Technical and Functional Assessment of 10 Office Spirometers: A Multicenter Comparative Study

Giuseppe Liistro, Carl Vanwelde, Walter Vincken, Jan Vandevoorde, Geert Verleden, Johan Buffels and on Behalf of the COPD Advisory Board

Chest 2006;130;657-665
DOI: 10.1378/chest.130.3.657

Technical and Functional Assessment of 10 Office Spirometers*

A Multicenter Comparative Study

Giuseppe Liistro, MD, PhD; Carl Vanwelde, MD; Walter Vincken, MD, PhD, FCCP; Jan Vandevoorde, MD; Geert Verleden, MD, PhD; Johan Buffels, MD; on Behalf of the COPD Advisory Board†

Study objectives: **To investigate the technical properties and user friendliness of 10 office spirometers devoted for use in general practice, and to compare the results with standard diagnostic spirometers.**

Design: **Multicenter study.**

Setting: **Ten spirometer models were tested independently in three pulmonary function laboratories and by three general practitioners (GPs).**

Measurements: **The laboratories studied the technical quality of the office spirometers in terms of precision and agreement with standard spirometers, whereas the three GPs assessed their user friendliness. The spirometers tested were as follows: Spirobank (Medical International Research; Rome, Italy); Simplicity (Puritan Bennett; Pleasanton, CA); OneFlow (Clement Clarke International; Harlow, Essex, UK); Datospir 70 (Sibelmed; Barcelona, Spain); Datospir 120 (Sibelmed); SpiroPro (SensorMedics; Yorba Linda, CA); EasyOne (NDD; Zurich, Switzerland); MicroLoop (Micro Medical; Chatham, Kent, UK); SpiroStar (Medikro; Kuopio, Finland); and Pneumotrac (Vitalograph; Maids Moreton, Buckingham, UK). FVC and FEV_1 were measured in 399 subjects. User friendliness was assessed by the three GPs using a questionnaire.**

Results: **The precision of FEV_1 of the office and standard spirometers was comparable, but three office spirometers had > 200 mL limits of precision for FVC. Some devices presented a proportional difference on the FEV_1 with standard spirometers, underestimating the small values. The limits of agreements between standard and some office spirometers for FEV_1/FVC ratio was > 10%. The overall user friendliness was estimated as good.**

Conclusions: **The global quality and user friendliness of several office spirometers make them acceptable for the detection of COPD, although differences between the laboratory and some of the office spirometers values suggest that the misclassification rates may be increased when using some models of office spirometers.** ***(CHEST 2006; 130:657–665)***

Key words: comparative study; COPD detection; pulmonary function tests; spirometer; spirometry

Abbreviations: ATS = American Thoracic Society; CI = confidence interval; ERS = European Respiratory Society; ET = expert technician; GOLD = Global Initiative for Chronic Obstructive Lung Disease; GP = general practitioner; PFT = pulmonary function test; Sw = within-subjects SD

There is a clear need for early diagnosis of COPD.[1] This lung disease is one of the leading causes of mortality and disability in developed countries, and only smoking cessation has proven its efficacy in changing the natural evolution of COPD.[2] One major problem with early detection of COPD is the fact that smokers rarely complain even if they have dyspnea. However, lung function changes are often detectable > 10 years before onset of dyspnea at rest.[3] Therefore, according to a consensus statement from the National Lung Health Education Program,[4] the screening of asymptomatic at-risk populations should start from the age of 45 years. The screening by general practitioners (GPs) using

office spirometry can double the number of early diagnoses in COPD patients.[5] Therefore, primary care providers should be encouraged to perform good quality spirometry. For that, a good spirometer is as important as good training. Previous studies[6–10] were conducted to assess the quality of some handheld spirometers. These studies[6–10] usually compared one small spirometer with a conventional spirometer, and some found significant differences between the devices. However, small electronic spirometers constantly improve, and it is difficult to have a precise opinion on the quality of all the models present on the market. Moreover, the poor technical quality of some office spirometers may be an obstacle for routine clinical use and for the interchangeability of the measurements.[6] According to their manufacturers, the majority of the modern office spirometers do not require a daily calibration check. If this were true, it would represent an advantage because calibration checks are seldom carried out in general practice.[11] The aims of the present study were to assess the technical properties and the user friendliness of 10 spirometers devoted for use in general practice, and to compare the results with standard diagnostic spirometers.

Materials and Methods

In 2002, we asked the sales representatives of office spirometers available in Belgium to propose one or two models for use in general practice. The sales representatives of the office spirometers were first contacted through GlaxoSmithKline Belgium. Table 1 presents the 10 devices tested, their type of flow sensor, and their manufacturers.

*From the Pneumology Unit (Dr. Liistro), Cliniques Universitaires Saint-Luc, Brussels; Department of General Practice (Dr. Vanwelde), Université Catholique de Louvain, Louvain-la-Neuve; Pneumology Unit (Dr. Vincken) and Department of General Practice (Dr. Vandevoorde), Vrije Universiteit Brussel, Brussels; and Pneumology Unit (Dr. Verleden) and Department of General Practice (Dr. Buffels), Katholieke Universiteit Leuven, Leuven, Belgium.
†Members of the COPD Advisory Board are given in the Appendix.
This work was performed at Cliniques Universitaires Saint-Luc, Vrije Universiteit Brussel, and Katholieke Universiteit Leuven, Belgium.
Financial support was provided by GlaxoSmithKline Belgium, which hosted some of the investigator meetings.
The authors of this study have no financial interest in the subject of this article.
Manuscript received July 26, 2005; revision accepted March 5, 2006.

Correspondence to: Giuseppe Liistro, MD, PhD, Pneumology Unit, Cliniques Universitaires Saint-Luc, Avenue Hippocrate 10, 1200 Bruxelles, Belgium; e-mail: Giuseppe.liistro@clin.ucl.ac.be
DOI: 10.1378/chest.130.3.657

Setting

The study was divided into an in-laboratory study and an assessment of user friendliness of the office spirometers. The in-laboratory study was performed in the pulmonary function laboratories of three academic hospitals in Belgium.

For the user-friendliness assessment, the office spirometers were also presented to three GPs working in the general practice department of the three universities. The same office spirometers were used successively in the three laboratories and by the three GPs according to their availability.

Methods

The sales representatives of the office spirometers were asked to demonstrate the devices in each center. At the moment of the study, Sibelmed (Barcelona, Spain) was not represented in Belgium and the two Sibelmed spirometers (Datospir 70 and Datospir 120) were not demonstrated. We scrupulously followed the instructions of the manufacturers, in particular concerning the handling of the devices and the need for calibration checks. According to instruction manuals, the majority of the office spirometers do not need calibration.

In-Laboratory Study

The 10 office spirometers were tested independently in the three laboratories following the same protocol. The office spirometers were compared to standard diagnostic spirometers: Vmax 20C (SensorMedics; Bilthoven, the Netherlands; software: 5/2A, 2002) in two centers and Morgan TLC (Morgan Medical; Rainham, UK; software: Mdas 4.01, 1999) in one center. These devices are calibrated daily and used by expert technicians (ETs) only. The Vmax 20C, a flow-sensing spirometer, was calibrated with a 3-L syringe at three different speeds; the Morgan TLC, a volume-sensing device, using a single speed, was calibrated as recommended by the manufacturer. The Morgan TLC was checked daily for leaks. A log of calibration results and leaks checks was maintained.

In each laboratory, three ETs tested all the office spirometers. We first checked the ability of the ETs themselves to perform reproducible pulmonary function tests (PFTs), and the interchangeability of the results between the three centers. Only technicians able to blow five times successively with maximum variations of FEV_1 or FVC of 200 mL or 6% were selected. All were nonasthmatic nonsmokers and were free of respiratory symptoms. The ETs were asked to perform PFTs in the two other centers to verify interchangeability of the results. There was no significant difference between centers in absolute values and reproducibility ($p > 0.05$, by one-way analysis of variance; Fig 1).

Precision of the Spirometers: The ETs were asked to perform on the same day five successive forced expiratory maneuvers with the standard spirometers (one by center) and with the office spirometers to compare their reproducibility. After verifying the interchangeability of the results, we pooled the values and reported the within-subjects SD (Sw).

Agreement Between the Standard and Office Spirometers: In each center, PFTs were performed by healthy naive subjects and 10 COPD patients with the office spirometers and with the standard spirometers. The healthy subjects were nonsmokers, free of any respiratory symptoms, and members of the hospital staff. These tests were done in a random order to avoid a learning effect. We asked the three centers to select the patients to represent various degrees of severity of COPD (Global Initiative for Chronic Obstructive Lung Disease [GOLD] stages I to IV). We retained only the patients and subjects able to perform the PFT according to American Thoracic Society (ATS) quality

Table 1—*The Office Spirometers Tested, Characteristics of Their Flow Sensor, and Manufacturers*

Models	Flow Sensor	Software	Manufacturer, Country
Datospir 120	Heated Fleisch	Sibelmed W-20	Sibelmed, Spain
Datospir 70	Turbine	Sibelmed W-10	Sibelmed, Spain
EasyOne	Disposable ultrasonic	EasyWare	NDD, Switzerland
Microloop	Turbine	Spida 5	Micro Medical, United Kingdom
OneFlow	Nonheated, differential pressure	OneFlow soft 1.2	Clement Clarke, United Kingdom
Pneumotrac	Nonheated Fleisch	Spirotrac IV	Vitalograph, United Kingdom
Simplicity	Disposable plastic screen	None	Puritan Bennett, United States
Spirobank	Turbine	WinspiroPRO	Medical International Research, Italy
SpiroPro	Nonheated, differential pressure	SpiroPro for Windows	SensorMedics, United States
SpiroStar	Disposable plastic screen	Spiro2000 1.5.2	Medikro, Finland

criteria.[12] We compared the best values of FVC, FEV_1, and FEV_1/FVC from all acceptable tests for all pairs of devices (each office spirometer vs standard spirometer).

Because of the multicentric nature of the study and the fact that the same devices were tested successively in each center, the healthy volunteers and the COPD patients were different for each apparatus and by center. Each office spirometer was tested with 48 subjects ([three ETs + 3 healthy subjects + 10 COPD patients] × 3), a sample size comparable to previous studies.[9,10]

Assessment of User Friendliness

We developed a novel questionnaire to assess the user friendliness of office spirometers (to view Tables A and B in the on-line supplementary data, go to www.chestnet.org). The questionnaire completed by the three GPs covered the following: general properties and parameters of the software, quality of the patient administrative data, features of the display and automated quality control, comparison of successive tests in the same subject, use at home visits, and export facilities.

Analysis

The precision of the office spirometers was assessed by the Sw of FEV_1 and FVC obtained from five successive maneuvers done by the nine ETs with the office spirometers and with the standard spirometers.[13] The larger the Sw value, the lower the precision. The variance of each set of five measurements (FVC, FEV_1) was computed, and the within-subject variance of the measurements was obtained by averaging the nine variances. To obtain the 95% error limits, the square root of the within-subject variance was multiplied by 1.96. We used goals for within-session repeatability for FEV_1 and FVC of 200 mL, as our upper limits of precision.

The preliminary analyses revealed no difference between the populations studied, so the results of the PFTs done in the three centers were pooled for each office spirometer. The agreement and the bias between the standard and office spirometers were examined using Bland and Altman[14] analysis. The bias (mean difference between the office spirometers and the standard spirometers), its 95% confidence interval (CI), and the lower and upper limits of agreement between office spirometers and stan-

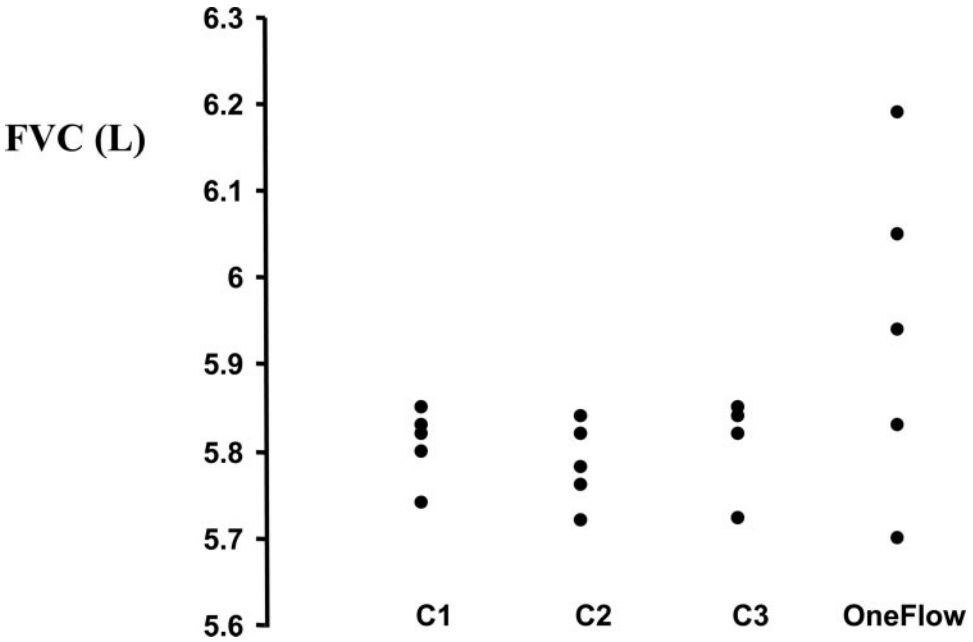

FIGURE 1. Plot of each set of five FVC measures in one ET in the three centers with the standard spirometers (C1 to C3) and with one office spirometer (OneFlow). The precision of FVC measurement was comparable between the three centers but differed with the office spirometer.

dard spirometers were reported. A statistical significant correlation ($p < 0.05$) indicates the presence of a proportional difference between the devices. We fixed the upper limits of acceptable bias at ± 100 mL according to the ATS accuracy criteria for diagnostic devices.[12] The acceptable limits of agreement between the office spirometers and the standard spirometers were 350 mL and 500 mL for FEV_1 and FVC, respectively. These limits were fixed according to the short-term coefficient of variability of FEV_1 and FVC measured in COPD patients.[15]

Results

In-Laboratory Study

A total of 399 different subjects (age range, 25 to 87 years; mean ± SD, 61.2 ± 14.6 years; 128 women) were studied in the three centers. The distribution of the 300 COPD patients according to the GOLD classification was as follows: stage 1, 10.3%; stage 2, 20.4%; stage 3, 59.0%; and stage 4, 6.3%. The nine ETs (three women) all had normal spirometric data (FVC, 4.68 ± 1.04 L; FEV_1, 3.80 ± 0.90 L).

Table 2 presents the precision of the standard diagnostic spirometers and each model of office spirometer. The precision was measured in ETs from the three centers (n = 9). The precision of FEV_1 was comparable between the standard and office spirometers, except for one device (Simplicity; Puritan Bennett; Pleasanton, CA), in which the limits of precision were > 200 mL. Three models—OneFlow (Clement Clarke International; Harlow, Essex, UK), EasyOne (NDD; Zurich, Switzerland), and Simplicity—showed precision limits > 200 mL for FVC.

Tables 3, 4 show the bias and the limits of agreement of each office spirometer for FVC, FEV_1, and FEV_1/FVC (n = 48). One spirometer presented a significant bias for FVC, four spirometers presented a significant bias for FEV_1, and five spirometers presented a significant bias for FEV_1/FVC. The biases remained within acceptable limits for FVC and FEV_1 (≤ 100 mL). The limits of agreement were different between the office spirometers according to the variable studied. Only three devices had acceptable limits of agreement for FVC (MicroLoop; Micro Medical; Chatham, Kent, UK; Pneumotrac; Vitalograph; Maids Moreton, Buckingham, UK; and Spiropro; SensorMedics; Yorba Linda, CA), and four devices had acceptable limits of agreement for FEV_1 (Datospir models 70 and 120, MicroLoop, and Pneumotrac). The Bland and Altman graphs of FVC, FEV_1, and FEV_1/FVC are submitted as on-line supplementary material.

Some office spirometers showed an apparent good agreement with the standard spirometers (low bias), but the visual inspection of the Bland and Altman plot revealed a proportional difference. In other words, there was a significant relationship between the mean difference and the average values, as indicated in Table 3 by a statistically significant coefficient of correlation. In the example of Figure 2, FEV_1 was underestimated for small values and overestimated for larger values. As shown in Tables 3 and 4, a proportional difference was observed in one device for FVC, in four devices for FEV_1, and in two devices for FEV_1/FVC.

Table 2—*The 95% Limits of Precision (1.96 × Sw) Obtained in Nine ETs With the Office and Standard Spirometers*

Devices	FVC	FEV_1
Standard	0.15	0.14
Datospir 120	0.18	0.14
Datospir 70	0.15	0.12
EasyOne	0.25*	0.19
Microloop	0.18	0.18
OneFlow	0.43*	0.15
Pneumotrac	0.18	0.13
Simplicity	0.34*	0.29*
Spirobank	0.15	0.15
SpiroPro	0.17	0.16
SpiroStar	0.19	0.17

*Limits of precision > 200 mL.

User-Friendliness Assessment

The pooled results of this survey are shown in the on-line supplementary material. The overall user friendliness of the tested devices was judged to be good. The information provided by the spirometers was in general very complete. The clusters of questions with strongly favorable answers were those about the general properties of the software, the completeness of the displayed information, and the automated tests of quality assurance. We noted intermediate positive answers for the following clusters: patient administrative data facilities, comparison of successive spirometrics and facilities for use at home visits. Rather unfavorable scores were given to the functions for export of data.

Discussion

We report the results of a multicentric study of 10 office spirometers compared according to both their technical and user-friendliness characteristics. The precision of FEV_1 measurement was good in the majority of these small spirometers and was comparable to the values obtained with the standard diagnostic devices. However, the repeatability of FVC was generally poorer, and the broad limits of agreement of FVC or FEV_1 observed between some office spirometers and the standard devices may preclude the interchangeability of the results.

Table 3—*Parameters From the Bland and Altman Analysis of the Office Spirometers

Variables	Bias (95%CI), L	Limits of Agreement	*r*	p Value
FVC				
Datospir 120	− 0.06 (− 0.13 to 0.01)	− 0.52 to 0.40	− 0.19	NS
Datospir 70	− 0.07 (− 0.16 to 0.02)	− 0.67 to 0.53	0.05	NS
EasyOne	0.07 (0.00 to 0.15)	− 0.43 to 0.57	− 0.15	NS
Microloop	− 0.03 (− 0.09 to 0.04)	− 0.47 to 0.42	0.08	NS
OneFlow	0.03 (− 0.12 to 0.19)	− 1.00 to 1.07	− 0.21	NS
Pneumotrac	− 0.03 (− 0.09 to 0.04)	− 0.46 to 0.40	0.36	< 0.05
Simplicity	0.03 (− 0.06 to 0.11)	− 0.57 to 0.62	− 0.24	NS
Spirobank	− 0.04 (− 0.12 to 0.04)	− 0.56 to 0.49	− 0.12	NS
SpiroPro	− 0.06 (− 0.12 to 0.00)	− 0.47 to 0.35	0.06	NS
SpiroStar	0.10 (0.02 to 0.19)*	− 0.49 to 0.70	− 0.27	NS
FEV_1				
Datospir 120	0.00 (− 0.04 to 0.05)	− 0.29 to 0.30	− 0.06	NS
Datospir 70	0.02 (− 0.02 to 0.05)	− 0.24 to 0.27	0.07	NS
EasyOne	0.08 (0.04 to 0.13)†	− 0.21 to 0.38	− 0.40	< 0.01
Microloop	0.06 (0.02 to 0.09)†	− 0.16 to 0.28	− 0.09	NS
OneFlow	0.07 (0.01 to 0.14)†	− 0.37 to 0.52	− 0.21	NS
Pneumotrac	0.08 (0.05 to 0.12)†	− 0.13 to 0.30	0.35	< 0.05
Simplicity	0.00 (− 0.09 to 0.08)	− 0.55 to 0.55	− 0.38	< 0.01
Spirobank	− 0.06 (− 0.11 to 0.00)	− 0.45 to 0.33	− 0.07	NS
SpiroPro	− 0.03 (− 0.08 to 0.02)	− 0.37 to 0.31	− 0.31	< 0.05
SpiroStar	− 0.08 (− 0.16 to 0.00)	− 0.63 to 0.47	− 0.10	NS

*NS = not significant.
†Bias significantly different from zero. A statistical significant correlation means that a proportional difference is present. The negative correlation coefficient implies that the office spirometer underestimates the small values whereas large values are overestimated.

All the office spirometers presented in this study received the label "meets ATS recommendations." This label supposes that the spirometers were checked by a series of predetermined flow-volume curves via a computer-driven piston pump.[12] In the standardization of spirometry published recently,[16] the ATS/European Respiratory Society (ERS) task force recommends that spirometers should be evaluated using a computer-driven mechanical syringe or its equivalent. The use of the 24 ATS standard waveforms for FVC and FEV_1 is recommended, but the recent ATS/ERS statement[16] does not require that these tests must be performed by an independent laboratory, such as the laboratory at LDS Hospital for example. We requested the results of these tests from the manufacturers of the office spirometers. Datospir models 70 and 120, Simplicity, OneFlow, Pneumotrac, and SpiroStar (Medikro; Kuopio, Finland) devices were tested by their manufacturers using 24 ATS standard waveforms, and the other devices were tested at LDS Hospital. All the spirometers tested at LDS Hospital using the 24 standard ATS waveforms had fewer than three accuracy errors and repeatability errors for FVC or FEV_1

Table 4—*Parameters From the Bland and Altman Analysis of the Office Spirometers for FEV_1/FVC

Variables	Bias (95%CI), %	Limits of Agreement	*r*	p Value
Datospir 120	0.1 (0.2 to 1.8)†	− 4.3 to 6.3	− 0.01	NS
Datospir 70	1.1 (0.0 to 2.3)	− 6.7 to 9.0	− 0.51	< 0.001
EasyOne	1.8 (0.6 to 3.0)†	− 6.4 to 10.0	− 0.26	NS
Microloop	2.3 (1.3 to 3.2)†	− 4.3 to 8.8	− 0.25	NS
OneFlow	2.0 (− 0.7 to 4.8)	− 16.2 to 20.3	− 0.07	NS
Pneumotrac	2.8 (1.8 to 3.7)†	− 3.8 to 9.4	− 0.24	NS
Simplicity	0.5 (− 1.1 to 2.0)	− 9.9 to 10.8	− 0.22	NS
Spirobank	− 0.9 (− 2.2 to 0.5)	− 9.7 to 7.9	0.10	NS
SpiroPro	0.4 (− 0.5 to 1.4)	− 5.9 to 6.7	− 0.43	< 0.01
SpiroStar	− 4.9 (− 6.1 to − 3.6)†	− 13.1 to 3.4	0.08	NS

*See Table 2 for expansion of abbreviation.
†Bias significantly different from placebo.

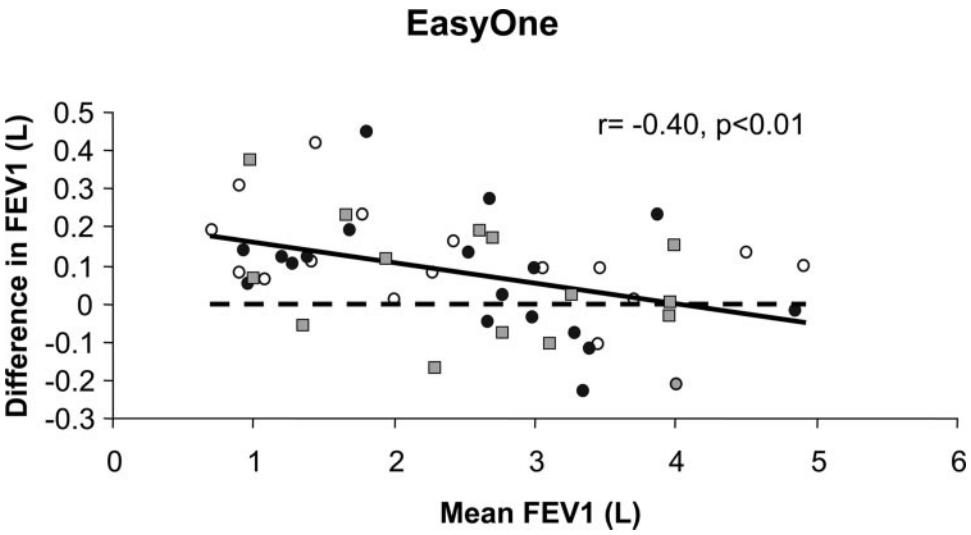

FIGURE 2. Bland and Altman plot of the relationship between mean FEV_1 and the difference in FEV_1 between the standard spirometers and the EasyOne. There was a significant linear relationship between these variables (straight oblique line), indicating a proportional difference between the devices. The origin of the points is identified (the three centers). For sake of clarity, the limits of agreement are not shown.

at ambient testing conditions. The OneFlow met the ATS criteria for monitoring devices but not for diagnostic spirometers. At variance with previous ATS quality criteria for spirometers, there is no longer distinction between diagnostic and monitoring devices in the recent ATS/ERS statement.[16] The same degree of accuracy is required for all devices, and corresponds to the former criteria (ATS) for diagnostic spirometers. Using the 24 ATS waveforms, the OneFlow had better accuracy results for FEV_1 than for FVC, and only the FEV_1 was within the criteria for diagnostic devices. Accordingly, in our study, the OneFlow performed better for FEV_1 measurement than for FVC. By contrast, the results of the SpiroStar tested by its manufacturer (Medikro) showed no error for accuracy and intradevice testing with FEV_1 and FVC, using the criteria for diagnostic devices, whereas this device presented large limits of agreement in our survey.

At variance with the 24 ATS waveforms, we did not use a fixed signal to test the devices, so the differences between *in vitro* and *in vivo* performances may be explained by the fact that different populations were studied. However, if we limit our analysis to the small sample (n = 9) of ETs who tested all the models, we can see that EasyOne presented unacceptable limits of precision for FVC. However, this device, tested by an independent laboratory (LDS Hospital), met the ATS (diagnostic devices) recommendations for accuracy and precision in measuring FVC and FEV_1, like the Spirobank, whereas the latter obtained acceptable limits of precision in our study.

The tests done with the computer-controlled mechanical syringe indicate that the spirometers working under ideal laboratory conditions, at least at ambient conditions, are reliable, but they do not guarantee that the same performances will be obtained in real-life conditions, as patients are tested throughout the day, as it was the case in our study. Therefore, we and other groups[17] propose that the spirometers should be tested both with waveform generators (on a bench) and with real patients.

The precision of FEV_1 measured by the office spirometers was comparable to that of the standard spirometers. Only one device showed unacceptable limits of precision (> 200 mL) for FEV_1 (± 0.295 L for the Simplicity).

All the office spirometers were less reproducible for the measurement of FVC when compared to the standard spirometers. FVC appears more difficult to measure by the small spirometers (Table 2). Three office spirometers measured FVC with a precision worse than 200 mL (Table 2).

Although statistically significant in some office spirometers, the biases observed for FEV_1 or FVC are probably not clinically relevant. Even if we consider the accuracy criteria of the ATS for diagnostic devices[12] as the limits of acceptable bias, all the office spirometers tested for FEV_1 and for FVC met this criteria. However, the SpiroStar had a bias

of – 4.9 for FEV_1/FVC (Table 4). This device underestimated FVC and consequently overestimated FEV_1/FVC ratio as compared to the diagnostic spirometers. Some stage 1 COPD patients could therefore be misclassified as stage 0 in that way.

The present data also show that the limits of agreements were generally larger for FVC than for FEV_1. We fixed the acceptable limits of agreement by analogy with the short-term coefficients of repeatability of FEV_1 (320 mL) and FVC (450 mL) measured in COPD patients.[15] These coefficients corresponded to the maximum absolute limit of agreement times 0.9 as computed using the Bland and Altman analysis with our data. Hence, the limits we chose were 350 mL (FEV_1) and 500 mL (FVC). Relatively wide limits of agreement were found with some office spirometers for FVC.

Underestimation of low FEV_1 and overestimation of high FEV_1 were also found with some office spirometers. This proportional difference may also result in an underestimation of the FEV_1/FVC ratio and COPD misclassification.

We devised a new questionnaire about the user friendliness of office spirometers. The results of this small and informative survey suggest large similarities among the majority of the spirometers. Almost all of them have an interface with a personal computer, can be connected to a printer, and have correct help functions. An important issue is the automated quality assessment.[17,18] Most of the devices check the reproducibility and the acceptability of the spirometric maneuvers. Many spirometers display immediately messages about the quality of the "blow." The software offers mostly a large choice of reference tables and provides an automated interpretation. According to some authors,[19,20] this feature is not without danger of misinterpretation for uninformed users. Large differences were encountered for the facility to use during home visits. Some devices are readily excluded by their large size. Some spirometers display a flow-volume curve on a small screen on the device itself. The need for calibration of the spirometers is a controversial aspect.[17,18] It is not realistic to expect a GP to handle on a daily basis a 3-L calibration syringe and perform a calibration check. Most manufacturers claim that their device does not need any further calibration or even calibration checks. In fact, the majority of the office spirometers cannot be calibrated by the user, and calibration can only be checked. The new ATS/ERS statement[16] on standardization of spirometry, published after the first submission of this study, requires that flow spirometer calibration be checked daily using different speeds of injection to verify that the spirometers measure accurately across a range of flows. More studies are needed to verify the stability of each device over months to years of use.

Another important property is the integration of the spirometric data into the patient data files. The management of the medical data is "core business" for the GP. However, technical specifications may vary largely from one country to the other. In the Belgian situation, the I-Med standard in XML (extensible markup language) format is becoming increasingly important as an interface between electronic medical data files and external information. Extensible markup language format is designed especially for Web documents and to enable the transmission, validation, and interpretation of data between applications and between organizations. There is plenty of room for improvement in international standardization in this domain, as it was proposed in the recent ATS/ERS guidelines for standardization of spirometry.[16]

We did not compare the quality and the cost of the mouthpieces, nor the facility of disinfecting the parts in contact with the exhaled air. We do not have any information on the long-term solidity of the devices, nor on their resistance to shocks during transportation, and to changes in humidity and temperature. We did not make any price/quality assessment. Last but not least, we have no feedback about the aftersales service of the different spirometer representatives. This could be of major importance in real-world circumstances.

Other limitations of the study should also be considered. Of course, all the office spirometers available on the market could not be tested during this study. However, we included office spirometers coming from Europe and the United States, showing different technical characteristics, and our sample was sufficiently diversified. It is noticeable that some of these devices (both hardware and software) are frequently upgraded or changed, so that the same brands sold now may differ from the devices tested. The devices tested were those available in Belgium in 2002.

Another issue is our assumption that standard diagnostic spirometers are accurate and precise. Indeed, if one standard diagnostic spirometer presents a proportional error, it will make the office spirometer look like it has a proportional error, even if this latter is accurate. In a multicentric study like the present one, the risk that all the reference spirometers are inaccurate is relatively low. Moreover, the office spirometers were compared to two types of laboratory spirometers: volumetric and hot wire. If only one reference spirometer presented a proportional error, the relation between the average and the difference of FEV_1 would not be significant. The clinically acceptable limits for the bias, the

precision, and the limits of agreement of FEV_1 and FVC were fixed by analogy with the quality criteria of the ATS for diagnostic spirometers and with the coefficients of repeatability observed in COPD patients, because these limits are not available in the literature.

At variance with standard spirometers, the majority of the office spirometers were not calibrated during the study (Table 1). This may seen somewhat surprising in a laboratory study, but we wanted to respect the instructions of the vendors/manufacturers also in terms of calibration. We must keep in mind that these devices are devoted for the use by GPs and not for laboratory purposes.

The limits of precision of the office spirometers reported in this study refer to short-term repeatability. In real circumstances, the repeatability may decrease after several months if the users do not check regularly the spirometers, especially with screen-type pneumotachographs that commonly become clogged with secretions.

In conclusion, the availability of accurate, precise, and easy-to-use office spirometers is a condition for widespread use of spirometry in general practice. This survey shows that several office spirometers have a high score of user friendliness, but their instrumental properties are variable. All but one of the devices had excellent reproducibility for FEV_1. Moreover, we observed low biases and acceptable limits of agreement for FEV_1 in four office spirometers. However, several devices presented a proportional difference for FEV_1, leading to underestimation or overestimation of results. Furthermore, the lack of precision of FVC and unacceptable limits of agreement between FEV_1 and FVC may lead to a risk of misclassification of patients according to the GOLD criteria for COPD.

For all the office spirometers tested, more attention should be paid to improve the accuracy and precision of FVC because it limits the interchangeability of the spirometric data between the GP and the PFT laboratory. Import and export facilities of the spirometric data should also be improved. More investigation is needed to assess the overall quality of these rapidly changing instruments. Finally, we confirm that *in vivo* testing of the spirometers gives additional information to the strict *in vitro* benchmark study. This study could be used as model for the users and the manufacturers wishing to test new devices.

ACKNOWLEDGMENT: The authors thank the technicians of the pulmonary function laboratories: J. Onrubia, R. Vanleemputten, D. Schuermans, Y. Dewandeleer, F. Rochette, R. Juste, N. Celis, and S. Van Poyer. One of the ETs is a coauthor of this article (G.L.); the vendors: Mrs. Janssens, Mr. Du Four, Mr. Knez, Mr. Reyskens, Mr. Moens, Mr. Moreno, Mr. Servais, and Mr. Tastenoy; and our sponsor: GlaxoSmithKline Belgium (Mr. Dumoulin, Mr. Thierry, and Mr. Cleymans).

Appendix

The COPD Advisory Board is composed of Belgian pneumologists and GPs with the sponsorship of GlaxoSmithKline Belgium. Members of the COPD Advisory Group, Detection and Diagnostic Cell include Professor Yernault (deceased), Professor Louis, Professor Vincken, Professor Rodenstein, Professor Demedts, Dr. Gillard, Dr. Dierckx, Dr. Coolen, Dr. Robience, Mr. Schuermans, Dr. Cnockaert, and Mr. Rochette.

References

1 NHLBI/WHO. Global Initiative for Chronic Obstructive Lung Disease workshop report (updated 2004). Available at: www.goldcopd.org/. Accessed January 31, 2005
2 Anthonisen NR, Connett JE, Kiley JP, et al. Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV_1: the Lung Health Study. JAMA 1994; 272:1497–1505
3 Enright PL, Hyatt RE, eds. Office spirometry: a practical guide to the selection and use of the spirometers. Philadelphia, PA: Lea & Febiger, 2003
4 Ferguson GT, Enright PL, Buist AS, et al. Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program. Chest 2000; 117:1146–1161
5 Buffels J, Degryse J, Heyrman J, et al. Office spirometry significantly improves early detection of COPD in general practice: the DIDASCO study. Chest 2004; 125:1394–1399
6 Rebuck DA, Hanania NA, Durzo AD, et al. The accuracy of a handheld portable spirometer. Chest 1996; 109:152–157
7 Wiltshire N, Kendrick AH. Evaluation of a new electronic spirometer: the Vitalograph Escort Spirometer. Thorax 1994; 49:175–178
8 Jones KP, Mullee MA. Lung function measurement in general practice: a comparison of the Escort Spirometer with the Micromed Turbine Spirometer and the Mini-Wright peak flow meter. Respir Med 1995; 89:657–663
9 Maree DM, Videler EA, Hallauer M, et al. Comparison of a new desktop spirometer (Diagnosa) with a laboratory spirometer. Respiration 2001; 68:400–404
10 Swart F, Schuurmans MM, Heydenreich JC, et al. Comparison of a new desktop spirometer (Spirospec) with a laboratory spirometer in a respiratory out-patient clinic. Respir Care 2003; 48:591–595
11 Dowson LJ, Yeung A, Allen MB. General practice spirometry in North Staffordshire. Monaldi Arch Chest Dis 1999; 54: 186–188
12 American Thoracic Society. Standardization of spirometry, 1994 update. Am J Respir Crit Care Med 1995; 152:1107–1136
13 Bias and measurement error. In: Daly CR, Bourke SC, eds. Interpretation and uses of medical statistics. Oxford, UK: Blackwell Science, 2000; 381–421
14 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307–310
15 Enright PL, Connett JE, Kanner RE, et al. Spirometry in the Lung Health Study: II. Determinants of short-term intraindividual variability. Am J Respir Crit Care Med 1995; 151(2 pt 1):406–411
16 Miller MR, Hankinson J, Brusasco V et al. Standardisation of

spirometry. Eur Respir J 2005; 26:319–338
17 Townsend MC, Hankinson JL, Lindesmith LA, et al. Is my lung function really that good? Flow-type spirometer problems that elevate test results. Chest 2004; 125:1902–1909
18 Enright PL, Kaminsky DA. Strategies for screening for chronic obstructive pulmonary disease. Respir Care 2003; 48:1194–1201
19 Eaton T, Withy S, Garrett JE, et al. Spirometry in primary care practice: the importance of quality assurance and the impact of spirometry workshops. Chest 1999; 116:416–423
20 Schermer TR, Jacobs JE, Chavannes NH, et al. Validity of spirometric testing in a general practice population of patients with chronic obstructive pulmonary disease (COPD). Thorax 2003; 58:861–866