality</td>
<td></td>
</tr>
<tr>
<td>Optimal: acceptability and reproducibility criteria met (≤0.15 L)</td>
<td>83 (58.5)</td>
</tr>
<tr>
<td>Acceptable: acceptability criteria but not reproducibility criteria met; last curve normal (FEV₁ and FVC ≥80%, FEV₁/FVC ≥70%)</td>
<td>35 (24.6)</td>
</tr>
<tr>
<td>Acceptable: acceptability criteria but not reproducibility criteria met; last curve with AO (FEV₁/FVC &lt;70%)</td>
<td>8 (5.6)</td>
</tr>
<tr>
<td>Inacceptable</td>
<td>16 (11.3)</td>
</tr>
</tbody>
</table>

Abbreviations: AO, airway obstruction; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; s, seconds.
The GP–pneumologist concordance $\kappa$ was 0.40 for the possibility to interpret the spirometry curve and 0.62 for AO diagnosis.

The mean duration of a primary-care consultation in France is ~15 minutes. Hence, it seems illusory to include spirometry into a simple office visit. The consultation time must be prolonged and a dedicated appointment scheduled, meaning the patient has to come back or the test must be delegated to a technician in the office if the practice’s organizational structure permits. Should the physician conduct a reversibility test, as recommended by GOLD to diagnose COPD, the duration would be much longer. For COPD screening, spirometry duration could, in theory, be shortened because measuring slow VC is not mandatory but to do so raises the problem of insurance reimbursement. Indeed, in France, to be validated for payment, spirometry must include slow VC determination in addition to the flow–volume curve. A change in this policy in the context of COPD screening may facilitate the implementation of spirometry in primary care. Moreover, in a different setting where daily tests can be performed, in particular with dedicated technician or nurse, the mean duration of the spirometry might be shortened.

With only 58.5% optimal curves, quality might seem insufficient, but this rate is similar to those obtained in other primary-care studies. For example, quality control of 1,271 examinations done in routine practice in the Netherlands, using close similar acceptability and reproducibility criteria (acceptable variation between the two best FEV$_1$ and FVC up to 0.2 L), achieved an optimal quality rate of 38.8%. Analysis of 275 patients in a New Zealand study achieved 33% acceptable and 13% reproducible (same reproducibility definition) spirometry curves. When the spirometries were performed by dedicated technicians, the frequency of tests meeting American Thoracic Society criteria could be higher. In Austria, the frequency reached 71% of spirometries performed in primary care by certified technicians in a COPD-screening program. In a regional Canadian program to improve the management of asthma, the reference rate was 68% in a primary-care center with a dedicated technician, 91% in a university center, and 71% for ten recently trained health care technicians involved in asthmatic patient education. These conditions are very different from the routine practice of general medicine in France where it is very unlikely that spirometry will be performed by a trained technician or nurse.

The small number of tests in some centers might account for the relatively poor quality results. A learning curve is likely despite the fact that no experience effect has been found in this study, may be due to the small number of tests.

Patients with one normal flow volume curve are unlikely to have AO, even if the reproducibility criterion is not fulfilled. This would raise the rate of optimal or acceptable rate to 88.7%. Indeed, it is highly unlikely that a patient with a normal spirometry curve fulfilling acceptability but not reproducibility criteria would have COPD. One could argue that, in the context of screening for COPD, a normal spirometry would exclude the diagnosis and that any abnormality could prompt the GP to refer the patient for complementary explorations.

GP–pneumologist concordance on the possibility to interpret the curve was moderate ($\kappa$=0.40), while that for AO diagnosis was substantial ($\kappa$=0.62). This better agreement can undoubtedly be explained by a simpler notion based on the sole criterion, prebronchodilator FEV$_1$/FVC <0.70, whereas interpretability relies on the quality and reproducibility criteria, which are more complex. These findings underscore the importance of being able to turn to a pneumologist to interpret curves that are not obvious.

A brief continuing education program, consisting of two 3-hour evening sessions focused on performing spirometry, was chosen to be compatible with GPs’ work schedules. It was slightly longer than that proposed by Eaton et al, which consisted of 2 hours followed by a 90-minute remedial session 3 months later. The duration of our training program was about the same as the duration of the training delivered to ten technicians in the Canadian program to improve asthma management, 2 hours twice, guidance for the first six examinations, then an instructor available upon demand; this program achieved spirometry curves satisfying quality criteria and American Thoracic Society quality for 71% of the 472 examinations analyzed. However, it is much less ambitious than that proposed by the European Respiratory Society HERMES program, whose objective went beyond COPD screening. This program consists of initial training for 9–12 hours, writing a personal procedure for pulmonary function test, a questionnaire of 30 multiple-choice questions to assess knowledge on spirometry, a second training session of 7–10 hours, and finally, practical evaluation. To our knowledge, the HERMES project did not assess spirometry curves obtained in primary-care settings after completion of training. The continuing education dispensed in our study enabled the acquisition of knowledge, attested by the improved questionnaire grades and 60% optimal quality spirometry curves.

Our study had several limitations. The evaluation of the quality of curves was centralized but included, in addition to the objective quantitative criteria (eg, reproducibility,
FEV \(_1\)/FVC, and expiration time), qualitative parameters relatively subjective (eg, the notion of a good start and absence of artifacts).

The definition criterion retained for AO for this study was prebronchodilator FEV \(_1\)/FVC <70%, whereas the GOLD Group recommends measuring postbronchodilator FEV \(_1\)/FVC. However, in the context of screening, both a fixed FEV \(_1\)/FVC threshold instead of a cut off according to age, and pre-bronchodilator results are considered acceptable to suspect AO presence, despite the risk of over-diagnosis linked to the physiologic FEV \(_1\)/FVC decline with age and the risk of confusion with asthma without a post-bronchodilator reversibility test.\(^{25}\) Indeed, a simple, rapidly run, inexpensive test is needed for routine use in primary care and has been used in previous studies.\(^{26–28}\) Alternatively one could suggest to perform only a postdilator test, but adding a bronchodilator test, which requires 10–15 minutes between inhalation of the bronchodilator and the spirometry, would significantly increase the duration of the test.

This study was performed in the French context of primary-care organization and may not be generalizable to other countries. In some countries, spirometry in primary care is not performed by the physician but by other health professionals. In addition, the French insurance system requires the measurement of slow VC, which is not mandatory in other countries and obviously lengthens the duration of the spirometry.

Our study participants are not representative of French GPs. They practice in the Paris region, and seven of them also train future GPs, thus are particularly involved in educating medical residents. One might accord them greater motivation than usual to integrate a new technique into their practices.

The 20 patients expected per center was not reached, with only three centers including at least 20 patients during the 8-month recruitment period and three centers enrolling less than ten patients. The explanation given spontaneously by the GPs was the lack of time, thereby corroborating the difficulty of introducing into their practices an examination lasting at least the mean duration of a standard consultation. The lack of commitment of doctors to a COPD-screening program using spirometry in their office confirms a previously observed obstacle. In a Canadian study, 37% of the private practice offices approached refused to participate in spirometry implementation; the frequency reached 50% in an Austrian study.\(^{22,29}\) One of the factors explaining this difficulty of implanting spirometry in primary-care settings could be the lack of early training in performing and interpreting PFTs in medical school. A recent inquiry concerning 1,261 general medicine residents in France showed that only 4.3% of them thought that they had the competence required to conduct spirometry. Notably, 62.9% deemed this training necessary during their initial medical education.\(^{30}\)

Finally, the relatively short observation period precluded determination of future use of this tool by our participating GPs or the evolution of the quality of these examinations over time.

**Conclusion**

The duration of spirometry is too long to be compatible with the present French primary-care practice and would require a dedicated consultation or delegation of the task to other health professionals. However, the educational effort toward GPs was valuable, and the competence of the GPs was acceptable. Based on the results of this pilot study, we think that COPD screening cannot be improved simply by providing GPs with a spirometer and training sessions. A more global strategy remains to be devised.

**Acknowledgments**

The authors thank Drs Yves Bertin, Julien Uzzan, Michel Courteaud, Robert Lestrat, Marie-Laure Megret-Gabeaud, Pierre Sebag, Linda Terrasse, Didier Gérard, and Christian Weiss for their participation in the study, Drs Elise Nedelec and Sébastien Tubiana for their participation in the design of the study and collection of data, and CARDIF for its grant.

**Disclosure**

The authors report no conflicts of interest for this work.

**References**


