

# The accuracy of a handheld “disposable pneumotachograph device” in the spirometric diagnosis of airway obstruction in a Chinese population

Guojun Chen<sup>1,\*</sup>  
 Longyuan Jiang<sup>1,\*</sup>  
 Liwen Wang<sup>1</sup>  
 Wei Zhang<sup>2</sup>  
 Carlos Castillo<sup>3,4</sup>  
 Xiangshao Fang<sup>1</sup>

<sup>1</sup>Department of Emergency Medicine, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China;

<sup>2</sup>Department of Respiratory Medicine, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China;

<sup>3</sup>Zhang Neuroscience Research Laboratories, School of Medicine, Loma Linda University, Loma Linda, CA, USA;

<sup>4</sup>Bioengineering Department, Gordon and Jill Bourns College of Engineering, California Baptist University, Riverside, CA, USA

\*These authors contributed equally to this work

**Background and aim:** It is desirable to facilitate the use of an affordable, reliable, and portable spirometer, for earlier diagnosis of COPD in China, particularly in rural areas. The aim of this study was to assess the agreement of a handheld “disposable pneumotachograph” (D-PNEU) spirometer with the gold standard spirometer and to evaluate its diagnostic accuracy of spirometric classification of airflow obstruction.

**Subjects and methods:** A total of 241 adult Chinese subjects ranging from healthy to those with mixed levels of pulmonary disease performed spirometry in a conventional body plethysmograph, and using a D-PNEU device in randomized order. The three best spirometric tests were recorded for comparative analysis. A Bland–Altman graph was created to assess the agreement between devices. Using  $FEV_1/FVC < 70\%$  as the “gold standard” for obstruction, the accuracy of classifying the severity of airway obstruction for all subjects was assessed. For the specific individuals ( $n=159$ ) able to exhale for at least 6 seconds, the accuracy of classifying airway obstruction was further assessed. For this purpose, a receiver operating characteristic curve was used to determine an optimal cutoff point of  $FEV_1/FEV_6$  ratio obtained by the D-PNEU device, which matched the global definition of  $FEV_1/FVC < 70\%$  by the traditional spirometer.

**Results:** The Bland–Altman analysis showed that the between-device agreement for key airflow metrics was within clinically acceptable limits. The D-PNEU device had 87.1% accuracy in the classification of severity of obstruction in all 241 subjects, when using  $FEV_1/FVC < 70\%$  as the “gold standard” for both devices. The D-PNEU device had 93.7% accuracy in the 159 individuals able to exhale for at least 6 seconds, when a cutoff point of  $FEV_1/FEV_6$  was 74%.

**Conclusion:** A disposable handheld spirometry device is capable of accurately identifying and quantifying airway obstruction in patients deemed to be at risk, however, caution should be exercised and all available brands should be tested.

**Keywords:** chronic obstructive pulmonary disease, handheld spirometry, Bland–Altman method, agreement analysis

## Introduction

In China, COPD ranks as the fourth leading cause of death in urban areas and third leading in rural areas in this middle-income country.<sup>1,2</sup> A recent spirometry-based survey revealed that the overall prevalence of COPD in China is 8.2% in individuals 40 years of age or older, which is a large proportion of the population.<sup>3</sup> Laboratory-based spirometry is essential as a basis for the proper staging of pathology and for

Correspondence: Xiangshao Fang  
 Department of Emergency Medicine,  
 Sun Yat-sen Memorial Hospital, Sun Yat-sen University, 107 Yan Jiang Xi Road,  
 Guangzhou 510120, China  
 Tel +86 20 8133 2084  
 Email fangxsh@syzu.edu.cn

follow-up of disease.<sup>4</sup> Yet, it has been reported that among Chinese patients who have COPD, only 6.5% have been tested with spirometry.<sup>3</sup>

Testing options for spirometry range from full body plethysmography to fully portable units that are wirelessly connected to mobile phones. One of the more popular methods for evaluating patients with COPD among providers is the use of clinical-grade, in-office, handheld spirometry solutions. Ensuring their accuracy is essential. Several large, previous studies have indicated that the quality and user-friendliness of several in-office spirometers make them acceptable for detection of COPD, while a few increase the risk of misclassification.<sup>5,6</sup> In China, there is insufficient literature to support the agreement and validity of handheld spirometry as compared to full body plethysmography for lung function performance. Such information is critical to advance the proper diagnosis and management of COPD, especially for those in rural China where COPD rates are higher. A clinical-grade handheld “disposable pneumotachograph” (D-PNEU) spirometer, represents an attractive option for outpatient diagnosis of COPD. Being relatively low-cost and easily deployed using a laptop, the device may be a viable alternative to the more expensive and much less portable full body plethysmography method.

An important parameter in determining the quality of a medical instrument is agreement with a gold standard. Systematic reviews have concluded that the Bland–Altman method is the preferred method to assess agreement between medical instruments measuring continuous variables.<sup>7–9</sup> Thus, the primary objective of this study was to assess the agreement between devices for key airflow metrics using the Bland–Altman method.

Our secondary aim, apart from the between-device agreement for key airflow metrics, was to measure the clinical accuracy of the D-PNEU spirometer as compared to a laboratory-based, traditional full body spirometer in a Chinese population. First, using  $FEV_1/FVC < 70\%$  as the “gold standard” for obstruction, the accuracy of classifying the severity of airway obstruction between devices for all enrolled subjects was assessed, regardless of expiratory times. The effort to empty the lungs fully, in order to reach FVC, can be particularly difficult for some patients, especially for older patients and those with severe respiratory diseases. Indeed, accumulated evidence has suggested that the ratio of  $FEV_1/FEV_6$  in 6 seconds ( $FEV_1/FEV_6$ ) can be used as a valid alternative for  $FEV_1/FVC$ .<sup>10–12</sup> Thus, the accuracy of classifying airway obstruction between devices for these specific individuals able to exhale for at least

6 seconds was further assessed. For this purpose, the fixed cutoff point of  $FEV_1/FEV_6$  ratio obtained by the D-PNEU spirometer was explored, which matched the global definition of  $FEV_1/FVC < 70\%$  by the traditional spirometer.

## Subjects and methods

### Subjects

The study was approved by the ethics committee at Sun Yat-sen Memorial Hospital of Sun Yat-sen University (no 2013–66) and performed in accordance with the Declaration of Helsinki. All subjects provided written informed consent to participate in the study. The study was registered with the Chinese Clinical Trial Registry, with the number ChiCTR-DDD-17013664. The study was carried out from October 2014 to February 2015 in Guangzhou, the capital city of China’s Guangdong province. A total of 241 patients referred to the pulmonary function test (PFT) laboratory, participated in the PFT by using two spirometers. The subjects were considered eligible for participation in the study if they were medically stable, capable of performing repeated spirometric measurements, and were without obvious medical contraindications prior to participation.

### Spirometers

The gold standard spirometer used within our center is the body plethysmograph (Elite DX Model NO-830001-005, MedGraphics Corp, St Paul, MN, USA).

The D-PNEU spirometer (Model IQspiro, Midmark Corp, Dayton, OH, USA) used in our study, is a portable unit, weighing only nine ounces and is comprised of a single-use disposable mouthpiece and a handle (digital transducer) (Figure 1). The unit connects to a computer via a USB cable, which also supplies the power for the unit. Among the measured parameters, the IQspiro includes FVC,  $FEV_3$ ,  $FEV_1$ ,  $FEV_3$ ,  $FEV_6$ ,  $FEV_1/FVC$ ,  $FEV_3/FVC$ ,  $FEV_1/FEV_6$ , peak flow and mid flow,  $FEV_{25\%}$ ,  $FEV_{50\%}$ ,  $FEV_{75\%}$ , and  $FEV_{25\%–75\%}$ . The measuring range of flow is  $\pm 14$  L/s and volume is  $\pm 8$  L. The features of the tested handheld spirometer indicated that it might be compatible with the plethysmograph. The management software provides the graphical interface for the operator to conduct an effective test. Once the test is completed, the software provides immediate visibility of the key airflow parameters and a flow-volume curve, along with an automatic and customizable interpretation of the pulmonary function results. Its precision and accuracy is reported by the manufacturer to meet or exceed the European Respiratory Society (ERS) and American Thoracic Society (ATS) standards issued in 2005.<sup>13</sup>



**Figure 1** The handheld “disposable pneumotachograph” spirometer.

Quality control of spirometry was based on the ERS/ATS recommendations.<sup>13</sup> A single discharge of 3 L calibrated syringe was used daily to check the volume accuracy of each spirometer. The percentage difference of volume needed to be within the range of  $\pm 3.5\%$ .<sup>13</sup> For IQspiro, an accurate calibration was automatically achieved prior to patient testing.

## Protocol

The subjects were asked to perform two sets of tests, one on the established gold standard spirometer and the other set on the D-PNEU spirometer device. The order of testing was randomized and conducted in single-blind fashion, meaning the subjects were unaware as to which machine was under study. The predicted equations for the population in South China (Guangdong province) selected in this study were from the nationwide normal lung function study, which was organized and sponsored by the Ministry of Health, China.<sup>14</sup>

The following indices were measured or derived using the two devices: FVC,  $FEV_1$ ,  $FEV_1/FVC$ , mean FEF calculated between 25% and 75% of FVC ( $FEF_{25\%-75\%}$ ), and PEF. In addition,  $FEV_1/FEV_6$  ratio was obtained by the disposable spirometer for individuals able to exhale for 6 seconds. Then, a fixed cutoff point of  $FEV_1/FEV_6$  ratio derived from the D-PNEU spirometer for the detection of obstruction was explored, which corresponded best to the global definition of  $FEV_1/FVC < 70\%$  by the traditional spirometer. The two

cutoffs were considered synonymous indices between the two units.

At least three but no more than eight maneuvers per device were performed on each subject. The three best spirometric maneuvers were further analyzed. As a quality control measure, the difference between the two largest FVC and the two largest  $FEV_1$  needed to be within 0.15 L or 5% in order for the overall test to be considered accurate and reproducible. Tests that achieved the largest sum of  $FEV_1$  plus FVC were chosen as the best studies for analysis. The FVC and  $FEV_1$  were recorded from the highest value among all accepted curves. The  $FEV_1/FVC$  ratio was calculated from the best  $FEV_1$  and the best FVC, whereas the  $FEF_{25\%-75\%}$  was recorded from the best curve with the largest sum of  $FEV_1$  and FVC.  $FEV_1/FEV_6$  derived from the D-PNEU spirometer was calculated from the best  $FEV_1$  and the best  $FEV_6$ .

## Spirometric diagnosis of obstruction

$FEV_1/FVC < 70\%$  and a fixed cutoff of 80% of the predicted value for FVC were used for the diagnosis of obstructive and a restrictive pattern, respectively. The obstructive group was further classified into subgroups according to the severity of airway obstruction in accordance with the GOLD guidelines:<sup>4</sup>  $FEV_1/FVC < 70\%$ , in combination with  $FEV_1 \geq 80\%$  predicted (Stage 1), or  $50\% \leq FEV_1 < 80\%$  predicted (Stage 2), or  $30\% \leq FEV_1 < 50\%$  predicted (Stage 3), or  $FEV_1 \leq 30\%$  predicted (Stage 4).

## Analyses

Test of normality was conducted using the previously mentioned parameters. Measurements of parameters were reported as mean  $\pm$  SD for normally distributed variables.

## Correlation analysis

The linear correlation and simple linear regression model were used to determine the strength of relationship for each parameter between devices. Correlation coefficient ( $r$ ) and the 95% CI were computed for each reading. A scatter plot with a regression line was graphed for each parameter.

## Agreement analysis

A Bland–Altman graph was further created to assess the agreement between devices. The graph illustrated the mean differences ( $\bar{d}$ ) of between-device readings (IQspiro-Elite DX) compared with the corresponding averages [(IQspiro+Elite DX)/2]. The mean differences between the two spirometers were regarded as the estimated bias. The 95% limits of agreement (LoA), which reflects random error,

is expressed as  $\bar{d} \pm 1.96$  SD. The upper limit of agreement (UL) is  $\bar{d} + 1.96$  SD, and the lower limit of agreement (LL) is  $\bar{d} - 1.96$  SD. According to the ATS/ERS standards,<sup>13</sup> the accuracy criteria (biases) needed to be within  $\pm 3\%$  of the reading or  $\pm 0.050$  L for FVC and FEV<sub>1</sub>,  $\pm 5\%$  of the reading or  $\pm 0.200$  L/s for FEF<sub>25%-75%</sub> and  $\pm 10\%$  of the reading or  $\pm 20$  L/min for PEF, whichever is greater. The judgment of whether LoA or reliability was acceptable was determined by clinical applications, as they could not be proven by a statistical test. The acceptable 95% LoA was preliminarily set as  $\pm 0.5$  L for FVC and  $\pm 0.35$  L for FEV<sub>1</sub>, as reported previously.<sup>5,15</sup>

### Receiver operating characteristic (ROC) curve analysis

An ROC curve was used to determine the FEV<sub>1</sub>/FEV<sub>6</sub> ratio derived from the D-PNEU spirometer that corresponded best to the commonly used fixed cutoff point for FEV<sub>1</sub>/FVC < 70% from the laboratory-based spirometer.

### Kappa statistics

Using FEV<sub>1</sub>/FVC < 70% as the “gold standard” for obstruction, overall agreement in classification of the severity of airway obstruction between devices was assessed using kappa statistical methodology. A Cohen’s kappa test was used to calculate the kappa coefficient ( $\kappa$ ), which indicates the strength of diagnosis agreement based on its magnitude.<sup>16</sup> The kappa values were qualified according to their level of agreement: 0.4–0.6, moderate agreement; 0.6–0.8, substantial agreement; and 0.8–1.0, almost perfect agreement.<sup>17</sup>

## Results

A total of 241 individuals (180 men, 61 women) aged from 14 to 81 years (mean age  $55 \pm 14.7$  years) completed PFTs with both spirometers. Overall, 190 (79%) subjects performed forced expiratory maneuvers meeting the reproducibility standard based on ATS/ERS standardization of spirometry using the D-PNEU spirometer, while 192 (80%) did so using the traditional spirometer. One hundred and nine (45%) were healthy, 97 (40%) subjects suffered from airway obstruction, 24 (10%) had mixed airway dysfunction, and eleven (5%) suffered from restrictive disease only. In the end, the total number of non-obstructed subjects was 120 [(109+11), 50%].

A strong linear relationship was found between devices for all parameters (Figure 2, Table 1). As shown in Figure 2, all parameters, with the exception of FVC, had a small tendency to be underestimated by the D-PNEU device.

The Bland–Altman plot was drawn to display the mean difference ( $\bar{d}$ ) or bias and 95% LoA ( $\pm 1.96$  SD) between devices for each value measured (Figure 3). The values measured or derived were shown in Table 2. There was no significant difference for the mean difference of FVC (within  $\pm 0.05$  L or  $\pm 3\%$ ) between spirometers, while significant differences were found for that of FEV<sub>1</sub>, FEV<sub>1</sub>/FVC, FEF<sub>25%-75%</sub> and PEF. Nevertheless, the biases remained within acceptable limits for FEV<sub>1</sub> (within  $\pm 0.05$  L or  $\pm 3\%$ ), FEF<sub>25%-75%</sub> (within  $\pm 0.2$  L/s or  $\pm 5\%$ ), and PEF (within  $\pm 20$  L/min or  $\pm 10\%$ ). As previously described, the evaluation criteria of the parameters measured were: met either the absolute value of the difference, or percentage of reading (accuracy) between the devices, whichever is greater. The 95% LoA showed that the LL and UL for FVC were  $-0.38$  L and  $0.42$  L (within  $\pm 0.5$  L),  $-0.29$  L and  $0.21$  L for FEV<sub>1</sub> (within  $\pm 0.35$  L), respectively. Taken together, the findings suggested that between-device agreement was well within clinically acceptable limits.

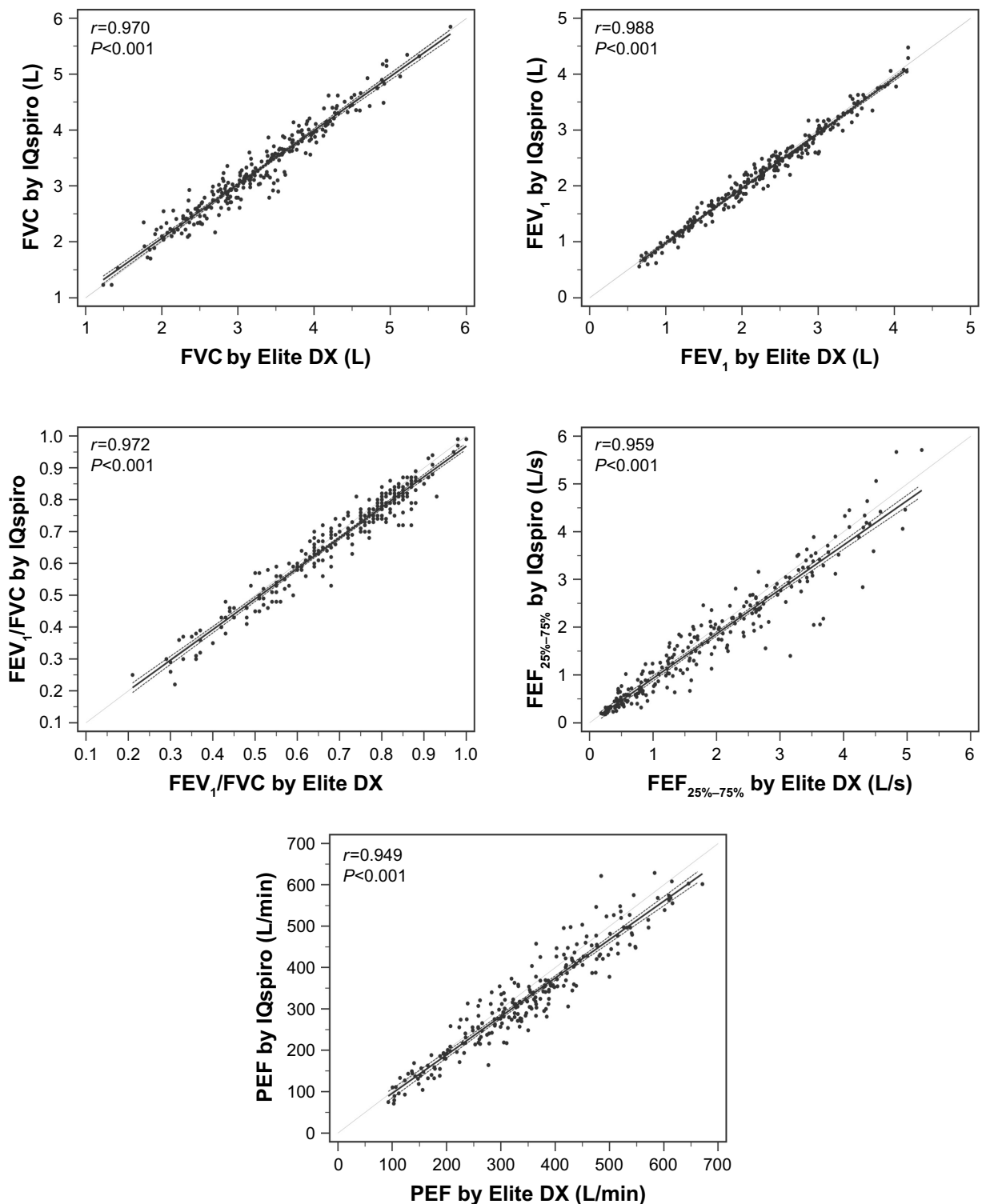
As shown in Table 3, the accuracy of spirometric classification of the stages of obstruction severity was 87.1% [(100+57+28+25)/241] for all enrolled subjects (n=241), regardless of expiratory times. The Kappa coefficient ( $\kappa$ ) was 0.869 (95% CI, 0.822–0.915).

A total of 159 individuals were able to exhale for at least 6 seconds using both devices. Their data were analyzed to determine the role of FEV<sub>1</sub>/FEV<sub>6</sub> ratio as alternative index to FEV<sub>1</sub>/FVC in the detection of airway obstruction. Using FEV<sub>1</sub>/FVC < 70% from conventional spirometry as the definition of airflow limitation, an ROC curve was analyzed to establish a cutoff point for FEV<sub>1</sub>/FEV<sub>6</sub> ratio obtained using the D-PNEU device, that corresponded to the optimal combination of sensitivity and specificity. The area under the ROC curve was 0.983 (95% CI: 0.949–0.997), and the sensitivity (94.6%) and specificity (92.4%) were at their maximum when the FEV<sub>1</sub>/FEV<sub>6</sub> obtained using D-PNEU device was 74% (Figure 4A). The maximum accuracy of spirometric classification was 93.7% [(61+88)/159], while the kappa coefficient was highest at 0.870 (Table 4).

In comparison to the FEV<sub>1</sub>/FEV<sub>6</sub> < 74% as optimal cutoff point, the ROC analysis showed less sensitivity (78.5%) and slightly greater specificity (97.0%), when FEV<sub>1</sub>/FEV<sub>6</sub> was set as 70% (Figure 4B). The accuracy of spirometric classification decreased to 86.2% [(64+73)/159], when kappa coefficient was 0.726 (Table 5).

## Discussion

The aim of the current study was to investigate the accuracy and validity of handheld spirometry, D-PNEU, as compared



**Figure 2** Correlation of key airflow metrics between devices (n=241).

**Note:** Line of identity (thin solid line), line of regression (thick solid line), and 95% CIs (dotted line).

**Abbreviation:** FEF<sub>25%-75%</sub> mean FEF calculated between 25% and 75% of FVC.

to full body plethysmography. The primary goal was to assess the agreement for key airflow metrics (FVC, FEV<sub>1</sub>, FEV<sub>1</sub>/FVC, FEF<sub>25%-75%</sub> and PEF) between the D-PNEU device and a laboratory-based spirometer. The second goal

was to evaluate the diagnostic accuracy of spirometric classification of airflow obstruction by using the D-PNEU device, as compared to a laboratory-based spirometer. Firstly, we compared diagnostic accuracy for all subjects, regardless

**Table 1** Correlation coefficient for all indices measured (n=241)

Index	Correlation coefficient (95% CI)	P-value
FVC	0.970 (0.961–0.977)	<0.001
FEV <sub>1</sub>	0.988 (0.984–0.990)	<0.001
FEV <sub>1</sub> /FVC	0.972 (0.964–0.978)	<0.001
FEF <sub>25%–75%</sub>	0.959 (0.948–0.968)	<0.001
PEF	0.949 (0.935–0.961)	<0.001

**Abbreviation:** FEF<sub>25%–75%</sub>, mean FEF calculated between 25% and 75% of FVC.

of the expiratory times of forced expiratory maneuvers. Secondly, we investigated the diagnostic accuracy for the specific subjects able to exhale for at least 6 seconds.

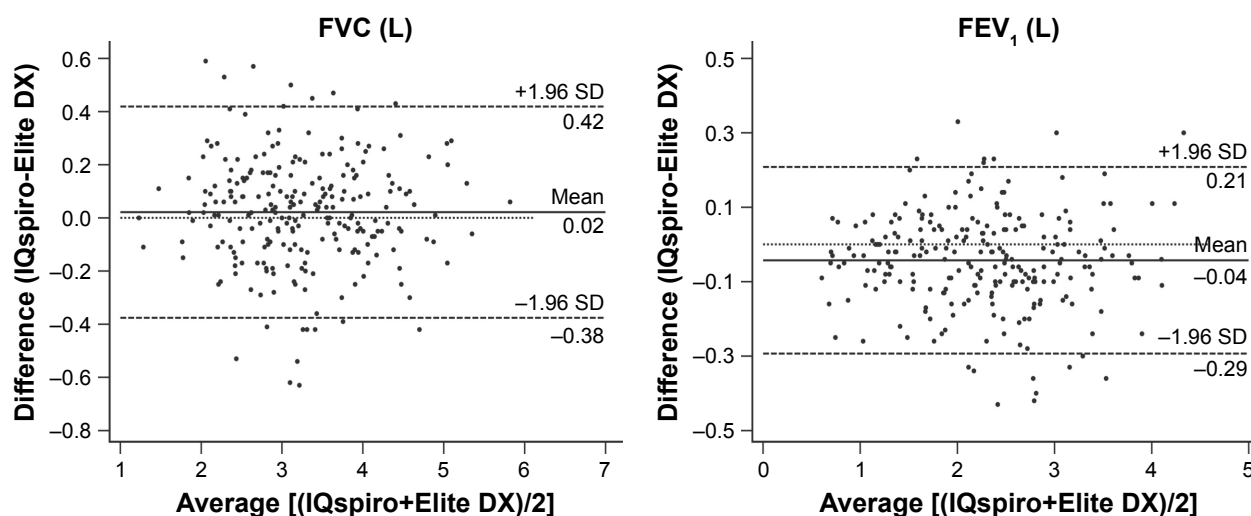
In the present study, the between-device agreement using Bland–Altman analysis showed that the biases (mean differences) of directly measured parameters were well within clinically acceptable limits. Our study demonstrates that using FEV<sub>1</sub>/FVC <70% as the “gold standard” for obstruction, the D-PNEU device was 87.1% accurate in the classification of severity of obstruction for all enrolled subjects. For the specific subjects able to exhale for at least 6 seconds, when a fixed cutoff of FEV<sub>1</sub>/FEV<sub>0</sub> ratio derived from the PNEU was 74%, the D-PNEU device was 93.7% accurate in identifying subjects with or without airflow limitation (defined as FEV<sub>1</sub>/FVC <70% using full body spirometry).

Spirometry is necessary to avoid misdiagnosis and to ensure proper determination of the severity of airflow limitation.<sup>4</sup> It is often underused in China, particularly in rural areas. A previous study reported that in China, only 6.5% of patients with COPD had ever been tested by spirometry.<sup>3</sup> This may be attributable to several factors, including large disparities in health care resources, negligence of primary

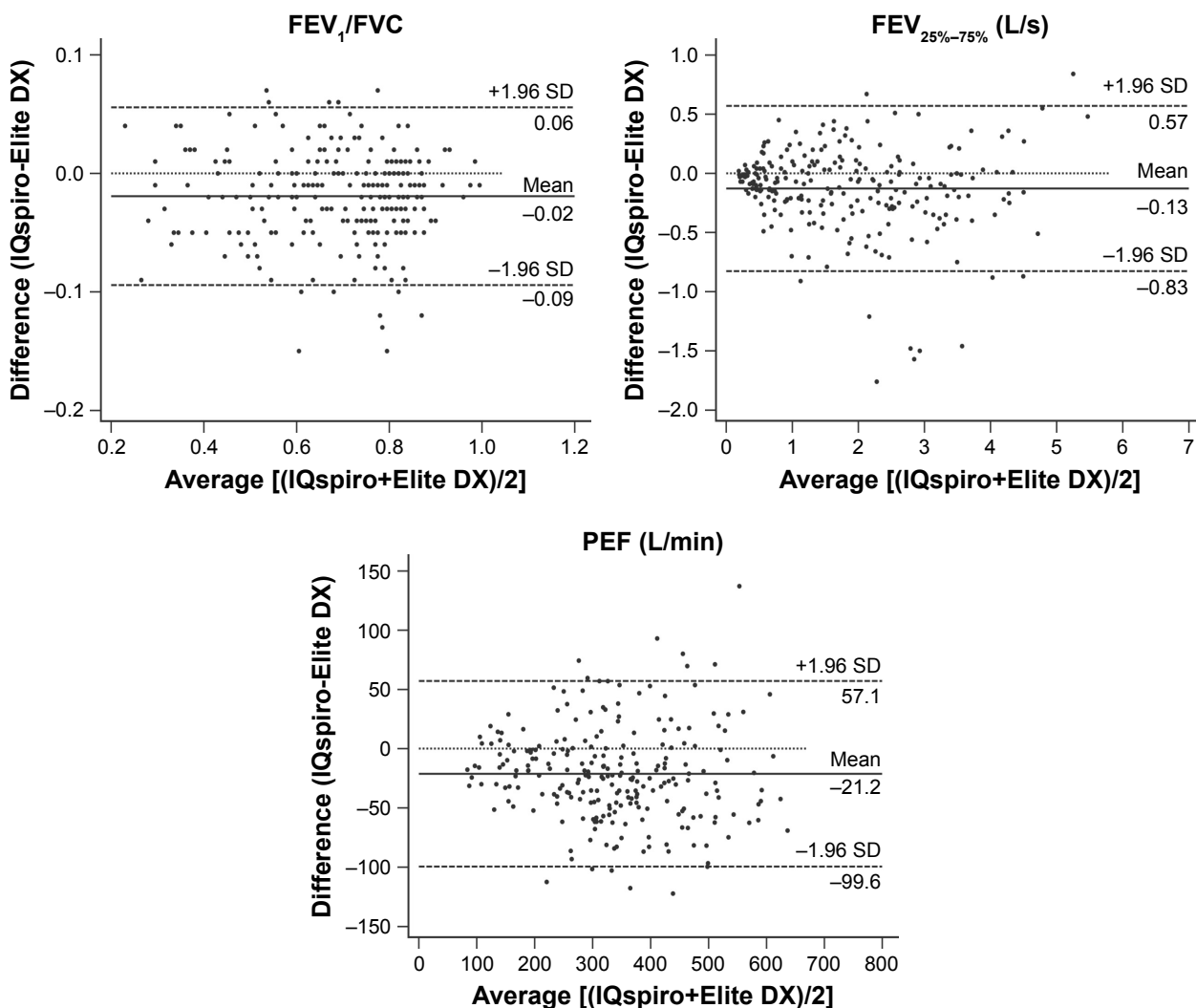
care physicians and the patient. At present, a disproportionate amount of China’s health care resources has traditionally been allocated to larger hospitals, particularly those in urban areas. More than 80% of health expenditures are allocated to urban areas even though 70% of the total population resides in rural areas.<sup>18</sup> The investigations highlight the urgent need for the physicians to leverage spirometry as a central tool in the proper diagnosis and management of COPD.

It is important to be sure that the new medical device is as accurate as the gold standard method. Therefore, the primary goal was to measure the agreement of the portable D-PNEU spirometer with the established gold standard spirometer. The old favorite for measuring agreement is the correlation coefficient. In this study, correlation analysis suggested that nearly all the key parameters had a small tendency to be underestimated by the D-PNEU spirometer. A strong correlation was found between devices for all parameters with the high correlation coefficient ranging from 0.949 to 0.988. However, this is inappropriate as correlation only measures the strength of linear association between variables, which has been discussed by Altman and Bland since the 1980s.<sup>9</sup> Therefore, merely using correlation coefficient is not enough for assessing agreement.

Several systematic reviews have identified that the Bland–Altman method is currently the appropriate and indeed the most popular method, that has been used to assess agreement between medical instruments measuring continuous variables.<sup>7,8</sup> The Bland–Altman plot analysis is a simple way to evaluate a bias between the mean differences, and to estimate an agreement interval, within which 95% of the differences of the second method, compared to the first one fall.



**Figure 3** (Continued)



**Figure 3** Bland–Altman plot of mean differences against averages of two readings for key airflow metrics (n=241).  
**Notes:** The 95% limits of agreement (LoA) are expressed as mean differences ( $\bar{d}$ ) $\pm$ 1.96 SD; the upper limit of agreement (UL) is  $\bar{d}$ +1.96 SD, and the lower limit of agreement (LL) is  $\bar{d}$ -1.96 SD. Line of mean difference (solid line), lines of 95% LoA (dashed line), and zero line (dotted line).  
**Abbreviation:** FEF<sub>25%-75%</sub> mean FEF calculated between 25% and 75% of FVC.

These studies also pointed out that the Bland–Altman method only defines the mean differences (biases) and LoA, it does not say whether those biases or limits are acceptable or not.<sup>7,8</sup> Acceptable limits must be defined a priori, based on clinical

**Table 2** The mean difference and 95% LoA for all parameters measured by the two spirometers (n=241)

Variable	$\bar{d}$ (bias) <sup>a</sup>	Accuracy	P-value	LoA (LL, UL) <sup>b</sup>
FVC	0.02L	0.61%	>0.05	(-0.38, 0.42)
FEV <sub>1</sub>	-0.04L	-1.67%	<0.05	(-0.29, 0.21)
FEV <sub>1</sub> /FVC	-0.02	-2.74%	<0.05	(-0.09, 0.06)
FEF <sub>25%-75%</sub>	-0.13 L/s	-6.73%	<0.05	(-0.83, 0.57)
PEF	-21.2 L/min	-6.00%	<0.05	(-99.6, 57.1)

**Notes:** <sup>a</sup>The mean differences ( $\bar{d}$ ) between the two spirometers were regarded as the estimated bias; <sup>b</sup>The 95% limits of agreement (LoA) are expressed as  $\bar{d}\pm 1.96$  SD; the upper limit of agreement (UL) is  $\bar{d}+1.96$  SD, and the lower limit of agreement (LL) is  $\bar{d}-1.96$  SD.  
**Abbreviation:** FEF<sub>25%-75%</sub> mean FEF calculated between 25% and 75% of FVC.

necessity. In the present study, the criteria for the acceptable biases and the 95% LoA were described in the “Analyses” section. We demonstrated that no significant difference of FVC was found for between-devices bias. Although significant differences of biases were found for the directly measured parameters, including FEV<sub>1</sub>, FEF<sub>25%-75%</sub>, and PEF. According to the accuracy criteria issued by ATS/ERS standards, described in the “Analyses” section,<sup>13</sup> these biases remained within clinically acceptable limits. The Bland–Altman analysis also revealed that except for FVC, the biases of all parameters were slightly negative, which confirmed our finding in the correlation analysis that the parameters were generally underestimated by D-PNEU spirometer when compared to conventional spirometer. In addition, we found that the 95% LoA for FVC and FEV<sub>1</sub> were within

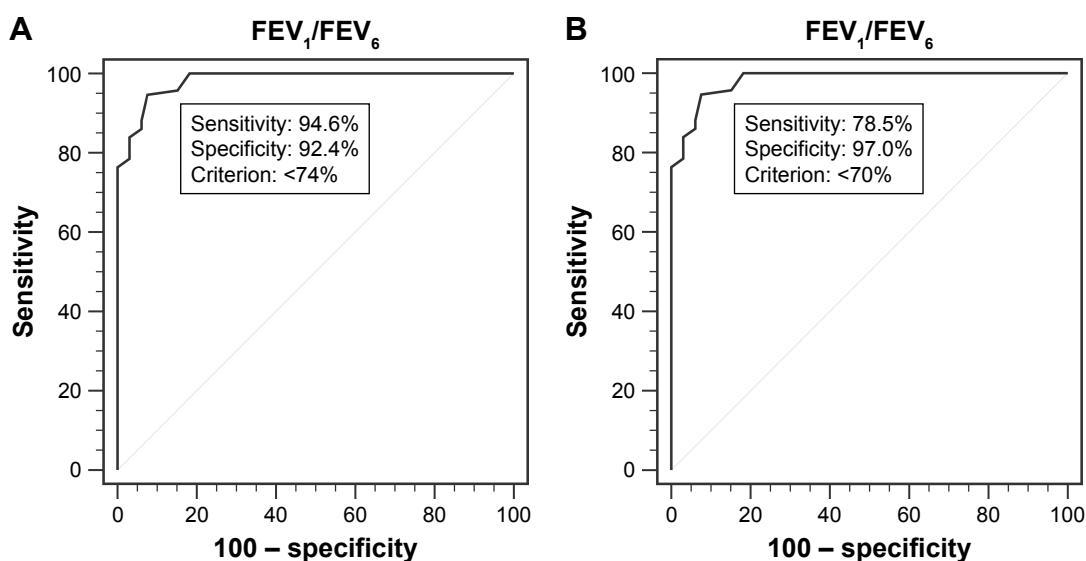
**Table 3** Spirometric classification of severity of airway obstruction by two spirometers for all enrolled subjects (n=241)

IQspiro (FEV <sub>1</sub> /FVC <70%)	Elite DX (FEV <sub>1</sub> /FVC <70%)				Total
	Non-obstruction	Obstruction (Stage 1)	Obstruction (Stage 2)	Obstruction (Stage 3 and 4)	
Non-obstruction	100	0	0	0	100
Obstruction (Stage 1)	17	57	4	0	78
Obstruction (Stage 2)	3	4	28	0	35
Obstruction (Stage 3 and 4)	0	0	3	25	28
Total	120	61	35	25	241

the acceptable range (“Analyses” section), as previously reported by other studies.<sup>5,15</sup> Based on the view of specialized field, the magnitude of the range for FEF<sub>25%-75%</sub> and PEF is more dependent on the force of initial portion of the FVC maneuver.<sup>19,20</sup> These observations deserve further investigation. Taken together, our study demonstrated clinically acceptable agreements for the key airflow metrics between the D-PNEU device and a laboratory-based spirometer. This is the prerequisite to the subsequent investigation of clinical accuracy of the D-PNEU device.

The second goal of current study was to evaluate the diagnostic accuracy of spirometric classification of airflow obstruction by using a D-PNEU device, as compared to a laboratory-based spirometer. First, we compared diagnostic accuracy for all subjects, regardless of the expiratory times of forced expiratory maneuvers. The accuracy of spirometric classification was 87.1% for the total 241 subjects referred to the PFT laboratory, when using FEV<sub>1</sub>/FVC <70% as the “gold standard” for obstruction for both devices. Secondly, we investigated the diagnostic

accuracy for specific subjects able to exhale for at least 6 seconds. As subjects with airflow obstruction have prolonged expiration, the use of FEV<sub>1</sub>/FVC alone may result in underdiagnosis of airflow obstruction in younger people and overdiagnosis in the elderly, particularly in patients with moderate-to-severe airflow obstruction or in elderly subjects.<sup>21-23</sup> FEV<sub>6</sub> has been proposed as a simplified alternative to an FVC maneuver.<sup>10,12,24</sup> Indeed, the ratio of the FEV<sub>1</sub>/FEV<sub>6</sub> has been found nearly equivalent to FEV<sub>1</sub>/FVC for the diagnosis of airway obstruction, but the former is simpler, causes less fatigue, and is possibly more reliable than FEV<sub>1</sub>/FVC because FVC varies with the duration of the forced exhalation.<sup>10,25,26</sup> The accuracy of spirometric classification was 93.7% for the 159 individuals able to exhale for at least 6 seconds, when a fixed cutoff point of 74% for FEV<sub>1</sub>/FEV<sub>6</sub> was calculated by an ROC curve analysis. The Kappa coefficient in both of these two analyses was greater than 0.8, which indicated that the diagnostic validity of airflow obstruction between the two devices was “almost perfect agreement beyond chance”.<sup>17</sup> Our findings thus

**Figure 4** The receiver operating characteristic curve that discriminates the optimal cutoff value of FEV<sub>1</sub>/FEV<sub>6</sub> measured using the IQspiro (n=159).

**Notes:** (A) The sensitivity and specificity when the FEV<sub>1</sub>/FEV<sub>6</sub> was <74%. (B) The sensitivity and specificity when the FEV<sub>1</sub>/FEV<sub>6</sub> was <70%.



**Table 4** Spirometric classification of airway obstruction by two spirometers for individuals able to exhale for at least 6 seconds ( $FEV_1/FEV_6 < 74\%$ ) (n=159)

IQspiro ( $FEV_1/FEV_6 < 74\%$ )	Elite DX ( $FEV_1/FVC < 70\%$ )		Total
	Negative	Positive	
Negative	61 (92.4%)	5 (5.4%)	66
Positive	5 (7.6%)	88 (94.6%)	93
Total	66 (100%)	93 (100%)	159

confirmed the reliability of the tested D-PNEU device as a useful diagnostic solution for the spirometric classification of various levels of airway obstruction.

Three core elements are essential to capturing valid pulmonary function data: accurate instrumentation; cooperative testing subject; pulmonary function technologist capable of coaching patients during pulmonary function testing; and properly interpreting test result.<sup>27</sup> The management software of the currently tested device is responsible for quality control of maneuvers with a choice of easy-to-see incentives on a full screen during spirometry. The results are also automatically interpreted and displayed on the computer. This study thus suggested that a handheld spirometer such as the D-PNEU spirometer, is an easy-to-use device from the perspective of the physician. Such tools are suitable for use in primary care as they are relatively accurate when compared to full body plethysmography.<sup>5,19,28–32</sup>

We admit several limitations in this study. First, no significant difference for the bias of FVC was observed between spirometers; however, it does not necessarily mean the measurement of FVC by two devices can be used interchangeably. Second, the study was designed to determine the agreement and validity of a portable spirometer, thus the bronchiectasis test was not routinely performed as part of our procedure. In addition, we did not compare the agreement and validity of other office spirometers available on the market. It is noticeable that some of these devices are more accurate and precise than others.

**Table 5** Spirometric classification of airway obstruction by two spirometers for individuals able to exhale for at least 6 seconds ( $FEV_1/FEV_6 < 70\%$ ) (n=159)

IQspiro ( $FEV_1/FEV_6 < 70\%$ )	Elite DX ( $FEV_1/FVC < 70\%$ )		Total
	Negative	Positive	
Negative	64 (97.0%)	20 (21.5%)	84
Positive	2 (3.0%)	73 (78.5%)	75
Total	66 (100%)	93 (100%)	159

## Conclusion

Our findings suggest the clinical-grade handheld spirometer, the D-PNEU device under investigation in current research, is a relatively accurate and easy-to-use device able to identify subjects who have airway obstruction. In addition, our study further supports that  $FEV_1/FEV_6$  ratio can be used as a valid alternative for  $FEV_1/FVC$ . Such kind of portable spirometers might likely facilitate earlier diagnosis and management of COPD in the rural areas of China.

## Acknowledgments

We express our sincere thanks to Yiqun Li, the technician in the pulmonary function laboratory of Sun Yat-sen Memorial Hospital. Dr Guojun Chen is now working at the Department of Emergency Medicine, The First People's Hospital of Foshan, Guangdong, China.

This study was supported by a grant from the National Natural Science Foundation of China (81272061).

## Author contributions

XF had full access to all the data in the study, was involved in conception, design of the study, revision of the manuscript for appropriateness of the contents, and takes responsibility for the integrity of the data and the accuracy of the data analysis. GC and LJ contributed substantially to the study concept and design, data collection, data analysis and interpretation, and the writing of the manuscript. LW, WZ and CC contributed substantially to the study design, data analysis and interpretation, and the writing of the manuscript. All authors have read and approved the final version of the manuscript.

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Fang X, Wang X, Bai C. COPD in China: the burden and importance of proper management. *Chest*. 2011;139(4):920–929.
- Yoon HI, Sin DD. Confronting the colossal crisis of COPD in China. *Chest*. 2011;139(4):735–736.
- Zhong N, Wang C, Yao W, et al. Prevalence of chronic obstructive pulmonary disease in China: a large, population-based survey. *Am J Respir Crit Care Med*. 2007;176(8):753–760.
- Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187(4):347–365.
- Liistro G, Vanwelde C, Vincken W, et al. Technical and functional assessment of 10 office spirometers: a multicenter comparative study. *Chest*. 2006;130(3):657–665.
- Kjeldgaard P, Lykkegaard J, Spillemoose H, Ulrik CS. Multicenter study of the COPD-6 screening device: feasible for early detection of chronic obstructive pulmonary disease in primary care? *Int J Chron Obstruct Pulmon Dis*. 2017;12:2323–2331.

7. Giavarina D. Understanding Bland Altman analysis. *Biochem Med.* 2015;25(2):141–151.
8. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. *PLoS One.* 2012;7(5):e37908.
9. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307–310.
10. Vandevoorde J, Verbanck S, Schuermans D, Kartounian J, Vincken W. FEV1/FEV6 and FEV6 as an alternative for FEV1/FVC and FVC in the spirometric detection of airway obstruction and restriction. *Chest.* 2005;127(5):1560–1564.
11. Lamprecht B, Schirnhofner L, Tiefenbacher F, et al. Six-second spirometry for detection of airway obstruction: a population-based study in Austria. *Am J Respir Crit Care Med.* 2007;176(5):460–464.
12. Akpınar-Elci M, Fedan KB, Enright PL. FEV6 as a surrogate for FVC in detecting airways obstruction and restriction in the workplace. *Eur Respir J.* 2006;27(2):374–377.
13. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J.* 2005;26(2):319–338.
14. Hou S, Zhan YG, Guo XC, He J. Normal values of lung function in the population of Guangdong. In: Mu KJ, Liu SW, editors. *Nationwide Normal Values of Lung Function.* 1st ed. PUMC & Beijing Medical University; 1990:67–80.
15. Gerbase MW, Dupuis-Lozeron E, Schindler C, et al. Agreement between spirometers: a challenge in the follow-up of patients and populations? *Respiration.* 2013;85(6):505–514.
16. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20(1):37–46.
17. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, For GG. Tips for learners of evidence-based medicine: 3. Measures of observer variability (kappa statistic). *Can Med Assoc J.* 2004;171(11):1369–1373.
18. IBM Institute for Business Value. Healthcare in China: Toward greater access, efficiency and quality. New York: IBM Institute for Business Value; 2006. Available from: <https://www-935.ibm.com/services/us/imc/pdf/g510-6268-healthcare-china.pdf>. Accessed June 26, 2018.
19. Rebeck DA, Hanania NA, D'Urzo AD, Chapman KR. The accuracy of a handheld portable spirometer. *Chest.* 1996;109(1):152–157.
20. Wiltshire N, Kendrick AH. Evaluation of a new electronic spirometer: the vitalograph “Escort” spirometer. *Thorax.* 1994;49(2):175–178.
21. Hansen JE, Sun X-G, Wasserman K. Discriminating measures and normal values for expiratory obstruction. *Chest.* 2006;129(2):369–377.
22. Bhatt NY, Wood KL. What defines abnormal lung function in older adults with chronic obstructive pulmonary disease? *Drugs Aging.* 2008;25(9):717–728.
23. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med.* 1999;159(1):179–187.
24. Swanney MP, Jensen RL, Crichton DA, Beckert LE, Cardno LA, Crapo RO. FEV<sub>6</sub> is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am J Respir Crit Care Med.* 2000;162(3):917–919.
25. Pérez-Padilla R, Hallal PC, Vázquez-García JC, et al. Impact of bronchodilator use on the prevalence of COPD in population-based samples. *COPD.* 2007;4(2):113–120.
26. Perez-Padilla R, Wehrmeister FC, Celli BR, et al. Reliability of FEV1/FEV6 to diagnose airflow obstruction compared with FEV1/FVC: The PLATINO Longitudinal Study. *PLoS One.* 2013;8(8):e67960.
27. Haynes JM. Quality assurance of the pulmonary function technologist. *Respir Care.* 2012;57(1):114–126.
28. Rothe T, Karrer W, Schindler C. Accuracy of the Piko-1 pocket spirometer. *J Asthma.* 2012;49(1):45–50.
29. Ching SM, Pang YK, Price D, et al. Detection of airflow limitation using a handheld spirometer in a primary care setting. *Respirology.* 2014;19(5):689–693.
30. Sichelidis L, Spyros D, Papaioannou M, et al. A combination of the IPAG questionnaire and PiKo-6® flow meter is a valuable screening tool for COPD in the primary care setting. *Prim Care Respir J.* 2011;20(2):184–189.
31. Nishimura K, Nakayasu K, Kobayashi A, Mitsuma S. Case identification of subjects with airflow limitations using the handheld spirometer “Hi-Checker™”: comparison against an electronic desktop spirometer. *COPD.* 2011;8(6):450–455.
32. Barr RG, Stemple KJ, Mesia-Vela S, et al. Reproducibility and validity of a handheld spirometer. *Respir Care.* 2008;53(4):433–441.

## International Journal of COPD

### Publish your work in this journal

The International Journal of COPD is an international, peer-reviewed journal of therapeutics and pharmacology focusing on concise rapid reporting of clinical studies and reviews in COPD. Special focus is given to the pathophysiological processes underlying the disease, intervention programs, patient focused education, and self management protocols.

Submit your manuscript here: <http://www.dovepress.com/international-journal-of-chronic-obstructive-pulmonary-disease-journal>

Dovepress

This journal is indexed on PubMed Central, MedLine and CAS. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.