Testing Spirometers: Are the Standard Curves of the American Thoracic Society Sufficient?

Quentin Lefebvre Eng, Thomas Vandergoten Eng, Eric Derom MD PhD,
Emilie Marchandise PhD, and Giuseppe Liistro MD PhD

BACKGROUND: The performance of spirometers is often measured only under ideal conditions, with a mechanical simulator reproducing the expiratory standard American Thoracic Society (ATS) curves generated by a computer. Studies have questioned the value of these results in real-life conditions. The aim of this study was to evaluate the accuracy and precision of 5 office spirometers with a flow-volume simulator using the ATS curves and using flow-volume curves obtained from patients. METHODS: We measured the FVC, peak expiratory flow, and FEV_1 by simulating different dynamic waveforms applied by a computer-driven syringe, the Hans Rudolph flow-volume simulator. In addition to testing standard curves recommended by the ATS, we also tested curves obtained with subjects. RESULTS: The precision of the office spirometers was good and comparable using the standard ATS curves. One device presented the best performances in terms of accuracy and precision according to the ATS recommendations, but we observed significant biases in all devices with Bland-Altman analysis, particularly with the curves obtained from subjects with severe COPD. CONCLUSIONS: The global quality of most spirometers makes them acceptable for the detection of pulmonary diseases. However, we demonstrated accuracy issues not shown by the standard testing procedure. We propose to improve the testing of spirometers by implementing more realistic flow-volume curves and to refine the analysis of the results. *Key words: chronic obstructive pulmonary disease; COPD; benchmarking; instrumentation; quality control; spirometry; technology assessment.* [Respir Care 2014;59(12):1–•. © 2014 Daedalus Enterprises]

Introduction

To extend and facilitate the diagnosis of respiratory diseases, office spirometers were introduced on the market that are specially adapted to the practice of general medicine.[1] A major objective of these office spirometers is to diagnose obstructive lung diseases such as asthma and COPD.[2]

The performance of spirometers is often assessed under ideal conditions with a mechanical flow-volume simulator designed to produce expiratory standard American Thoracic Society (ATS) curves, the standard test waveform set,[3] generated by a computer following a standardized procedure, outlined by the ATS/European Respiratory Society (ERS) in 2005.[4] The validity of these results has previously been studied by comparing the performance of some office spirometers with standard laboratory devices in patients tested under real-life conditions. We showed

Mr Lefebvre, Mr Vandergoten, and Dr Marchandise are affiliated with the Institute of Mechanics, Materials and Civil Engineering, Louvain School of Engineering, Université Catholique de Louvain, Louvain-la-Neuve, Belgium. Dr Derom is affiliated with the Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium. Dr Liistro is affiliated with the Institut de Recherche Expérimentale et Clinique, Pôle de Pneumologie, ORL et Dermatologie, Cliniques Universitaires Saint Luc, Université Catholique de Louvain, Brussels, Belgium.

Dr Liistro presented a version of the paper at the American Thoracic Society 2011 International Conference, held May 13–18, 2011, in Denver, Colorado.

The authors have disclosed a relationship with the MEC Group.

Correspondence: Giuseppe Liistro MD PhD, Pneumology Department, Cliniques Universitaires Saint Luc, Université Catholique de Louvain, 10 avenue Hippocrate, 1200 Brussels, Belgium. E-mail: giuseppe.liistro@uclouvain.be.

that many of the investigated spirometers revealed defaults, despite ATS certification.[5] We therefore hypothesized that the standard ATS curves were not representative of some of the features routinely seen in flow-volume loops originating from patients.

The aim of this study was to evaluate the accuracy and precision of 5 different office spirometers using the Hans Rudolph flow-volume simulator (Hans Rudolph, Shawnee, Kansas) to generate both standard ATS curves and flow-volume curves obtained from healthy volunteers and subjects with respiratory disorders. We also aimed to refine the analysis of the data using the Bland-Altman method.[6]

Methods

We asked the sales representatives of office spirometers available in Belgium to demonstrate one production model of their devices and to make them available for research. Table 1 presents the 5 devices tested. They were all reported by the manufacturer to comply with the ATS/ERS guidelines.

All devices report the values at body temperature and pressure saturated (BTPS) condition. In this study, we analyzed the performances of the spirometers only during the expiratory part of an FVC maneuver, for which it is important to know the BTPS correction factor. Among the 5 spirometers under investigation, only the PocketSpiro (MEC Group, Brussels, Belgium) allows switching between BTPS and ambient temperature and pressure condition.

We scrupulously followed the instructions of the manufacturers, in particular concerning the handling of the devices and the need for calibration checks. Each spirometer except the SpiroScout (Ganshorn, Niederlauer, Germany) requests a calibration before the measurements. Calibration involves either calibration checks (MicroLoop, Micro Medical, Rochester, United Kingdom) or recalibrations according to ATS standards. Both procedures consist of delivering different flows with a calibrated 3-L syringe stored at room temperature.

Table 1. Characteristics of the Equipment Under Investigation

Model	Flow Sensor	Software
PocketSpiro	Pneumotach VOM	PDI
MicroLoop	Turbine	Spida 5
SpiroScout	Ultrasonic	LF8
SpiroBank	Turbine	WinspiroPRO
SpiroStar	Screen pneumotach	Spiro 2000

VOM = variable orifice membrane

QUICK LOOK

Current knowledge

The performance of spirometers is evaluated using a mechanical simulator that reproduces the standard set by the American Thoracic Society curves generated by a computer. These evaluations are generally performed under ideal conditions and may not reflect the performance of spirometers in typical clinical use.

What this paper contributes to our knowledge

Independent testing of 5 office spirometers with a flow-volume simulator revealed important accuracy issues when using flow-volume loops obtained from actual patients. The data suggest that the standards be re-evaluated using more realistic flow-volume curves. The results cannot determine whether these changes will result in improvements that would alter diagnostic results.

To evaluate the different spirometers, we used the Hans Rudolph series 1120 flow-volume simulator. This device is able to generate flows of up to 16 L/s and volumes of up to 8.5 L. It can impose any kind of flow-volume curves previously sampled via the software Waveform Editor.

Protocol

After calibration or a calibration check according to the manufacturer's instructions, we first injected ambient air into the spirometers using a reference syringe (Hans Rudolph 3-L calibration syringe, volume of 2,993.32 mL [-0.22%] at 22.6°C). This test was performed 3 times and at different speeds (at a maximum flow of 1, 2, and 3 L/s). During this step, the volume measured by the spirometers may slightly exceed the volume of air originating from the calibration syringe. Indeed, any volume of air introduced in a spirometer will be artificially increased by the BTPS factor, which corrects for the difference in temperature and water vapor pressure between the airways and outside. This also applies for the 3 L of ambient air originating from the calibration syringe volume. We obtained the expiratory BTPS factors from the manufacturers, and we applied these factors to correct the volumes measured by the spirometers because we observed that, except for the PocketSpiro, the devices use a fixed expiratory BTPS factor. In other words, the volumes measured during expiration do not change according to the ambient variables. According to the ATS/ERS guidelines, the volumes measured with the 3-L syringe should meet the accuracy requirement of $\pm 3.5\%$.

Thereafter, the instrument resistance of each device was calculated by submitting each spirometer to a range of
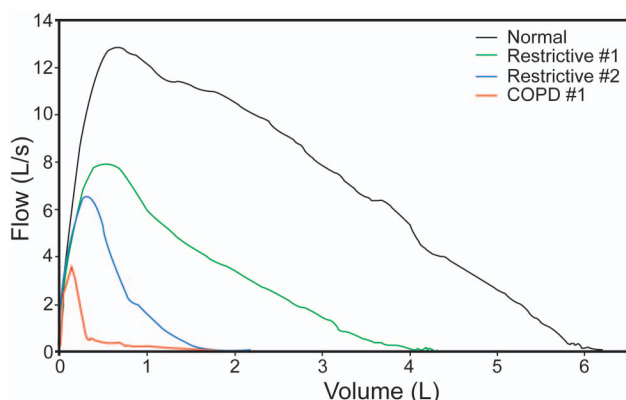
Fig. 1. Human flow-volume curves used with the flow-volume simulator. For the sake of clarity, only one COPD curve is represented.

constant flows (between 0.5 and 11 L/s) generated by the Hans Rudolph simulator. Pressure drop across the spirometer was measured using the pressure transducer of the simulator. The resistance of the tubing at a given flow was subtracted from the total resistance to obtain the instrument resistance. These measurements were done in triplicate.

Different waveforms from the computer-controlled simulator were introduced into the spirometers, each curve being imposed 3 times on each spirometer. The peak expiratory flow (PEF), FVC, and FEV_1 measured by the spirometer were compared with the actual values delivered by the simulator. For this part of the study, 9 curves from the standard test waveform set of the ATS were selected according to their complexity and to cover a broad range for all studied parameters. Five of these were volume-time curves (curves 1–4 and 12), and 4 of these were flow-time curves (curves 2, 8, 9, and 25). In addition to these standard curves, we also selected 6 flow-time curves obtained with subjects: one volunteer, 3 subjects with stage IV COPD, and 2 subjects with restrictive lung disease (Fig. 1 and Table 2). These spirometries, provided by the lung function laboratory of Ghent University Hospital (Ghent, Belgium), were performed on a Vmax Spectra (software version 12-7; SensorMedics, Yorba Linda, California). Flow, time, and integrated volume signals were recorded at a sampling rate of 500 Hz, saved in an Excel sheet, and subsequently downloaded on the Hans Rudolph simulator.

All the ATS and subject curves were tested under BTPS conditions. An external humidifier (HC150 with Ambient Tracking, Fisher & Paykel Healthcare, Auckland, New Zealand) and the heater of the lung simulator were used to obtain heated (37°C) and humidified (100% relative humidity) air. This setup allowed us to mimic human conditions without exceeding 2 min between each of the different maneuvers. Temperature and humidity were checked before each generation of flow (Dicon SM, Jumo, East Syracuse, New York).

Analysis

The readings provided by the different spirometers with the 3-L syringe at 3 different flows were individually compared with the expected values. Instrument resistances were expressed in cm H_2O/L/s and presented as means of 3 measurements for each different flow.

The repeatability and accuracy of the 3 measured variables (PEF, FEV_1, and FVC) were calculated for each of the 5 spirometers in accordance with the 2005 ATS/ERS guidelines only under BTPS conditions. To investigate repeatability, we computed the span between the maximum and minimum values for the 3 variables in each of the 15 curves. For a given variable, that span was defined as the difference between the maximum and minimum values of the successive maneuvers and expressed as: span = maximum − minimum. The percentage span was defined as: percentage span = 100 × (maximum − minimum)/average. The average is the mean of 3 successive maneuvers.

The accuracy of the spirometers was assessed in different ways. We first calculated the absolute and percentage deviation of the 15 curves for the 3 variables, and for each of the spirometers, we computed the number of accuracy errors as defined by the ATS/ERS guidelines: deviation = average − standard; percentage deviation = 100 × (average − standard)/standard. The average is the mean of 3 successive maneuvers, and the standard is the value of the standard ATS curve.

We took the average of 3 measurements for each parameter for the Bland-Altman analysis. We calculated the bias, which was defined as the mean difference between the volume (or flow) measured by a spirometer and the volume (or flow) generated by the simulator (virtually the target value of the ATS and subject curves) for each curve and each spirometer. The bias or mean difference between the spirometers and the Hans Rudolph simulator was calculated separately for the ATS and subject curves. $P < .05$ was considered significant, meaning that a bias was present or significantly different from zero. A positive bias means that, on average, the tested spirometer overestimates the measurements with respect to the simulator and vice versa.

This study was approved by the ethics committee (B403201318648) of the Cliniques Universitaires Saint Luc (Brussels, Belgium).

Results

Volume Measurements With Calibration Syringe at Different Flows

Table 3 presents the volumes measured by the different spirometers after injection of 3 L of air with the calibration

Table 2.   Characteristics of Subject Curves

Subject	FVC (L)	FEV$_1$ (L)	PEF (L/s)	RT (ms)	EV (% FVC)	Forced T$_E$ (s)
Restrictive 1	1.78	1.47	6.30	63	2.1	6.5
Restrictive 2	4.33	3.21	7.87	63	1.7	10.8
COPD 1	2.18	0.61	3.41	47	0.4	11.2
COPD 2	2.81	0.64	2.24	125	0.8	10.8
COPD 3	2.90	0.71	1.93	47	0.3	14.9
Healthy	6.30	5.52	12.58	55	1.8	6.3

PEF = peak expiratory flow
RT = rise time from 10 to 90% of PEF
EV = extrapolated volume
T$_E$ = expiratory time

Table 3.   Volume Measurements After Injection of 3 L of Air at Different Flows at 20°C Corrected to Ambient Conditions

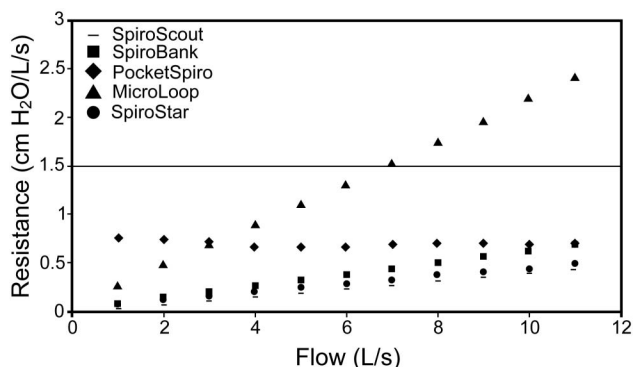|	Model				
	PocketSpiro (L)	SpiroScout (L)	SpiroBank (L)	MicroLoop (L)	SpiroStar (L)
Flow					
1 L/s	2.98	3.01	2.99	3.03	2.96
2 L/s	2.99	3.02	2.99	3.00	2.93
3 L/s	2.98	2.99	3.01	2.99	2.94
Maximum − minimum FVC	0.01	0.03	0.02	0.04	0.03

Fig. 2. Resistance of the spirometers in relation to flow. The horizontal line shows the upper limit of resistance according to the American Thoracic Society/European Respiratory Society (ATS/ERS) guidelines.

syringe at different flows under ambient conditions. Volume measurements were slightly affected by air flow but remained within the calibration limits of 3.5%, proposed by the ATS/ERS guidelines.

Instrument Resistance

Figure 2 shows the resistance values of each spirometer for different flows. Resistance increased with flow except for the PocketSpiro, which uses a variable orifice membrane sensor. Four of the 5 spirometers exhibited a resis-

tance of < 1.5 cm H$_2$O/L/s at a flow of 14 L/s. The MicroLoop presented a resistance exceeding 1.5 cm H$_2$O/L/s for flows above 7 L/s and did not comply with the recommendations of the ATS/ERS guidelines.

Repeatability

The repeatability of all spirometers tested under BTPS conditions is shown in Table 4. None of measured values fell outside the limits defined by the ATS/ERS guidelines.

Accuracy

The results of the 9 ATS and 6 subject curves were analyzed separately. Table 5 presents the FVC data, and Table 6 shows the mean ± SD for FEV$_1$, FVC, and PEF generated by the simulator for each spirometer and the mean ± SD for the biases in absolute values and as a percentage. The number of tests that were outside the ATS/ERS standards is summarized (errors).

Of the 5 spirometers, only the SpiroStar (Medriko, Kuopio, Finland) was not in accord with the ATS/ERS requirements for accuracy of PEF measurement (see Table 6). In contrast, the PocketSpiro and SpiroScout failed to measure accurately the FVC of all the ATS curves. Figure 3 shows that the SpiroStar underestimated all the FEV$_1$ values while meeting all the ATS/ERS criteria when the curves were measured under BTPS conditions. In this case, the bias

Table 4.   Repeatability of the Spirometer Results

Parameter	MicroLoop		SpiroBank		SpiroStar		PocketSpiro		SpiroScout	
	Span	%	Span	%	Span	%	Span	%	Span	%
ATS 1/24										
PEF	0.06	0.90	0.01	0.16	0.03	0.48	0.18	2.78	0.11	1.68
FVC	0.02	0.34	0.03	0.49	0.03	0.51	0.09	1.54	0.03	0.50
FEV_1	0.01	0.23	0.05	1.15	0.01	0.24	0.04	0.96	0.02	0.47
ATS 2/24										
PEF	0.11	1.07	0.06	0.61	0.06	0.62	0.07	0.69	0.32	3.12
FVC	0.02	0.39	0.06	1.18	0.03	0.61	0.01	0.20	0.02	0.41
FEV_1	0.01	0.22	0.02	0.43	0.02	0.45	0.01	0.22	0.02	0.44
ATS 3/24										
PEF	0.02	1.46	0.00	0.00	0.00	0.00	0.05	3.69	0.02	1.39
FVC	0.01	0.29	0.09	2.65	0.08	2.32	0.07	2.09	0.06	1.83
FEV_1	0.02	1.67	0.01	0.85	0.01	0.88	0.05	4.27	0.03	2.56
ATS 4/24										
PEF	0.05	1.64	0.03	0.99	0.00	0.00	0.02	0.69	0.11	3.58
FVC	0.01	0.66	0.02	1.31	0.01	0.69	0.04	2.79	0.01	0.68
FEV_1	0.02	1.44	0.03	2.16	0.01	0.76	0.03	2.26	0.02	1.47
ATS12/24										
PEF	0.01	0.25	0.02	0.52	0.02	0.54	0.12	3.14	0.08	1.97
FVC	0.03	1.52	0.06	3.08	0.10	5.03	0.03	1.55	0.01	0.52
FEV_1	0.00	0.00	0.03	1.86	0.01	0.64	0.03	1.90	0.02	1.23
ATS 2/26										
PEF	0.05	0.45	0.03	0.28	0.03	0.30	0.03	0.28	0.02	0.18
FVC	0.02	0.46	0.03	0.69	0.01	0.24	0.02	0.48	0.03	0.72
FEV_1	0.03	0.77	0.03	0.77	0.01	0.27	0.02	0.52	0.03	0.79
ATS 8/26										
PEF	0.06	2.71	0.01	0.46	0.01	0.48	0.02	0.88	0.15	6.14
FVC	0.01	0.64	0.03	2.06	0.01	0.70	0.01	0.70	0.01	0.72
FEV_1	0.00	0.00	0.02	2.08	0.01	1.08	0.01	1.06	0.01	1.06
ATS 9/26										
PEF	0.04	0.72	0.01	0.19	0.04	0.81	0.10	1.86	0.11	2.02
FVC	0.02	0.74	0.02	0.75	0.03	1.17	0.02	0.78	0.03	1.16
FEV_1	0.01	0.45	0.02	0.90	0.01	0.47	0.00	0.00	0.03	1.37
ATS 25/26										
PEF	0.27	1.84	0.00	0.00	0.03	0.22	0.04	0.29	0.15	1.06
FVC	0.03	0.45	0.04	0.60	0.01	0.16	0.07	1.10	0.05	0.79
FEV_1	0.02	0.33	0.02	0.33	0.00	0.00	0.05	0.86	0.06	1.03
Restrictive 1										
PEF	0.13	1.98	0.02	0.31	0.07	1.22	0.08	1.27	0.11	1.69
FVC	0.02	1.09	0.05	2.91	0.02	1.16	0.02	1.16	0.01	0.60
FEV_1	0.01	0.67	0.02	1.36	0.01	0.70	0.01	0.69	0.01	0.70
Restrictive 2										
PEF	0.04	0.49	0.06	0.76	0.02	0.27	0.08	0.99	0.18	2.15
FVC	0.02	0.46	0.09	2.12	0.01	0.24	0.03	0.72	0.03	0.74
FEV_1	0.01	0.31	0.01	0.31	0.00	0.00	0.02	0.64	0.02	0.61
COPD 1										
PEF	0.03	0.81	0.02	0.52	0.04	1.35	0.13	3.76	0.20	5.32
FVC	0.02	0.87	0.03	1.41	0.07	3.13	0.01	0.47	0.02	1.01
FEV_1	0.00	0.00	0.03	4.29	0.00	0.00	0.00	0.00	0.00	0.00
COPD 2										
PEF	0.00	0.00	0.01	0.47	0.01	0.50	0.02	0.91	0.02	0.88
FVC	0.00	0.00	0.01	0.35	0.07	2.37	0.01	0.36	0.02	0.75
FEV_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.62
COPD 3										
PEF	0.03	1.57	0.00	0.00	0.01	0.58	0.03	1.51	0.18	8.61
FVC	0.01	0.35	0.09	3.30	0.10	3.33	0.01	0.35	0.02	0.95
FEV_1	0.00	0.00	0.03	4.21	0.01	1.44	0.00	0.00	0.01	1.42
Healthy										
PEF	0.17	1.37	0.05	0.41	0.04	0.34	0.04	0.33	0.04	0.31
FVC	0.02	0.31	0.03	0.48	0.02	0.33	0.02	0.33	0.03	0.49
FEV_1	0.03	0.54	0.02	0.36	0.01	0.19	0.02	0.37	0.02	0.36

The absolute span and percentage span are indicated (see text for details). Span is expressed in L/s for peak expiratory flow (PEF) and in L for FVC and FEV_1.
ATS = American Thoracic Society

Table 5.   Accuracy of the Spirometers Regarding FVC Measured During BTPS Conditions

	Target (L)	MicroLoop		SpiroBank		SpiroStar		PocketSpiro		SpiroScout	
		L	%	L	%	L	%	L	%	L	%
ATS 1/24	6.000	−0.039	−0.7	0.101	1.7	−0.166	−2.8	−0.136	−2.5	−0.047	−0.8
ATS 2/24	4.999	0.070	1.4	0.079	1.6	−0.110	−2.2	−0.064	−1.5	−0.068	−1.4
ATS 3/24	3.498	0.005	0.2	−0.101	−2.9	−0.048	−1.4	−0.201*	−5.5*	−0.220*	−6.3*
ATS 4/24	1.498	0.014	1.0	0.028	1.8	−0.056	−3.7	−0.053	−4.4	−0.036	−2.4
ATS 12/24	2.002	−0.029	−1.4	−0.053	−2.6	−0.012	−0.6	−0.060	−3.6	−0.080	−4.0
ATS 2/26	4.247	0.086	2.0	0.096	2.3	−0.104	−2.4	−0.044	−1.0	−0.077	−1.8
ATS 8/26	1.453	0.099	6.8	−0.001	0.0	−0.019	−1.3	−0.031	−2.1	−0.057	−3.9
ATS 9/26	2.617	0.099	3.8	0.049	1.9	−0.051	−2.0	−0.048	−1.8	−0.038	−1.5
ATS 25/26	6.502	0.124	1.9	0.110	1.7	−0.200	−3.1	−0.166	−2.6	−0.143	−2.2
Restrictive 1	1.784	0.052	2.9	−0.055	−3.6	−0.061	−3.4	−0.064	−3.6	−0.122	−6.8
Restrictive 2	4.333	0.015	0.3	−0.078	−1.9	−0.125	−2.9	−0.149	−3.4	−0.279*	−6.4*
COPD 1	2.181	0.106	4.8	−0.048	−2.2	0.056	2.6	−0.038	−1.7	−0.201*	−9.2*
COPD 2	2.814	0.066	2.4	0.053	1.5	0.136	4.8	−0.008	−0.3	−0.145	−5.2
COPD 3	2.901	−0.047	−1.6	−0.202*	−7.0*	0.101	3.5	−0.059	−2.0	−0.792*	−27.3*
Healthy	6.298	0.121	1.9	−0.025	−0.4	−0.148	−2.4	−0.188	−3.0	−0.145	−2.3

Target is the FVC value programmed by the software of the simulator. The deviations (L and %) are the average value of 3 measurements (see text for details).
* Deviation beyond the accuracy limits
BTPS = body temperature and pressure saturated
ATS = American Thoracic Society

was proportional to the measured FEV_1 ($r^2 = 0.92$), indicating a proportional error. In fact, almost all spirometers had a significant bias, so they have a tendency to overestimate or to underestimate the values. Table 6 shows that significant biases were also present for the FVC and PEF measurements.

A large difference in bias appeared to exist between the ATS curves and the subject curves, with the biases of the ATS curves remaining more within acceptable limits. Particularly the FVC of the SpiroScout (Fig. 4) and the PEF of the SpiroStar were definitely beyond the ATS/ERS limits. Moreover, the biases of the curves originating from subjects with COPD exceeded those originating from healthy subjects and subjects with restrictive disorders, as illustrated with the Bland-Altman plot (see Fig. 3 and Table 6).

Discussion

In this study, an independent test of 5 office spirometers using a validated flow-volume simulator revealed significant accuracy issues with some devices under BTPS conditions. Moreover, we highlighted some important disagreements in volumes and flows between the simulator and the office spirometers, which became apparent with curves obtained from subjects with severe COPD but remained unnoticed with the currently recommended ATS curves.

This study has some limitations that have to be addressed before any further discussion and contextualiza-

tion of the data. First, we tested only one device for each model of spirometer, except for the MicroLoop, for which we requested another flow sensor to double-check the unexpected high resistance. The number of each spirometer to be tested is not debated in the ATS/ERS guidelines, where it is stated that a production spirometer should be submitted to the test with the pump system. Future studies could answer this question by testing several spirometers instead of only one. We followed scrupulously the recommendations of the manufacturers in performing a true calibration of the devices that require it and only a calibration check if this was recommended or feasible, mimicking real-life conditions as much as possible.

Another limitation is that we did not update the software of the devices during the course of the study; the new devices that reached the market after the start of the study were not included. From a theoretical point, we might have missed some of the defaults spirometers may generate because we tested only a subset of the 50 curves currently recommended by the ATS/ERS guidelines. The 9 selected curves were among the most complex and extreme curves found in the ATS set. If our data are not ideally suited to make true comparisons between office spirometers, they are still exploitable from a methodological point of view because the spirometers were tested with a sample of curves representing patients currently seen in daily practice, in addition to the recommended ATS curves. Admittedly, it might have been interesting to extend the present observations to a wider range of subject curves. Indeed, the characteristics of the curves obtained from our

Table 6.   PEF, FVC, and FEV$_1$ Measured With 5 Different Spirometers for 9 ATS Curves and 6 Subject Curves

	Simulator (Mean ± SD)	Bias (Mean ± SD)	Bias (Mean ± SD, %)	Errors
PEF (L/s)				
PocketSpiro				
ATS	6.511 ± 4.420	−0.176 ± 0.134*	−3.61 ± 3.07	0
Subject	5.833 ± 4.117	−0.153 ± 0.269	−2.35 ± 2.33	0
SpiroScout				
ATS	6.502 ± 4.423	0.004 ± 0.080	0.32 ± 1.64	0
Subject	5.852 ± 4.101	0.118 ± 0.163	1.93 ± 3.00	0
SpiroBank				
ATS	6.491 ± 4.417	−0.189 ± 0.125*	−3.75 ± 3.22	0
Subject	5.821 ± 4.111	−0.072 ± 0.247	−1.44 ± 6.64	1
MicroLoop				
ATS	6.446 ± 4.356	0.073 ± 0.209	−0.62 ± 4.27	0
Subject	5.743 ± 3.924	0.074 ± 0.172	−0.40 ± 4.55	0
SpiroStar				
ATS	6.503 ± 4.436	−0.462 ± 0.321*	−7.81 ± 3.64	2
Subject	5.821 ± 4.134	−0.548 ± 0.273*	−11.04 ± 3.44	2
FVC (L)				
PocketSpiro				
ATS	3.647 ± 1.907	−0.089 ± 0.052*	−2.66 ± 1.21	1
Subject	3.384 ± 1.670	−0.084 ± 0.069*	−2.34 ± 1.26	0
SpiroScout				
ATS	3.646 ± 1.906	−0.085 ± 0.060*	−2.69 ± 1.74	1
Subject	3.386 ± 1.670	−0.281 ± 0.257*	−9.53 ± 8.99	3
SpiroBank				
ATS	3.647 ± 1.907	0.034 ± 0.074	0.59 ± 2.01	0
Subject	3.384 ± 1.670	−0.059 ± 0.072	−2.11 ± 2.61	1
MicroLoop				
ATS	3.647 ± 1.907	0.048 ± 0.060*	1.66 ± 2.48	0
Subject	3.383 ± 1.669	0.052 ± 0.062	1.79 ± 2.21	0
SpiroStar				
ATS	3.647 ± 1.907	−0.085 ± 0.065*	−2.16 ± 0.97	0
Subject	3.385 ± 1.671	−0.007 ± 0.121	0.37 ± 3.66	0
FEV$_1$ (L)				
PocketSpiro				
ATS	2.878 ± 1.803	−0.045 ± 0.041*	−1.64 ± 0.99	0
Subject	2.037 ± 1.969	−0.039 ± 0.050	−1.43 ± 0.81	0
SpiroScout				
ATS	2.877 ± 1.802	−0.027 ± 0.045	−0.71 ± 0.84	0
Subject	2.038 ± 1.969	−0.013 ± 0.031	−1.55 ± 1.87	0
SpiroBank				
ATS	2.878 ± 1.802	0.033 ± 0.036*	1.01 ± 0.90	0
Subject	2.037 ± 1.969	0.026 ± 0.027	1.60 ± 1.55	0
MicroLoop				
ATS	2.880 ± 1.805	0.040 ± 0.022*	1.55 ± 0.55	0
Subject	2.037 ± 1.972	0.033 ± 0.016*	2.66 ± 1.67	0
SpiroStar				
ATS	2.878 ± 1.802	−0.082 ± 0.049*	−3.00 ± 0.52	0
Subject	2.038 ± 1.970	−0.055 ± 0.069	−1.86 ± 1.65	0

* Bias significantly different from zero (Bland-Altman analysis)
PEF = peak expiratory flow
ATS = American Thoracic Society

subjects with COPD were similar to each other in terms of FEV$_1$, precluding any generalization of our conclusions. Nevertheless, with only 3 COPD subjects randomly selected from a database of an out-patient clinic, we identified problems in 3 spirometers, which were partly or completely overlooked with the conventional ATS curves.

We observed that both the simulator and the spirometers provided highly reproducible results and that the calibra-
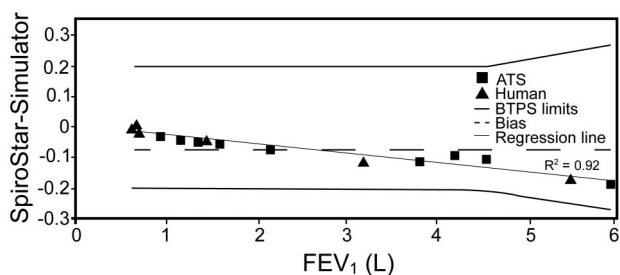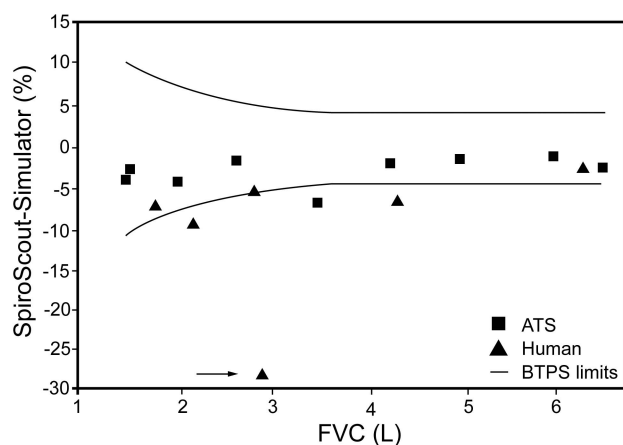
Fig. 3. Bland-Altman plot showing the relation between the absolute values of FEV_1 and the mean bias measured by the SpiroStar using American Thoracic Society (ATS) curves under body temperature and pressure saturated (BTPS) conditions. The subject curves (in BTPS) are also represented. The lines show the upper and lower bounds of the ATS/European Respiratory Society specifications for BTPS testing. The dashed line shows the bias (or mean difference) calculated for both ATS and subject curves.

Fig. 4. Bland-Altman plot showing the relation between the absolute values of FVC and the bias in percent measured by the SpiroScout using American Thoracic Society (ATS) curves under body temperature and pressure saturated (BTPS) conditions. The subject curves (in BTPS) are also represented. The lines show the upper and lower bounds of the ATS/European Respiratory Society specifications for BTPS testing. The arrow indicates a FVC value largely out of range.

tion check with the 3-L syringe remained within the expected limits. Conversely, none of the devices tested was perfect in terms of accuracy for volumes and flows during a forced expiratory maneuver.

Indeed, the ATS/ERS guidelines define acceptable spirometer performance as < 3 accuracy errors for either FVC or FEV_1 across the 4 BTPS waveforms. Although this was always the case for the FEV_1, 3 spirometers produced individual deviations between the target FVC value and those measured by the spirometer. Moreover, the Bland-Altman analysis revealed a tendency for 3 spirometers to overestimate the real value. In other words, a spirometer can present a significant bias even if it perfectly meets the ATS/ERS requirements, as seen with the SpiroStar. The

existence of a systematic bias and/or proportional errors may limit the interchangeability of the measurements between different spirometers and questions the reliability of data generated with portable spirometers in epidemiologic studies. The clinical importance of the errors depends on the variable tested. The largest errors were found in the FVC measurements (see Table 5). FVC is one major outcome parameter in several clinical studies on the treatment of idiopathic pulmonary fibrosis.[7] This parameter is a robust predictor of the survival of patients with interstitial lung diseases. If we assume that the long-term repeatability of all the devices we tested is good, the presence of biases may preclude their interchangeability during the follow-up of patients.

As an example, using COPD curve 3 with a target FVC of 2.901 L, we obtained 2.109 L with the SpiroScout and 3.002 L with the SpiroStar. When we look at the average bias (see Table 6), we observe a significant negative bias with the subject set of curves with the SpiroScout (−0.281 L), but no significant bias with the SpiroStar (−0.002 L). During a longitudinal study or follow-up of patients, the biases we observed may become a clinically important issue if we have to replace the spirometer. Therefore, when a device must be replaced, if the same model is not available, we suggest choosing a model with the same characteristics, that is, the same bias.

A new and surprising finding of this study was that the SpiroScout, SpiroBank (MIR, Rome, Italy), and SpiroStar failed to accurately measure some of the FVC or PEF values using curves obtained from subjects. This questions whether the currently recommended ATS curves should be supplemented by a series of curves characterized by a prolonged and very low expiratory flow, as seen in patients with severe COPD.

Manufacturers claim that all spirometers presented in this study comply with the requirements of the ATS/ERS. This label supposes that their accuracy and repeatability were checked with the standard curves using a computer-driven piston pump. The ATS/ERS guidelines state that, under ambient conditions, acceptable spirometer performance is defined as < 3 accuracy errors for either FVC or FEV_1 across the 24 waveforms (< 5% error rate) and that, under BTPS conditions, all 4 curves should be measured within the limits (± 4.5% or ± 200 mL, whichever is greater). As these errors often occur in the most complex curves, problems occurring in very demanding curves with a steep rise time or low expiratory flows, as seen in patients with severe COPD, may be easily missed, and an ATS/ERS compliance label may be provided to devices that do not deserve it.

Careful analysis of our data indicated that some office spirometers experienced difficulties in measuring flows and volumes correctly if the curves presented the following features: low-rise time, high-frequency oscillations,
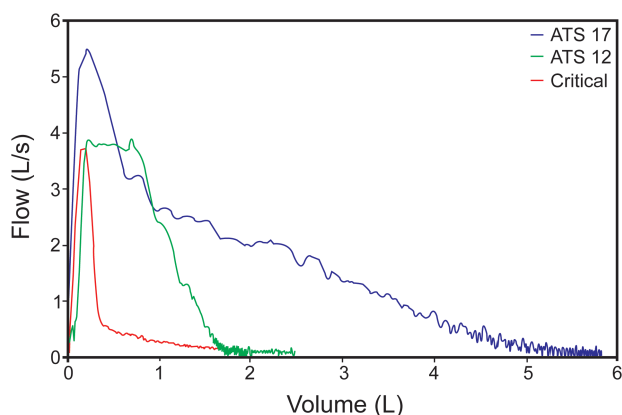
Fig. 5. The critical flow-volume curve compared with American Thoracic Society (ATS) curves 12 and 17.

Table 7. Characteristics of the Critical Curve

Parameter	Critical Curve
PEF, L/s	3.608
FVC, L	2.448
FEV_1, L	0.701
RT, ms	31
V_E, %	1.50
Forced T_E, s	14.32
f_{max}, Hz	12

PEF = peak expiratory flow
RT = rise time
V_E = extrapolated volume
T_E = expiratory time
f_{max} = maximal frequency of the oscillations

and low flows over a long expiratory time. Other variables such as the extrapolated volume and a rapidly decreasing PEF might also influence the results. These characteristics are more frequently found in patients with severe COPD (Global Initiative for Chronic Obstructive Lung Disease Stage IV). As it is difficult to find one real-life COPD patient presenting all these features in one curve, such a curve was recently created by combining the critical issues found in our COPD patients, calling it the critical curve. Its flow-volume representation and its characteristics are given in Figure 5 and Table 7. It is characterized by a low-rise time, a rapid decrease in flow after the PEF, and very low flows over a few seconds with high-frequency oscillations. In a preliminary study, we tested all the spirometers with this new curve under BTPS conditions, and we very easily identified the 2 devices that also presented problems in the present study. Whether this critical curve might be useful in the initial evaluation of newly marketed office spirometers remains to be investigated in a prospective study.

In light of these exciting findings, we could imagine a new procedure to certify spirometers. If our concept of a critical curve could be implemented, a test procedure could start with 3 curves, each with one particular feature such as low flow, low-rise time, or high-frequency oscillations, and a fourth curve would combine all these critical parameters. In a second step, some of the current ATS curves (maybe 8–10 curves) would be kept to test the spirometers throughout the range of spirometric values measured in a sample of patients and healthy volunteers. Even if it is more time-consuming, we favor testing under BTPS conditions, which are closer to real-life conditions. We also propose to include a Bland-Altman analysis for expression of the results, not mentioned in the ATS/ERS guidelines, because this method helps to identify problems due to proportional errors and bias. Finally, the spirometers should be tested in real-life conditions in a comparative assessment with standard validated spirometers in a significant sample of patients and healthy volunteers. There is no consensus on the number of healthy volunteers and patients required to test spirometers, but published studies were performed using from 15[8] to 75[9] subjects. Such a study should not be performed using spirometers connected in series because we have demonstrated that this method is not valid.[10]

Conclusions

Independent testing of 5 office spirometers with a flow-volume simulator revealed important accuracy issues that became particularly apparent using flow-volume loops obtained in patients. These findings invite the scientific community to reconsider the currently accepted procedures to evaluate spirometers. We propose to improve the testing of spirometers by implementing new flow-volume curves that are more realistic than the original ones. By refining the analysis of the results, we may identify some biases not shown by the ATS/ERS procedure. However, whether such improvements may eventually affect the quality of the diagnostic process in patients with respiratory symptoms remains to be demonstrated.

REFERENCES

1. Ferguson GT, Enright PL, Buist AS, Higgins MW. Office spirometry for lung health assessment in adults: a consensus statement from the National Lung Health Education Program. Chest 2000;117(4):1146-1161.
2. Buffels J, Degryse J, Heyrman J, Decramer M. Office spirometry significantly improves early detection of COPD in general practice: the DIDASCO Study. Chest 2004;125(4):1394-1399.

3. Standardization of Spirometry, 1994 Update. American Thoracic Society. Am J Respir Crit Care Med 1995;152(3):1107-1136.

4. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. Eur Respir J 2005;26(2):319-338.

5. Liistro G, Vanwelde C, Vincken W, Vandevoorde J, Verleden G, Buffels J. Technical and functional assessment of 10 office spirometers: a multicenter comparative study. Chest 2006;130(3):657-665.

6. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307-310.

7. Collard HR, King TE Jr, Bartelson BB, Vourlekis JS, Schwarz MI, Brown KK. Changes in clinical and physiologic variables predict survival in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2003;168(5):538-542.

8. Abboud S, Bruderman I. Assessment of a new transtelephonic portable spirometer. Thorax 1996;51(4):407-410.

9. Rebuck DA, Hanania NA, D'Urzo AD, Chapman KR. The accuracy of a handheld portable spirometer. Chest 1996;109(1):152-157.

10. Lefebvre Q, Vandergoten T, Derom E, Marchandise E, Liistro G. Technical assessment of spirometers connected in series. Respir Care 2012;57(8):1273-1277.