

Reliability of Pulmonary Function Tests in Patients with Chronic Obstructive Pulmonary Disease

Fatma Arslan¹, Elif Şen¹, Burak Şirin², Nurdan Köktürk³, Gaye Ulubay⁴, Selma Aydoğan Eroglu⁵, Bilun Gemicioglu⁶, Özlem Sönmez⁷, İrem Şerifoğlu⁸, Asli Gorek Dilektasli⁹, Burcu Öztürk Şahin¹⁰, İlkey Doğan¹¹, Nermin Gürhan¹², Hakan Günen⁵, Nihal Ensen⁶

¹Department of Pulmonary Medicine, Ankara University Faculty of Medicine, Ankara, Türkiye; ²Faculty of Health Sciences, Department of Nursing, Tokat Gaziosmanpaşa University, Tokat, Türkiye; ³Department of Pulmonary Medicine, Gazi University Faculty of Medicine, Ankara, Türkiye; ⁴Department of Pulmonary Medicine, Başkent University Faculty of Medicine, Ankara, Türkiye; ⁵Department of Pulmonary Medicine, University of Health Sciences, İstanbul Sureyyapasa Chest Diseases and Thoracic Surgery Training and Research Hospital, İstanbul, Türkiye; ⁶Department of Chest Diseases, İstanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, İstanbul, Türkiye; ⁷Department of Chest Diseases, Ankara Atatürk Sanatorium Training and Research Hospital, Ankara, Türkiye; ⁸Department of Chest Diseases, Ankara Bilkent City Hospital, Ankara, Türkiye; ⁹Department of Chest Diseases, Uludağ University Faculty of Medicine, Bursa, Türkiye; ¹⁰Department of Chest Diseases, Gülhane Faculty of Medicine, University of Health Sciences, Ankara, Türkiye; ¹¹Department of Biostatistics and Medical Informatics, Gaziantep University Faculty of Medicine, Gaziantep, Türkiye; ¹²OKA Congress and Health Tourism Institute, Tokat, Türkiye

Correspondence: Fatma Arslan, Department of Pulmonary Medicine, Ankara University Faculty of Medicine, Cebeci Campus, Ankara, 06220, Türkiye, Email fciftci@ankara.edu.tr

Introduction: Chronic Obstructive Pulmonary Disease (COPD) is frequently underdiagnosed due to limited use and suboptimal quality of spirometry.

Methods: This multicenter, prospective observational study aimed to assess the quality of spirometry tests in COPD patients across six tertiary care centers in Turkey. A total of 560 spirometry tests were independently evaluated by two pulmonologists, and inter-rater reliability was analyzed using Cohen's Kappa.

Results: The study included 124 women and 436 men, with a mean age of 65.9±10.2 years. Spirometry patterns were interpreted as obstructive in 79.1% of tests, restrictive in 5.7%, mixed in 2.7%, normal in 6.7%, inconclusive in 5.5%, and as upper airway obstruction in 0.18% of cases ($K=0.890$). Routine post-bronchodilator testing was performed in 87.14% of the cases, while 12.85% were post-reversibility tests. Graphical data included volume-time curves (88.5%), flow-volume loops (99.6%), and inspiratory limbs (98.7%). Reference values were predominantly based on ECSC (86.4%), followed by GLI (8.9%) and NHANES (0.25%). Mean spirometric values included FEV1 (% predicted) 57.2%, FVC (% predicted) 72.7%, and FEV1/FVC 61.9%. Cough in the first second of FVC was observed in 17 (3%) ($K=0.677$), variable or insufficient effort 13.4% ($K=0.563$), mouth leak 1.8% ($K=0.305$), mouthpiece obstruction 0.5% ($K=0.332$), rapid expiration 1.8% ($K=0.354$), increased concavity 83% ($K=0.683$), and small airway obstruction 87% ($K=0.709$) were reported by interpreters.

Conclusion: Spirometry reports frequently lack key information such as preliminary diagnoses, test indication, technician, and device details. ECSC is the most commonly used reference. Rates of inconclusive and erroneous tests—due to variable or insufficient effort, mouth leak, mouthpiece obstruction, and rapid expiration—are low.

Keywords: chronic obstructive pulmonary disease, COPD, pulmonary function tests, spirometry

Background

Chronic Obstructive Pulmonary Disease (COPD) is a heterogeneous lung disease characterized by persistent respiratory symptoms such as dyspnea, coughing, sputum production, and/or exacerbations. It results from structural changes in the airways and alveoli that lead to persistent, often progressive airflow limitation.¹ In the appropriate clinical context, airflow obstruction that is not fully reversible, demonstrated by spirometry, is the most important diagnostic finding. Although COPD is a well-established lung disease, most of the subjects remain underdiagnosed in population based studies.² The most significant reason of underdiagnosis is the low rate of pulmonary function tests (PFT), especially

spirometry utilization. Additionally, technical errors related to the quality of spirometric tests frequently result in incorrect diagnoses.³

Previous research has identified that the proportion of high-quality spirometry in COPD cases was only 22.6% of the study group.⁴ These findings highlight significant challenges in access to spirometry and the proper execution of spirometry tests on a national scale.

Recent reports have emphasized that the reliability of spirometry largely depends on strict adherence to international standards for performance, calibration, and quality control.^{5–7} Studies from different countries have shown that sub-optimal test quality, inadequate technician training, and inconsistent use of reference equations significantly affect diagnostic accuracy and the estimation of COPD prevalence.^{8,9}

Therefore, evaluating the technical quality and reporting standards of pulmonary function tests in routine clinical practice remains crucial for ensuring diagnostic validity and improving patient outcomes.

This study aims to assess the quality of spirometric testing in patients diagnosed with COPD. Understanding the errors in PFT reporting and interpretation could be valuable in developing and targeting interventions to improve COPD diagnosis.

Methods

This multicenter, prospective, and observational study was conducted by collecting spirometry test reports from six tertiary healthcare centers across three different cities in Turkey that agreed to participate between May and July 2020. The PFT laboratories of the participating centers met the standards, were qualified, and staffed with experienced PFT technicians who were well-trained and skilled. The study included post-bronchodilator spirometry reports of patients over 40 years of age, with a history of smoking >10 pack-years, and who were previously followed with a confirmed diagnosis of COPD based on clinical and pulmonary function tests (Table 1). These test reports were evaluated by two independent pulmonary specialists with expertise in the field.

The reports were reviewed by the interpreters to collect information on patient demographics, body composition, technician details, device specifications, flow rates, reference values, graphs, test noncompliance, and errors. These details were then recorded in patient follow-up forms.

Statistical analyses were performed using SPSS version 22.0, with a significance level set at $p < 0.05$. Descriptive statistics were employed to present data with a normal distribution, expressed as numbers (percentage) and mean \pm standard deviation or median (min-max). The statistical reliability between the two interpreters was measured using Cohen's Kappa Test (K). A Kappa coefficient 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, 0.81–1.00 almost perfect agreement was interpreted as indicating agreement between the two interpreters.¹⁰

In cases with very low event frequency or high agreement by chance, prevalence-adjusted bias-adjusted kappa (PABAK) values were also computed to account for the potential influence of prevalence and bias on K estimates.

Table 1 Distribution of Pulmonary Function Tests Evaluated from the Participating Centers

Participating PFT Laboratories	n
Gazi University Hospital	100 (17.85)
Uludağ University Hospital	50 (8.93)
Süreyyapaşa Hospital	141 (25.17)
Başkent University Hospital	70 (12.5)
İstanbul University Cerrah paşa Hospital	100 (17.85)
Ankara Atatürk Sanatory Education and Research Hospital	99 (17.68)
Results were given as n (%)	

These adjusted values were used only for error categories with limited occurrences, such as mouth leak and mouthpiece obstruction.

Ethics Committee approval was granted by the Gazi University (Assessment and Evaluation Ethics Sub-Working Group) Human Research Ethics Committee with no 2019–106.

Results

A total of 560 spirometry tests reports of distinct patients were included in the study were reviewed by two independent interpreters. The mean age of the patients was found to be 65.92 ± 10.25 years. The study population consisted of 124 (22.14%) female and 436 (77.85%) male cases, with a mean body mass index (BMI) of 27.81 ± 6.14 kg/m².

The interpreters classified 443 (79.1%) tests as showing obstruction, 32 (5.7%) as suggestive of restriction, 15 (2.68%) as a mixed pattern, 38 (6.7%) as normal, 31 (5.53%) as non-compliant, and 1 (0.18%) as upper airway obstruction (Table 2).

When the spirometry test reports were reviewed, no written diagnosis was found in 289 (51.6%) tests. Obstruction was reported in 5 (1.84%) cases, and 4 (1.47%) cases were documented as erroneous tests. For 15 (2.68%) cases, it was indicated that the test was the first one, 291 (51.96%) cases were identified as follow-up tests, and no information regarding test sequence was provided for 254 (45.35%) cases.

Of all the spirometry tests, 488 (87.14%) were routine post-bronchodilator tests, while 72 (12.85%) were post-reversibility tests. Among the post-reversibility test group (n=72), the mean increase in FEV₁ (% predicted) was 7 ± 1.2 .

As for the reference values, 343 (86.4%) tests used ECSC, 36 (8.93%) tests used GLI, and 1 (0.25%) test used NHANES. In 180 (32.14%) reports, there was no information provided regarding which reference values were used. Some tests included both ECSC and GLI reference values.

The volume-time graph was present in 495 (88.5%) tests, while the flow-volume graph was available in 556 (99.64%) tests. The inspiratory limb was demonstrated in 550 (98.74%) tests. Increased concavity was reported in 83.2% of the tests ($K=0.683$), and small airway obstruction was noted in 87.1% of the tests ($K=0.709$) by the interpreters.

The mean FEV₁ and FVC values of the study group were 1.49 ± 0.7 L (57.17 ± 23.09 , % predicted) and 2.41 ± 0.95 L (72.7 ± 24.96 , % predicted), respectively, with a mean FEV₁/FVC ratio of 61.91 ± 14.31 . The mean forced expiratory time (FET) was calculated as 5.61 ± 1.88 seconds (Table 3).

Participants were stratified according to WHO BMI categories as underweight (n = 33, <18.5 kg/m²), normal weight (n = 290, 18.5–24.9 kg/m²), overweight (n = 148, 25.0–29.9 kg/m²), and obese (n = 89, ≥ 30 kg/m²). When all cases were categorized according to standard BMI classifications, no significant differences were observed in the overall spirometry values across these categories. The obese group had significantly lower mean FVC, L and FVC % predicted values compared with the non-obese participants (2.07 ± 0.34 vs 2.63 ± 0.85 L, $p = 0.04$; 70.1 ± 27.34 vs 73.5 ± 14.62 , $p=0.03$). Although the mean FEV₁/FVC ratio was higher in the obese group (63.91 ± 11.44 vs 60.23 ± 15.32), the difference was not statistically significant.

Table 2 Pulmonary Function Tests Interpretation Results

PFT Evaluation	n=560	K
Obstruction	443 (79.1)	0.877
Possibility of restriction	32 (5.7)	0.765
Mixt	15 (2.68)	0.856
Normal	38 (6.7)	0.803
Inconclusive	31 (5.53)	0.832
Upper airway obstruction	1 (0.18)	0.864

Notes: Results were given as n (%). A Kappa coefficient 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, 0.81–1.00 almost perfect agreement between two interpreters.

Abbreviation: PFT, Pulmonary function test.

Table 3 Maximal Expiratory Flow Rates and FET

PFT Results	n=560
FEV ₁ (L)	1.49 ± 0.7
FEV ₁ (predicted %)	57.17 ± 23.09
FVC (L)	2.41 ± 0.95
FVC (predicted %)	72.7 ± 24.97
FEV ₁ /FVC (%)	61.91 ± 14.31
FEF ₂₅₋₇₅ (predicted %)	31.81 ± 24.85
FET (sn)	5.61 ± 1.88

Notes: The results are presented as mean±SD.

Abbreviations: FEV₁, Forced expiratory volume in one second; FVC, Forced vital capacity; FEF 25–75, forced expiratory flow between 25 and 75% of vital capacity; FET, forced expiratory time.

Table 4 The Rates of Inconclusive and Erroneous Tests in Pulmonary Function Tests Evaluation

PFT Evaluation	n=560	K	95% CI (lower)	95% CI (upper)	PABAK
Cough in the first second of FVC	17 (3)	0.677	0.469	0.657	0.95
Variable or insufficient effort	67 (12)	0.563	0.435	0.646	0.89
Early termination	43 (7.7)	0.852	0.674	0.932	0.97
Mouth leak	17 (3)	0.305	0.278	0.435	0.94
Mouthpiece obstruction	5 (0.9)	0.332	0.290	0.421	0.96
Rapid expiration	19 (3.4)	0.354	0.306	0.412	0.93

Notes: The results are presented as n (%). A Kappa coefficient 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, 0.81–1.00 almost perfect agreement between two interpreters.

Abbreviations: PABAK, Prevalence-adjusted bias-adjusted kappa; PFT, Pulmonary function test; FVC, Forced vital capacity.

Regarding the inconclusive and erroneous test results identified by the two interpreters, variable and insufficient effort was the most frequently reported cause, noted in 67 (12%) tests ($K=0.563$). This was followed by early termination in 43 (7.7%) tests ($K=0.852$), rapid expiration in 19 (3.4%) tests ($K=0.354$), mouth leak in 17 (3%) tests ($K=0.305$), cough in the first second of respiration in 17 (3%) tests ($K=0.677$), and mouthpiece obstruction in 5 (0.9%) tests ($K=0.332$) (Table 4). Since the frequencies of some test errors were low, prevalence-adjusted bias-adjusted kappa (PABAK) analysis was performed as a more appropriate measure of inter-rater reliability. The PABAK values for mouth leak, mouthpiece obstruction, and rapid expiration were 0.94, 0.96, and 0.93, respectively (Table 4).

Discussion

In this study, spirometry reports of patients diagnosed with COPD from various centers were evaluated to assess the quality of spirometry tests across the country. COPD is defined as a chronic lung disease characterized by persistent, and often progressive, airflow obstruction resulting from abnormalities in the airways (bronchitis, bronchiolitis) and/or alveoli (emphysema), which lead to chronic respiratory symptoms such as dyspnea, cough, sputum production, and/or exacerbations.¹

The post-bronchodilator FEV₁/FVC ratio of less than 70% in spirometry testing, which measures forced expiratory flow rates, is a mandatory criterion for the diagnosis of COPD. In this study, the evaluation of spirometry reports for the diagnosis of COPD was based on a FEV₁/FVC ratio of less than 70%. The 2022 ATS/ERS guideline emphasizes the importance of PFTs for classifying physiological impairments rather than solely for establishing clinical diagnoses.⁵

Spirometry is the most reproducible and objective test for confirming airway obstruction and diagnosing COPD. It is non-invasive, inexpensive, relatively simple to perform, and widely accessible.

However, the number of patients diagnosed with COPD is significantly lower than the estimated prevalence. One of the main reasons for this discrepancy is the inadequate performance of spirometry tests, including the use of improper techniques and poor test quality.

The rationale for selecting tertiary care centers, emphasizing that these PFT labs meet international standards and have experienced technicians, thus highlighting the significance of finding quality issues even in these high-standard settings.

During patient follow-ups, it is crucial to distinguish whether a change in lung function over time reflects a true respiratory variation or merely test variability. Therefore, the reproducibility of parameters monitored during follow-up should be periodically verified using biological controls to exclude laboratory-induced variability. Variations originating from laboratory conditions can largely be eliminated by adhering to national and international consensus guidelines regarding equipment, personnel, and laboratory standards.⁹

In this study, spirometry tests performed in tertiary healthcare facilities across different cities were analyzed. The advanced mean age and higher proportion of males among the cases align with COPD's typical demographic characteristics. The slightly lower smoking rate among women might explain this male predominance. However, it is essential to note that women often have higher rates of biomass exposure and increased susceptibility to smoking and air pollutants. Nevertheless, significant barriers to the effective treatment of women with COPD include a higher rate of underdiagnosis compared to men, fewer spirometry tests performed, and less frequent medical consultations.^{11,12}

The presence of airflow obstruction in spirometry is essential for the diagnosis of COPD. In this study, the interpretation of some spirometry results as restrictive or mixed ventilatory impairment emerged as a point of debate. In such cases, factors such as inadequate effort, obesity, cardiovascular system diseases, newly developed pleural effusion, neuromuscular defects or other comorbidities may contribute to the restrictive pattern. Additionally, COPD patients with a prior history of obstructive spirometry who later demonstrate normal spirometric findings may exhibit reversible airflow limitation. In cases where there is a significant improvement in symptoms and airflow limitation at postbronchodilator spirometry, the COPD diagnosis should be reconsidered. It should be noted that different spirometry reference norms can yield varying obstruction rates within the same population, and evaluations based on the fixed ratio criterion in the GOLD guidelines may lead to misdiagnosis in more than 20% of cases.¹³

A similar finding was reported in the Sator L. et al study, where a substantial proportion of individuals with a prior COPD diagnosis did not exhibit obstructive spirometry. Among 16,177 participants, 919 (5.7%) had a self-reported COPD diagnosis; however, post-bronchodilator spirometry was unobstructed in 569 (61.9%), indicating a high rate of false-positive diagnoses. Similarly, when using the fixed FEV₁/FVC ratio criterion, overdiagnosis was observed in 55.3% of cases. These findings align with our study, suggesting that a considerable number of patients diagnosed with COPD may not have persistent airflow obstruction upon objective testing.¹⁴

Pulmonary function test reports are required to clearly indicate the reference equations utilized for each parameter.¹⁵ When interpreting results, it is important to exercise caution if different reference equations or a combination of reference equations are applied, as discrepancies can occur due to variations in the populations from which these equations were derived. Expiratory airflow is typically evaluated using spirometry, with key indices including FEV₁, FVC, and the FEV₁/FVC ratio. In healthy lungs, airflow is influenced by the strength of expiratory driving pressure (generated by expiratory muscles and elastic recoil) as well as the size and viscoelastic properties of the lungs and airways.

The criterion proposed by the Global Initiative for Chronic Obstructive Lung Disease (GOLD)¹ and the ATS/ERS⁵ COPD guidelines, which defines obstructive ventilatory defects based on a fixed FEV₁/FVC ratio of 70%, was adopted in this study. Alternatively, an obstructive ventilatory defect may also be diagnosed when the FEV₁/FVC ratio falls below the lower limit of normal (LLN), corresponding to the 5th percentile of a healthy population as recommended by the 1991 ATS⁶ and 2005 ATS/ERS¹⁶ guidelines.

In this study, it was determined that spirometry reports predominantly used ECSC as the reference value, while a portion utilized GLI in combination. Although the 2005 ATS/ERS interpretation document does not provide specific recommendations,¹⁶ the European Community for Steel and Coal (ECSC) equations have been predominantly used as

reference values. Differences exist between the predicted values of ECSC and GLI, with GLI consistently providing higher predicted values compared to ECSC.^{14–17}

The 2005 ATS/ERS interpretation strategy¹⁶ recommends the use of National Health and Nutrition Examination Survey (NHANES) III spirometry reference equations for individuals in North America. While NHANES III spirometry data are included within the GLI equations, there are some notable distinctions between the two. NHANES III derived separate equations for Mexican Americans and Caucasians, whereas the GLI equations do not make this distinction. Furthermore, GLI equations cover a broader age range (3–95 years) compared to NHANES III (8–80 years). Significant differences in predicted values have been observed between the two equations, particularly for adults over 65 years of age.^{18,19}

GLI reference ranges for the FEV₁/FVC ratio appear largely unaffected by population differences, enabling more consistent interpretations across various groups.

Z-scores quantify how far a measured lung function value deviates from the predicted value, taking into account factors such as sex, age, height, and ancestry, and are expressed in standard deviation units. This approach is recommended for defining the lower limit of normal and assessing the degree of lung function impairment. Percentile values, derived from z-scores, indicate the probability that a healthy individual would have a result below a given level and show where the individual's result falls within the healthy population. While percentile values are particularly useful for interpreting results near the normal range, they are less effective for evaluating extreme values.²⁰

The high prevalence of technical errors found in our study, particularly variable or insufficient effort and early termination, necessitates a stronger discussion on their clinical consequences. According to recent literature, substandard spirometry performance is not merely a technical failing; it directly compromises diagnostic accuracy. Studies have shown that poor effort or premature cessation often leads to an underestimation of FEV₁ and overestimation of the FEV₁/FVC ratio, potentially causing both false-positive diagnoses of obstruction (when effort is variable) and, more critically, false-negative results in patients with mild COPD (when the maneuver is too short).²¹ This confirms the urgency of our finding regarding the need for standardized reporting and enhanced technical training. Furthermore, our finding that the ECSC reference equation is overwhelmingly the most used highlights a significant methodological issue. The 2012 GLI equations are now widely endorsed by the ATS/ERS as the gold standard, overcoming the age, ethnic, and height limitations of the older ECSC and NHANES III sets. Multiple comparative analyses have demonstrated that switching from ECSC to GLI often leads to a reclassification of patients, typically resulting in a higher prevalence of both obstructive and restrictive patterns, particularly in older adults.^{22,23} Since the ECSC equation may overestimate predicted values in certain demographics, its continued widespread use, as observed in our cohort, increases the risk of under-diagnosing airflow obstruction (COPD), thereby delaying crucial treatment. This emphasizes that standardizing the reference equations used in PFT labs is as critical for diagnostic accuracy as improving the quality of patient maneuver performance.

The volume-time graph and flow-volume graph are essential components of a spirometry report, and in our study, they were present in the majority of the reports. Forced vital capacity is measured following a deep inhalation to total lung capacity (TLC) and is performed through a rapid and forceful maximal exhalation, with the results displayed graphically as volume-time or flow-volume curves.

During spirometric measurement, the volume-time graph for each maneuver should be carefully examined. This graphical analysis helps determine whether the patient's effort was sufficient at the start of the test. A rapid and forceful initiation of exhalation is crucial for the accuracy of FEV₁ and other parameters derived from the expiratory curve.

The shape of the flow-volume loop provides detailed insights into the underlying pathology.²⁴ In patients with intrathoracic airway obstruction, the expiratory curve initially loses its linearity and adopts a curvilinear appearance. As the severity of obstruction increases, the peak expiratory flow (PEF) decreases, and the terminal portion of the curve becomes progressively prolonged.²⁵

Increased concavity and small airway obstruction are critical spirometric findings in COPD cases. In COPD, airflow obstruction results from a combination of small airway disease, which raises airway resistance, and parenchymal destruction (emphysema), which diminishes the lung's normal elastic recoil. The extent to which each factor contributes varies among individuals.

Inconclusive and erroneous spirometry tests are a significant challenge in pulmonary function testing, as they can impact the reliability of results and the accuracy of clinical interpretations.²⁶ These issues can arise due to various factors related to patient effort, testing technique, and equipment functionality. Regarding the inconclusive and erroneous test, variable and insufficient effort was the most frequently reported. This was followed by early termination, rapid expiration, mouth leak, cough in the first second of respiration, and mouthpiece obstruction. The presence of volume-time and flow-volume graphs in the test report is important for identifying erroneous and inappropriate tests. Analyzing the expiratory flow-volume loop can be particularly valuable for detecting upper airway obstructions. Although some error types such as mouth leak and mouthpiece obstruction showed relatively lower kappa values, this is likely due to the small number of these events. Therefore, these findings should be interpreted as limited by event frequency rather than reflecting true disagreement between raters.

Technical and procedural errors during spirometry can have important clinical implications. Insufficient expiratory effort may lead to an underestimation of FVC and FEV₁, resulting in a false-positive diagnosis of airway obstruction. Conversely, premature termination of exhalation can underestimate disease severity in patients with true obstruction. Inaccurate results may therefore lead to inappropriate treatment decisions and misclassification of disease. Ensuring optimal patient cooperation and technician training is crucial for improving test reliability and clinical interpretation.^{5,27}

Volume-time graph helps determine whether the patient completed the maneuver correctly and sustained effort throughout the expiration. Flow-volume graph provides insights into specific patterns of airway obstruction or restriction and highlights errors such as abrupt terminations or irregular flow patterns.³

The limitations of this study include a few factors that could influence the interpretation of the findings. One limitation is the use of different spirometry devices across various PFT units, which may result in variability in test results due to differences in calibration, sensitivity, or performance of the equipment.

Another limitation is the lack of detailed information about how the diagnosis of COPD was established in the study population. Without standardized criteria for diagnosis, there is a possibility of variability in the inclusion of patients.

Additionally, data on patients' comorbidities were not available, which could have provided valuable context for interpreting spirometry results and their clinical implications.

Obesity is a common factor that can influence pulmonary function results. Increased body mass index (BMI) may lead to a restrictive spirometric pattern due to mechanical limitations of the chest wall and diaphragm rather than true parenchymal restriction. Excess adipose tissue around the thoracic and abdominal regions reduces chest wall compliance and functional residual capacity, leading to proportionally decreased FEV₁ and FVC values while the FEV₁/FVC ratio remains normal or mildly increased. Several clinical and physiological studies have demonstrated this relationship, and therefore reductions in lung volumes associated with obesity should be interpreted with caution, as they do not necessarily indicate intrinsic restrictive lung disease.²⁸

Despite these limitations, the primary focus of this study was to assess and reveal the quality and state of spirometry testing, rather than to investigate clinical outcomes or diagnostic methods in detail. This objective aligns with the study's scope and highlights the importance of spirometry as a critical tool in pulmonary function assessment.

Conclusions

The reports frequently lack adequate information on preliminary diagnoses, whether the test was the first one, as well as details about the technician and the equipment. Reference values are predominantly based on ECSC values. The rates of non-compliant and erroneous tests are low due to variable effort, insufficient effort, leaks from the mouthpiece, mouthpiece closure, and rapid expiratory flow. There is a high level of agreement between interpreters in identifying erroneous tests. Obstructive tests are predominant, and in tests with obstruction, increased concavity and small airway obstruction are the most commonly observed findings. To enhance the diagnostic validity of spirometry for COPD, mandatory reporting standards must be implemented, explicitly requiring the inclusion of technician and device details, systematic flagging of technical errors, and physician awareness of clinically significant graphical patterns like increased concavity.

Abbreviations

COPD, Chronic Obstructive Pulmonary Disease (COPD); *K*, Cohen's Kappa Test; ECSC, European Coal and Steel Community; GLI, Global Lung Function Initiative; NHANES, National Health and Nutrition Examination Survey; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; FEF₂₅₋₇₅, forced mid-expiratory flow; PFT, pulmonary function tests; BMI, body mass index; FET, forced expiratory time; GOLD, Global Initiative for Chronic Obstructive Lung Disease; ATS, American Thoracic Society; ERS, European Respiratory Society; TLC, Total lung capacity.

Data Sharing Statement

We have research data to declare.

Ethics and Consent

This research was conducted in accordance with the Declaration of Helsinki (<https://www.wma.net/policies-post/wma-declaration-of-helsinki/>).

Written informed consent was obtained from all research participants prior this study.

Ethics approval: Ethics Committee approval was granted by the Gazi University (Assessment and Evaluation Ethics Sub-Working Group) Human Research Ethics Committee with no 2019-106.

Acknowledgments

This paper/The abstract of this paper was presented at the European Respiratory Congress 2024 as a poster presentation 290 with interim findings. The poster's abstract was published in "Poster Abstracts" in European Respiratory Journal 2024 64 (suppl 68): PA1455; DOI: [10.1183/13993003.congress-2024.PA1455](https://doi.org/10.1183/13993003.congress-2024.PA1455). We thank all the technicians working in the pulmonary function test laboratory.

Author Contributions

All authors made a significant contribution to the work reported, whether in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

There is no funding to report.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: 2025 report. Available from: <https://goldcopd.org/2025-gold-report/>. Accessed January 12, 2026.
2. Agarwal AK, Raja A, Brown BD. Chronic obstructive pulmonary disease. In: *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing; 2025. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK559281/>
3. Diab N, Gershon AS, Sin DD, et al. Underdiagnosis and overdiagnosis of chronic obstructive pulmonary disease. *Ame Resp Critical Care Med*. 2018;198(9):1109–P18. doi:10.1164/rccm.201804-0621CI
4. Turkish chronic diseases and risk factors prevalence study. publication No: 909. Ankara; 2013. Türkiye Kronik Hastalıklar ve Risk Faktörleri Sıklığı Çalışması. Sağlık Bakanlığı Yayın No: 909, Ankara, 2013.
5. Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Euro Resp J*. 2022;60(1):2101499. doi:10.1183/13993003.01499-2021
6. Miller A. Lung function testing: selection of reference values and interpretative strategies. American thoracic society. *Am Rev Respir Dis*. 1991;144(5):1202–1218. doi:10.1164/ajrccm/144.5.1202
7. Pellegrino R, Viegi G, Brusasco, et al. ATS/ERS task force: interpretative strategies for lung function tests. *Eur Respir J*. 2005;26:948–968.
8. Enright PL. How to make sure your spirometry tests are of good quality. *Respir Care*. 2003;48(8):773–776.

9. Johns DP, Pierce RJ. Spirometry: the measurement and interpretation of ventilatory function in clinical practice. *MJA*. 2018;208(1):27–33.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi:10.2307/2529310
11. Zysman M, Raheison-Semjen C. Women's COPD. *Front Med*. 2022;8:600107. doi:10.3389/fmed.2021.600107
12. Zhang Y, Wang L, Mutlu GM, Cai H. More to explore: further definition of risk factors for COPD - differential gender difference, modest elevation in PM_{2.5}, and e-cigarette use. *Front Physiol*. 2021;5(12):669152. doi:10.3389/fphys.2021.669152
13. Quanjer PH, Brazzale DJ, Boros PW, Pretto JJ. Implications of adopting the global lungs initiative 2012 all-age reference equations for spirometry. *Eur Respir J*. 2013;42(4):1046–1054. doi:10.1183/09031936.00195512
14. Sator L, Horner A, Studnicka M, et al. Overdiagnosis of COPD in subjects with unobstructed spirometry: a BOLD analysis. *Chest*. 2019;156(2):277–288. doi:10.1016/j.chest.2019.01.015
15. Culver BH, Graham BL, Coates AL, et al. Recommendations for a standardized pulmonary function report. An official American thoracic society technical statement. *Am J Respir Crit Care Med*. 2017;196(11):1463–1472. doi:10.1164/rccm.201710-1981ST
16. Miller MR, Hankinson J, Brusasco V, et al. ATS/ERS Task force. Standardisation of spirometry. *Eur Respir J*. 2005;26(2):319–338. doi:10.1183/09031936.05.00034805
17. Huprikar NA, Holley AB, Skabelund AJ, et al. A comparison of global lung initiative 2012 with third national health and nutrition examination survey spirometry reference values. Implications in defining obstruction. *Ann Am Thorac Soc*. 2019;16(2):225–230. doi:10.1513/AnnalsATS.201805-317OC
18. Miller MR, Thinggaard M, Christensen K, Pedersen OF, Sigsgaard T. Best lung function equations for the very elderly selected by survival analysis. *Eur Respir J*. 2014;43(5):1338–1346. doi:10.1183/09031936.00100313
19. Linares-Perdomo O, Hegewald M, Collingridge DS, et al. Comparison of NHANES III and ERS/GLI 12 for airway obstruction classification and severity. *Eur Respir J*. 2016;48(1):133–141. doi:10.1183/13993003.01711-2015
20. Hulo S, de Broucker V, Giovannelli J, et al. Global lung function initiative reference equations better describe a middle-aged, healthy French population than the European community for steel and coal values. *Eur Respir J*. 2016;48(6):1779–1781. doi:10.1183/13993003.00606-2016
21. Yawn BP, Mintz ML, Doherty DE. GOLD in practice: chronic obstructive pulmonary disease treatment and management in the primary care setting. *Int J Chron Obstruct Pulmon Dis*. 2021;12(16):289–299. doi:10.2147/COPD.S222664
22. Belo J, Palmeiro T, Caires I, et al. Reference values for spirometry in elderly individuals: a cross-sectional study of different reference equations. *Multidiscip Respir Med*. 2018;13(1):4. doi:10.1186/s40248-017-0112-5
23. Zhang J, Hu X, Tian X, Xu KF. Global lung function initiative 2012 reference values for spirometry in Asian Americans. *BMC Pulm Med*. 2018;18(1):95. doi:10.1186/s12890-018-0658-9
24. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40(6):1324–1343. doi:10.1183/09031936.00080312
25. Gold WM, Koth LL. Pulmonary function testing. Murray and Nadel's textbook of respiratory medicine. 2016:407–435.e18. doi:10.1016/B978-1-4557-3383-5.00025-7
26. Ulubay G, Görek Dilektaşlı A, Börekçi Ş, et al. Turkish thoracic society consensus report: interpretation of spirometry. *Turk Thorac J*. 2019;20(1):69–89. doi:10.5152/TurkThoracJ.2018.180175
27. Miller MR, Hankinson JL, Brusasco V, et al. Standardization of spirometry quality assurance: common errors and their impact. *Am J Respir Crit Care Med*. 2019;199(7):e41–e52. doi:10.1164/rccm.201806-1170S
28. Jones RL, Nzekwu MMU. The effects of body mass index on lung volumes. *Chest*. 2006;130(3):827–833. doi:10.1378/chest.130.3.827

International Journal of Chronic Obstructive Pulmonary Disease

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. 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.

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

Dovepress
Taylor & Francis Group