Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Four-quadrant quantitative inhalation assessment and one-year changes in COPD Assessment Test scores in newly diagnosed COPD A descriptive longitudinal study

Yuqi Shen
Xiaoshan Affiliated Hospital of Wenzhou Medical University(The First People's Hospital of Xiaoshan District

Jing Wang
Xiaoshan Affiliated Hospital of Wenzhou Medical University(The First People's Hospital of Xiaoshan District

Haiyan Wu
Xiaoshan Affiliated Hospital of Wenzhou Medical University(The First People's Hospital of Xiaoshan District

Yanwen Zheng
johnzheng2009@163.com

Xiaoshan Affiliated Hospital of Wenzhou Medical University(The First People's Hospital of Xiaoshan District

Research Article

Keywords: Chronic obstructive pulmonary disease, COPD Assessment Test, Quantitative inhalation assessment, Peak inspiratory flow, Breath-holding, Longitudinal study

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11220014/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Inhalation performance includes device-dependent inspiratory flow, post-inhalation breath-holding and actuation–inhalation coordination. We characterised baseline performance using a four-quadrant classification and examined the association between baseline classification and one-year change in COPD Assessment Test (CAT) score in patients with newly diagnosed chronic obstructive pulmonary disease (COPD).

Methods

This single-centre descriptive longitudinal study included 484 patients with newly diagnosed, stable COPD who completed one year of regular inhaled therapy without an acute exacerbation. Baseline peak inspiratory flow (PIF) and breath-hold duration were graded from A (highest assessment priority) to D (lowest). Hand–mouth coordination contributed to the classification only when the recorded impedance grade was 0. Each patient was assigned the highest-priority grade among the contributing components. PIF thresholds were study-defined and device-specific. All patients received routine verbal inhaler education. The primary outcome was one-year change in CAT score (ΔCAT); a reduction of at least 2 points defined a CAT responder. Associations between baseline quadrant and CAT outcomes were examined using the Kruskal–Wallis test, chi-square test and baseline-adjusted analysis of covariance.

Results

PIF below the study-defined threshold occurred in 69/158 dry-powder inhaler users (43.7%), 16/91 soft-mist inhaler users (17.6%) and 13/235 pressurised metered-dose inhaler users (5.5%; $P < 0.001$). Composite quadrants A–D included 259, 84, 58 and 83 patients, respectively. Breath-hold duration was < 3 s in 254 of 259 patients in quadrant A. Reported CAT scores were 11.1 (interquartile range [IQR], 7.0–14.0) at baseline and 7.9 (IQR, 5.0–10.0) at one year; the reported ΔCAT was − 3.2 (IQR, − 6.0–0.0) points. Neither ΔCAT nor responder proportions differed significantly across quadrants ($P = 0.711$ and $P = 0.909$, respectively). The association between quadrant and CAT change remained non-significant after adjustment for baseline CAT ($P = 0.487$).

Conclusions

Short post-inhalation breath-hold duration was the most frequent component assigned the highest assessment priority. The proportion below the study-defined PIF threshold differed by inhaler class, but this comparison does not establish an independent effect of device resistance. No statistically significant association between baseline quadrant and one-year CAT change was detected in this

selected cohort. Repeated performance assessments and external validation are needed before the classification can guide prognosis or intervention.

Background

Inhaled bronchodilators are central to the pharmacological management of chronic obstructive pulmonary disease (COPD), with inhaled corticosteroids used in selected patients. Treatment effectiveness depends partly on appropriate device selection and correct inhaler use [1]. Inhaler errors remain common in clinical practice [2]. In the PIFotal study, specific inhalation errors were associated with higher COPD-related healthcare utilisation and costs [3]. However, a COPD-specific systematic review found that evidence linking individual errors to clinical outcomes was limited by study quality and inadequate assessment of adherence [4]. These limitations support examining individual components of inhalation performance in relation to clearly defined clinical outcomes.

Inhalation performance comprises several related components, including peak inspiratory flow (PIF), inhaled volume, post-inhalation breath-holding and actuation–inhalation coordination. In dry-powder inhalers (DPIs), adequate inspiratory flow is needed for effective drug delivery, and PIF measurements must be interpreted in relation to the device resistance and the assessment procedure [5]. Pressurised metered-dose inhalers (pMDIs) require a slow inhalation coordinated with actuation, whereas DPIs generally require a more forceful inhalation [6]. Assessment therefore needs to distinguish device-specific performance requirements from a general ability to generate inspiratory flow.

PIF measured against the resistance of the prescribed device can identify patients whose inspiratory manoeuvre does not meet an operational flow threshold [7]. Nevertheless, PIF alone does not capture all determinants of effective inhaler use. Device selection also requires consideration of cognitive function, manual dexterity and the ability to perform the complete inhalation sequence [8]. A structured profile combining flow, breath-holding and coordination may help identify components requiring further assessment; its association with subsequent health status requires empirical evaluation.

We applied a four-quadrant classification comprising device-specific PIF thresholds, post-inhalation breath-hold duration and hand–mouth coordination when the recorded inhaler impedance grade was 0. Each contributing component was graded from A (highest assessment priority) to D (lowest), and the highest-priority grade determined the patient's composite classification. We aimed to describe the distribution of baseline component and composite grades, compare attainment of the study-defined PIF thresholds across inhaler classes, and examine the association between baseline quadrant and one-year change in COPD Assessment Test (CAT) score. The classification was evaluated as a descriptive framework; its prognostic validity has not been established.

Methods

Study design and participants

This single-centre descriptive longitudinal study used routinely collected clinical data from Xiaoshan District First People's Hospital, Hangzhou, China. Patients with newly diagnosed, stable COPD receiving inpatient or outpatient care between June 2022 and August 2026 were considered for inclusion. The study cohort comprised 484 patients who maintained regular inhaled therapy throughout one year of follow-up without an acute exacerbation and had complete baseline and one-year CAT scores. All patients received a single session of verbal inhaler education at baseline as part of routine care; no additional structured study intervention was provided. There was no untreated comparison group. Accordingly, changes in CAT score describe the observed course of this selected cohort and cannot be attributed specifically to the education session. Reporting was guided by the STROBE statement [9]. The institutional review board waived written informed consent for analysis of de-identified routine-care data, and the study was conducted in accordance with the Declaration of Helsinki.

Quantitative inhalation assessment

Inhalation parameters and respiratory muscle pressures were measured once, at baseline, using the PF810 multifunctional pulmonary function measurement system (Zhejiang E-LinkCare Medical Technology Co., Ltd., China).DPI measurements were performed against simulated resistance corresponding to the recorded device (0.02–0.05 $kPa^{0.5}$/(L/min)); measurements in pMDI and soft-mist inhaler (SMI) users used near-zero resistance (< 0.01 $kPa^{0.5}$/(L/min)). Recorded parameters included PIF (L/min), inspiratory volume, effective inspiratory volume, inspiratory time, mean effective inspiratory flow, post-inhalation breath-hold duration (s) and hand–mouth coordination score. Maximal inspiratory pressure (MIP) and maximal expiratory pressure (MEP) were recorded as indices of respiratory muscle function; these effort-dependent pressure measurements assess a different physiological property from inspiratory flow measured against device resistance [10].

Composite four-quadrant classification

The term four-quadrant classification refers to four ordered categories, A–D, rather than four regions defined by two continuous axes. Each contributing component was graded separately, and the highest-priority component grade determined the composite category. Grade A denotes the highest assessment priority and grade D the lowest; these categories do not represent validated levels of clinical risk.

Peak inspiratory flow. The study-defined reference thresholds (T) were 30 L/min for pMDI and SMI, 50 L/min for Breezhaler, and 60 L/min for Ellipta, Diskus, Turbuhaler and Accuhaler. PIF < 0.5625T was graded A; 0.5625T ≤ PIF < 0.75T, B; 0.75T ≤ PIF <T, C; and PIF ≥ T, D. For T = 30, 50 and 60 L/min, the respective boundaries between A/B, B/C and C/D were 16.875/22.5/30, 28.125/37.5/50 and 33.75/45/60 L/min. These thresholds define the study classification and should not be interpreted as universally validated requirements for effective drug delivery. In particular, assigning grade D to all pMDI/SMI flows ≥ 30 L/min does not identify excessively rapid inhalation [6].

Post-inhalation breath-hold duration. Duration was classified as A for < 3 s, B for 3 to < 6 s, C for 6 to < 10 s, and D for ≥ 10 s. These intervals were study-defined classification boundaries.

Hand–mouth coordination. Coordination contributed to the composite classification only for patients with a recorded inhaler impedance grade of 0. This subgroup comprised 235 pMDI users, of whom 234 had an available coordination assessment. The recorded categories were missing (0), fair (1), good (2), excellent (3), poor (4) and very poor (5). Poor and very poor coordination were assigned grade A; fair, grade B; good, grade C; and excellent, grade D. Code 0 denoted missing data and was not a performance grade.

Composite grade. Component grades were ranked in the priority order A, B, C and D. Each patient received one composite grade corresponding to the highest-priority contributing component. When multiple components shared the highest-priority grade, each contributed to the same composite assignment.

Outcome

Health status was assessed using the CAT, an eight-item instrument scored from 0 to 40, with higher scores indicating greater impact of COPD [11]. The primary outcome was one-year change in CAT score, calculated as ΔCAT = CAT at one year − CAT at baseline; negative values indicate improvement. A reduction of at least 2 points (ΔCAT ≤ − 2) defined a clinically meaningful improvement in CAT score (CAT response), based on the established minimal clinically important difference [12]. Inhalation performance was not reassessed during follow-up.

Statistical analysis

Analyses used available observations for each variable, with invalid measurements treated as missing. Continuous variables were summarised using measures of central tendency and dispersion, and categorical variables as counts and percentages. Baseline characteristics were compared across composite quadrants using the Kruskal–Wallis test for continuous variables and the chi-square or Fisher exact test for categorical variables. Distributions of components used to define the composite quadrant were treated as descriptive, because their relationship with quadrant assignment is inherent in the classification rule.

The association between composite quadrant and ΔCAT was assessed using the Kruskal–Wallis test, one-way analysis of variance (ANOVA) and a non-parametric trend test across ordered quadrants. CAT responder proportions were compared using chi-square and Cochran–Armitage trend tests. Analysis of covariance assessed the quadrant association after adjustment for baseline CAT. Eight supplementary analyses were reported: (1) correlation of baseline CAT with ΔCAT; (2) comparison of baseline-adjusted ΔCAT residuals across quadrants; (3–4) separate analyses in patients with baseline CAT ≥ 10 and < 10; (5) percentage change in CAT; (6) the proportion with worsening CAT (ΔCAT ≥ 2); (7) cross-tabulation of

quadrant and baseline CAT stratum; and (8) associations of quadrant with age and baseline CAT. The correlation between baseline CAT and CAT change was interpreted descriptively and was not used to establish regression to the mean. Attainment of the study-defined PIF threshold was compared across inhaler classes using the chi-square test. All tests were two-sided, with $P < 0.05$ considered statistically significant. Analyses were performed in R version 4.4.0.

Results

Participant characteristics and quadrant distribution

The study cohort comprised 484 patients with newly diagnosed, stable COPD (Fig. 1). Median age was 68.6 years (IQR, 63.0–75.0), 64 patients (13.2%) were female, and median BMI was 23.9 kg/m² (IQR, 21.6–25.2). The recorded inhaler class was pMDI in 235 patients (48.6%), SMI in 91 (18.8%) and DPI in 158 (32.6%). Composite quadrants A–D included 259 (53.5%), 84 (17.4%), 58 (12.0%) and 83 (17.1%) patients, respectively (Fig. 2). Characteristics and CAT outcomes are presented in Table 1. Between-quadrant comparisons identified statistically significant differences in age, sex distribution, BMI, inhaler class and MIP. Median age was highest in quadrant A (70.1 years) and lowest in quadrant D (65.1 years; overall $P = 0.001$).

Table 1
Characteristics and CAT outcomes by baseline composite quadrant

Characteristic	All patients (n = 484)	Quadrant A (n = 259)	Quadrant B (n = 84)	Quadrant C (n = 58)	Quadrant D (n = 83)	P value
Age, years	68.6 (63.0–75.0)	70.1 (65.0–75.0)	68.7 (64.0–75.2)	67.2 (60.2–75.0)	65.1 (60.0–72.0)	0.001
Female, n (%)	64 (13.2%)	37 (14.3%)	13 (15.5%)	11 (19.0%)	3 (3.6%)	0.030
BMI, kg/m²	23.9 (21.6–25.2)	24.4 (22.2–26.0)	22.7 (20.7–24.4)	24.7 (21.1–25.2)	23.1 (21.1–24.6)	< 0.001
MIP, cmH_2O	44.8 (28.7–59.1)	41.3 (25.2–54.0)	45.3 (30.7–58.7)	46.9 (30.8–59.3)	53.5 (38.9–66.7)	< 0.001
MEP, cmH_2O	52.5 (37.3–65.3)	50.0 (35.3–62.9)	55.7 (37.5–66.3)	51.8 (39.8–64.0)	57.4 (40.8–72.1)	0.057
Inhaler device class, n (%)						< 0.001
pMDI	235 (48.6%)	154 (59.5%)	45 (53.6%)	28 (48.3%)	8 (9.6%)	
SMI	91 (18.8%)	28 (10.8%)	12 (14.3%)	11 (19.0%)	40 (48.2%)	
DPI	158 (32.6%)	77 (29.7%)	27 (32.1%)	19 (32.8%)	35 (42.2%)	
Recorded impedance grade, n (%)						< 0.001
Grade 0 (none)	235 (48.6%)	154 (59.5%)	45 (53.6%)	28 (48.3%)	8 (9.6%)	
Grade 1	72 (14.9%)	44 (17.0%)	4 (4.8%)	3 (5.2%)	21 (25.3%)	
Grade 2	58 (12.0%)	26 (10.0%)	11 (13.1%)	11 (19.0%)	10 (12.0%)	
Grade 3	26 (5.4%)	7 (2.7%)	12 (14.3%)	3 (5.2%)	4 (4.8%)	
Grade 4	93 (19.2%)	28 (10.8%)	12 (14.3%)	13 (22.4%)	40 (48.2%)	
Recorded coordination, n (%)						< 0.01
Missing (0)	3 (0.6%)	0 (0.0%)	1 (1.2%)	1 (1.7%)	1 (1.2%)	

Characteristic	All patients (n = 484)	Quadrant A (n = 259)	Quadrant B (n = 84)	Quadrant C (n = 58)	Quadrant D (n = 83)	P value
Fair (1)	219 (45.2%)	143 (55.2%)	49 (58.3%)	9 (15.5%)	18 (21.7%)	
Good (2)	161 (33.3%)	84 (32.4%)	15 (17.9%)	37 (63.8%)	25 (30.1%)	
Excellent (3)	84 (17.4%)	16 (6.2%)	18 (21.4%)	11 (19.0%)	39 (47.0%)	
Poor (4)	13 (2.7%)	13 (5.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Very poor (5)	4 (0.8%)	3 (1.2%)	1 (1.2%)	0 (0.0%)	0 (0.0%)	
Baseline CAT	11.1 (7.0–14.0)	11.6 (7.0–15.0)	11.2 (7.8–14.0)	10.6 (7.0–12.8)	9.9 (6.0–12.0)	0.050
1-year CAT	7.9 (5.0–10.0)	8.2 (5.0–10.5)	7.9 (4.0–11.0)	7.9 (5.0–9.0)	7.1 (4.0–9.0)	0.276
CAT change (1 year − baseline)	−3.2 (−6.0–0.0)	−3.4 (−6.0–0.0)	−3.3 (−6.0–0.0)	−2.7 (−6.0–0.0)	−2.8 (−5.0–0.0)	0.711
CAT responder (ΔCAT ≤ − 2), n (%)	293 (60.5%)	159 (61.4%)	52 (61.9%)	33 (56.9%)	49 (59.0%)	0.909

Values are median (interquartile range) unless otherwise stated. Categorical variables are presented as n (%) of the column total. BMI, body mass index; CAT, COPD Assessment Test; DPI, dry-powder inhaler; MEP, maximal expiratory pressure; MIP, maximal inspiratory pressure; pMDI, pressurised metered-dose inhaler; SMI, soft-mist inhaler. P values compare the four quadrants using the Kruskal–Wallis test for continuous variables and the chi-square test for categorical variables. The hand–mouth coordination rows describe recorded scores in the entire cohort; coordination contributed to the composite grade only for patients with recorded impedance grade 0. A between-quadrant P value is not reported for coordination because it contributes to the classification in the impedance-grade-0 subgroup. Quadrant A denotes the highest assessment priority and quadrant D the lowest.

Distribution of component grades

The component-grade distributions showed distinct descriptive patterns (Fig. 2 and Table 2). PIF was at or above the study-defined threshold (grade D) in 386/484 patients (79.8%), whereas 6 (1.2%) received grade A. Breath-hold duration was < 3 s (grade A) in 254 patients (52.5%) and ≥ 10 s (grade D) in 179 (37.0%). Among 234 patients with impedance grade 0 and a non-missing coordination score, coordination was most often graded B (n = 128, 54.7%) or C (n = 76, 32.5%).

Short breath-hold duration was the most frequent grade-A component among patients in quadrant A: 254/259 (98.1%) had a breath-hold duration < 3 s. PIF grades matched the composite grade in 6 patients

in quadrant A, 16 in quadrant B and 27 in quadrant C. Coordination grades matched the composite grade in 33 patients in quadrant B and 22 in quadrant C. These counts are not mutually exclusive because multiple components can share the highest-priority grade.

Table 2
Component grade distributions in the four-quadrant assessment

Assessment component	Grade A, n (%)	Grade B, n (%)	Grade C, n (%)	Grade D, n (%)	Gradable, n
Peak inspiratory flow	6 (1.2%)	29 (6.0%)	63 (13.0%)	386 (79.8%)	484
Breath-hold duration	254 (52.5%)	38 (7.9%)	13 (2.7%)	179 (37.0%)	484
Hand-mouth coordination	12 (5.1%)	128 (54.7%)	76 (32.5%)	18 (7.7%)	234

Percentages were calculated using the number of patients with an available assessment for the corresponding component as the denominator. Each component was graded independently from A (highest assessment priority) to D (lowest). Coordination contributed only for the 235 patients with recorded impedance grade 0; one assessment was missing, leaving 234 gradable. Percentages may not sum to 100 because of rounding.

Inhaler class and attainment of the study-defined PIF threshold

The proportion below the study-defined PIF threshold differed across inhaler classes (Fig. 3 and Table 3): 13/235 pMDI users (5.5%), 16/91 SMI users (17.6%) and 69/158 DPI users (43.7%; $P < 0.001$). Proportions within the recorded device categories were 13/235 for pMDI (5.5%), 9/72 for Breezhaler (12.5%), 16/91 for Respimat (17.6%), 35/53 for Ellipta (66.0%), 19/26 for Diskus (73.1%), 4/5 for Turbuhaler (80.0%) and 2/2 for Accuhaler (100%). Median PIF was 54.0, 43.4 and 58.0 L/min in pMDI, SMI and DPI users, respectively, against thresholds of 30, 30 and 50–60 L/min. Median MIP was 42.6, 51.2 and 43.8 cmH_2O in these groups; the overall three-group comparison was not statistically significant (Kruskal–Wallis $P = 0.064$). Because inhaler class, measurement resistance and classification threshold were not varied independently, these comparisons cannot isolate an effect of resistance or exclude a contribution from inspiratory muscle function.

Table 3
Attainment of study-defined PIF thresholds by inhaler class and recorded device

Device or class	n	Study PIF threshold, L/min	Median PIF (IQR), L/min	Below threshold, n (%)
All patients	484	—	53.7 (41.7–68.3)	98 (20.2%)
By inhaler class				
pMDI	235	30	54.0 (42.2–73.3)	13 (5.5%)
SMI	91	30	43.4 (33.6–54.2)	16 (17.6%)
DPI	158	60 or 50	58.0 (46.7–70.7)	69 (43.7%)
By individual device				
pMDI	235	30	54.0 (42.2–73.3)	13 (5.5%)
Respimat	91	30	43.4 (33.6–54.2)	16 (17.6%)
Breezhaler	72	50	65.7 (56.2–80.2)	9 (12.5%)
Ellipta	53	60	54.0 (45.7–64.9)	35 (66.0%)
Diskus	26	60	46.5 (42.6–61.9)	19 (73.1%)
Turbuhaler	5	60	44.2 (37.9–55.5)	4 (80.0%)
Accuhaler	2	60	51.8 (48.6–55.1)	2 (100.0%)

PIF, peak inspiratory flow; IQR, interquartile range; pMDI, pressurised metered-dose inhaler; SMI, soft-mist inhaler; DPI, dry-powder inhaler. Thresholds were study-defined: 30 L/min for pMDI and SMI, 50 L/min for Breezhaler, and 60 L/min for the other recorded DPI categories. Below-threshold proportions differed across inhaler classes (chi-square $P < 0.001$). Meeting these thresholds does not establish optimal drug delivery. Device-specific estimates, especially for Accuhaler (n = 2) and Turbuhaler (n = 5), are imprecise because of small samples.

One-year CAT outcomes according to baseline quadrant

As reported in Table 1, CAT scores were 11.1 (IQR, 7.0–14.0) at baseline and 7.9 (IQR, 5.0–10.0) at one year, with a reported ΔCAT of − 3.2 (IQR, − 6.0–0.0) points. At the individual level, 293 of 484 patients (60.5%) achieved a CAT reduction of at least 2 points and met the responder definition. The observed within-cohort improvement cannot be attributed specifically to verbal education.

The unadjusted comparisons did not identify a statistically significant difference in ΔCAT across composite quadrants (Table 1 and Fig. 4). The ΔCAT values reported in Table 1 were − 3.4, − 3.3, − 2.7 and − 2.8 points in quadrants A–D, respectively (Kruskal–Wallis $P = 0.711$; ANOVA $P = 0.735$; trend $P = 0.366$). Responder proportions were 61.4%, 61.9%, 56.9% and 59.0%, respectively ($P = 0.909$; trend P =

0.584). The association between quadrant and CAT change remained non-significant after adjustment for baseline CAT (P = 0.487).

Supplementary analyses

Supplementary outcome analyses did not identify a statistically significant association between quadrant and CAT change. Baseline CAT score was inversely correlated with ΔCAT (r = − 0.692). The comparison of baseline-adjusted CAT-change residuals across quadrants was non-significant (P = 0.768), as were analyses restricted to baseline CAT ≥ 10 (n = 264; P = 0.935) or < 10 (n = 220; P = 0.614), percentage change (P = 0.900) and worsening proportions (P = 0.896). The between-quadrant comparisons yielded P = 0.001 for age and P = 0.050 for baseline CAT (Table 1). Across quadrants, patients in the higher baseline CAT stratum showed larger mean reductions in CAT score (Fig. 4). These descriptive patterns do not establish a causal explanation for the observed change.

Discussion

This descriptive longitudinal study characterised baseline inhalation performance in 484 patients with newly diagnosed, stable COPD and examined the association between baseline quadrant and one-year CAT change. More than half were assigned to quadrant A, predominantly because their breath-hold duration was < 3 s. The proportion below the study-defined PIF threshold was higher among DPI users than among pMDI or SMI users. Neither CAT change nor responder proportion differed significantly across composite quadrants in the unadjusted analyses. The association between quadrant and CAT change also remained non-significant after adjustment for baseline CAT. These findings describe the performance profile of a selected cohort receiving routine inhaled therapy and education; they do not establish a causal effect of inhaler class, education or quadrant assignment.

Short post-inhalation breath-hold duration was the most frequent high-priority finding in our cohort, highlighting the value of describing individual steps of the inhalation sequence. In the PIFotal study, errors in breath-holding and inhalation were associated with higher (worse) CAT and Clinical COPD Questionnaire scores among DPI users [13]. This provides a clinical rationale for assessing breath-holding, although it does not validate our < 3 s boundary or the four-quadrant scheme. In our cohort, 254 of 259 patients in quadrant A had a breath-hold duration < 3 s, indicating that this component largely determined the highest-priority category. The concentration of observations in the < 3-s and ≥ 10-s categories could reflect several factors, including instruction, symptoms, physiological tolerance or measurement procedures. Without repeated assessment, this pattern cannot be attributed specifically to habitual behaviour, and its responsiveness to training remains uncertain.

Interpretation of the PIF findings requires separating measurement conditions from the classification threshold. In this study, PIF was measured under the resistance condition assigned to each device, and the resulting value was classified using a device-specific operational threshold. DPI measurements used simulated device resistance and thresholds of 50 or 60 L/min, whereas pMDI and SMI measurements

used near-zero resistance and the same threshold of 30 L/min. Differences in the proportions below threshold therefore cannot be attributed independently to device resistance. Moreover, the three-group MIP comparison (P = 0.064) does not establish equivalence of measured inspiratory muscle pressure or rule out confounding. Separate measurement of MIP and PIF does not establish which physiological factors explain the observed difference between inhaler classes. These distinctions are particularly relevant when interpreting small device subgroups.

The direction of the grading rule also warrants caution. Higher PIF is not uniformly preferable across inhaler types: pMDIs and SMIs require a slow, controlled inhalation, whereas effective DPI use depends on generating an appropriate manoeuvre against the device resistance [6]. The present scheme classifies every pMDI/SMI value ≥ 30 L/min as grade D and therefore cannot detect excessive flow or confirm optimal aerosol delivery. Consequently, the device comparisons describe attainment of study-defined thresholds and do not establish relative device suitability. These results alone cannot identify the most suitable device for an individual patient, because the classification captures only selected aspects of inhaler performance.

The absence of a statistically significant quadrant–CAT association should be interpreted alongside the heterogeneous evidence on inhaler errors and clinical outcomes in COPD [4]. A randomised trial by Huang and colleagues reported improvements in inhaler technique and adherence with a gamified educational programme compared with routine education in patients with COPD [14]. A video-based teach-to-goal trial also reported better technique and adherence than verbal instruction, although most participants had asthma and the COPD subgroup was small [15]. These trials evaluated structured educational interventions, whereas our study examined the association between a baseline performance classification and subsequent CAT change. Their findings therefore address a different question. In our cohort, all patients received verbal education, but technique was not reassessed. Subsequent improvement or deterioration in technique could have attenuated the association between baseline classification and one-year CAT change; this possibility remains untested.

Differences in study populations, measurement procedures and endpoints also complicate comparisons. In a hospital-based study, Mahler and colleagues found that suboptimal PIF was common and its prevalence varied across DPI resistance categories [16]. In stable outpatients, the same research group reported a shorter time to first exacerbation among patients with suboptimal PIF [17]. A PIFotal analysis linked suboptimal PIF to poorer health status but found no significant association with exacerbations in the preceding year after correction for multiple testing [18]. These findings indicate that health-status and exacerbation outcomes should be considered separately. The analysed cohort comprised patients who remained free of acute exacerbations during follow-up, and the outcome was CAT change. The study therefore cannot evaluate associations with exacerbation risk, and its findings may not generalise to patients who experience exacerbations.

The inverse correlation between baseline CAT and ΔCAT (r = − 0.692) requires a separate statistical interpretation. Because baseline CAT is included with a negative sign in the calculation of ΔCAT, the two

variables are mathematically coupled [19]. Regression to the mean, measurement variability and less opportunity for improvement among patients with low initial scores may also contribute; the correlation alone cannot quantify these effects. The interpretation of baseline-adjusted estimates depends on the temporal and causal roles of baseline variables and on measurement error [20]. Accordingly, our baseline-adjusted analysis describes an association conditional on baseline CAT and does not by itself establish a causal effect. Although the adjusted quadrant association remained non-significant, P values alone do not establish equivalence among quadrants. Effect estimates and confidence intervals are needed to determine which magnitudes of association remain compatible with the data.

Strengths include quantitative baseline assessment, explicit component-grading rules, a cohort of 484 patients and complete paired CAT measurements within the analysed sample. Reporting both component and composite-grade distributions helps identify which components contribute most frequently to the composite classification. The supplementary outcome analyses also allow the main finding to be examined under several alternative definitions and baseline CAT strata. These strengths support descriptive interpretation, but do not establish the validity of the classification or the absence of a clinically important association.

Several limitations affect interpretation. First, this was a single-centre observational analysis without an untreated comparator. Residual confounding by treatment selection, adherence, comorbidity and baseline disease characteristics remains possible, and CAT improvement cannot be attributed specifically to the education session. Restricting analysis to patients who continued regular therapy and had no acute exacerbation during follow-up may also introduce selection bias. Second, the four-quadrant thresholds have not been externally validated, and the strong influence of breath-hold duration may obscure associations involving other components. Coordination contributed to the composite grade only in the impedance-grade-0 subgroup, limiting comparability of the classification across inhaler classes. Third, a single baseline assessment cannot characterise changes in inhalation performance during follow-up. Lung deposition, changes in treatment and repeated technique measurements were not available to explain the observed CAT changes. Fourth, some device subgroups were very small, and women accounted for only 13.2% of the sample, limiting precision and generalisability. Finally, incomplete MIP and MEP data limit analyses of respiratory muscle pressures. The absence of reported between-quadrant effect estimates with confidence intervals limits interpretation of the non-significant CAT findings.

The component profile and CAT scores describe different aspects of the patient assessment: observed inhalation performance and patient-reported COPD impact, respectively. In practice, component-level findings can indicate which steps of the manoeuvre merit further review. The absence of a statistically significant quadrant–CAT association does not negate the potential value of assessing individual components. Future studies should evaluate whether repeated component assessment and structured education improve technique, adherence and relevant clinical outcomes, with clear diagnostic criteria, independently validated thresholds and inclusion of patients who experience exacerbations [4, 15].

Conclusions

In this selected cohort of patients with newly diagnosed, stable COPD, short post-inhalation breath-hold duration was the most frequent high-priority finding in the four-quadrant assessment. The proportion below the study-defined PIF threshold differed by inhaler class, but an independent effect of device resistance was not established. No statistically significant association between baseline quadrant and one-year CAT change was detected. The framework may provide a structured description of inhalation performance; evaluation of its thresholds, reproducibility and clinical relevance is required before it can be used to guide prognosis or evaluate the effectiveness of education.

Abbreviations

ANOVA
analysis of variance
BMI
body mass index
CAT
COPD Assessment Test
CI
confidence interval
COPD
chronic obstructive pulmonary disease
DPI
dry-powder inhaler
IQR
interquartile range
MCID
minimal clinically important difference
MEP
maximal expiratory pressure
MIP
maximal inspiratory pressure
NS
not statistically significant
PIF
peak inspiratory flow
pMDI
pressurised metered-dose inhaler
SD
standard deviation
SE

standard error
SMI
soft-mist inhaler
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology
ΔCAT
one-year CAT score minus baseline CAT score.

Declarations

Ethics approval and consent to participate

The Ethics Committee of The First People's Hospital of Xiaoshan District, Hangzhou, approved the study (Xiaoshan Medical Ethics Approval 2022-Science-072). The institutional review board waived written informed consent for analysis of de-identified routine-care data. The study followed the Declaration of Helsinki.

Clinical trial number: not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request, subject to approval by the Ethics Committee of the First People's Hospital of Xiaoshan District, Hangzhou.

Competing interests

The authors declare that they have no competing interests.

Funding

Zhejiang Provincial Medical and Health Science and Technology Program (2023XY121, 2023XY126 and 2025HY0797).

Authors' contributions

Yuqi Shen: Conceptualization, Data curation, Formal analysis, Methodology, Project administration.Yanwen Zheng:Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.Jing Wang: Data curation, Formal analysis, Investigation,

Methodology, Resources, Software, Validation, Visualization. Haiyan Wu: Writing - original draft, Writing - review & editing.

Acknowledgements

The authors thank the respiratory nursing and pulmonary function laboratory staff of Xiaoshan District First People's Hospital for their assistance with quantitative inhalation assessment and data collection.

References

1. Global Initiative for Chronic Obstructive Lung Disease. Global strategy for prevention, diagnosis and management of COPD: 2026 report. 2026. https://goldcopd.org/2026-gold-report-and-pocket-guide/. Accessed 28 Sep 2026.
2. Grandmaison G, Grobéty T, Vaucher J, Hayoz D, Suter P. Prevalence of critical errors and insufficient peak inspiratory flow in patients hospitalized with COPD in a department of general internal medicine: a cross-sectional study. Chronic Obstr Pulm Dis. 2024;11(4):406–15. 10.15326/jcopdf.2024.0505.
3. Leving MT, van Boven JFM, Bosnic-Anticevich SZ, van Cooten J, Correia de Sousa J, Cvetkovski B, et al. Suboptimal peak inspiratory flow and critical inhalation errors are associated with higher COPD-related healthcare costs. Int J Chron Obstruct Pulmon Dis. 2022;17:2401–15. 10.2147/COPD.S380736.
4. Halpin DMG, Mahler DA. Systematic review of the effects of patient errors using inhaled delivery systems on clinical outcomes in COPD. BMJ Open Respir Res. 2024;11(1):e002211. 10.1136/bmjresp-2023-002211.
5. Ohar JA, Ferguson GT, Mahler DA, Drummond MB, Dhand R, Pleasants RA, et al. Measuring peak inspiratory flow in patients with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis. 2022;17:79–92. 10.2147/COPD.S319511.
6. Rigby D. Inhaler device selection for people with asthma or chronic obstructive pulmonary disease. Aust Prescr. 2024;47(5):140–7. 10.18773/austprescr.2024.046.
7. Leving MT, Bosnic-Anticevich S, van Cooten J, de Sousa JC, Cvetkovski B, Dekhuijzen R, et al. Clinical recommendations for dry powder inhaler use in the management of COPD in primary care. NPJ Prim Care Respir Med. 2022;32(1):59. 10.1038/s41533-022-00318-3.
8. Mahler DA, Halpin DMG. Consideration and assessment of patient factors when selecting an inhaled delivery system in COPD. Chest. 2024;165(2):323–32. 10.1016/j.chest.2023.11.011.
9. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. 10.1371/journal.pmed.0040296.

10. Silveira BMF, Pereira HLA, Chaves G, Ho DGC, Parreira VF. Reliability and validity of maximal respiratory pressures. Respir Care. 2024;69(7):881–90. 10.4187/respcare.10641.
11. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Leidy NK. Development and first validation of the COPD Assessment Test. Eur Respir J. 2009;34(3):648–54. 10.1183/09031936.00102509.
12. Kon SS, Canavan JL, Jones SE, Nolan CM, Clark AL, Dickson MJ, et al. Minimum clinically important difference for the COPD Assessment Test: a prospective analysis. Lancet Respir Med. 2014;2(3):195–203. 10.1016/S2213-2600(14)70001-3.
13. Kocks J, Bosnic-Anticevich S, van Cooten J, Correia de Sousa J, Cvetkovski B, Dekhuijzen R, et al. Identifying critical inhalation technique errors in Dry Powder Inhaler use in patients with COPD based on the association with health status and exacerbations: findings from the multi-country cross-sectional observational PIFotal study. BMC Pulm Med. 2023;23(1):302. 10.1186/s12890-023-02566-6.
14. Huang X, Jiang Z, Dai Y, Liu Y, Dai Z, Wang J, et al. Effect of gamification on improved adherence to inhaled medications in chronic obstructive pulmonary disease: randomized controlled trial. J Med Internet Res. 2025;27:e65309. 10.2196/65309.
15. Al-Kharouf MS, Abdeljalil MH, Obeidat NM, Oweidat KA, Awwad O. Video-based teach-to-goal intervention on inhaler technique on adults with asthma and COPD: a randomized controlled trial. PLoS ONE. 2023;18(6):e0286870. 10.1371/journal.pone.0286870.
16. Mahler DA, Demirel S, Hollander R, Gopalan G, Shaikh A, Mahle CD, et al. High prevalence of suboptimal peak inspiratory flow in hospitalized patients with COPD: a real-world study. Chronic Obstr Pulm Dis. 2022;9(3):427–38. 10.15326/jcopdf.2022.0291.
17. Mahler DA, Niu X, Deering KL, Dembek C. Prospective evaluation of exacerbations associated with suboptimal peak inspiratory flow among stable outpatients with COPD. Int J Chron Obstruct Pulmon Dis. 2022;17:559–68. 10.2147/COPD.S353441.
18. Kocks JWH, Wouters H, Bosnic-Anticevich S, van Cooten J, Correia de Sousa J, Cvetkovski B, et al. Factors associated with health status and exacerbations in COPD maintenance therapy with dry powder inhalers. NPJ Prim Care Respir Med. 2022;32(1):18. 10.1038/s41533-022-00282-y.
19. Tennant PWG, Arnold KF, Ellison GTH, Gilthorpe MS. Analyses of change scores do not estimate causal effects in observational data. Int J Epidemiol. 2022;51(5):1604–15. 10.1093/ije/dyab050.
20. Glymour MM, Commentary. Modelling change in a causal framework. Int J Epidemiol. 2022;51(5):1615–21. 10.1093/ije/dyac151.

Figures

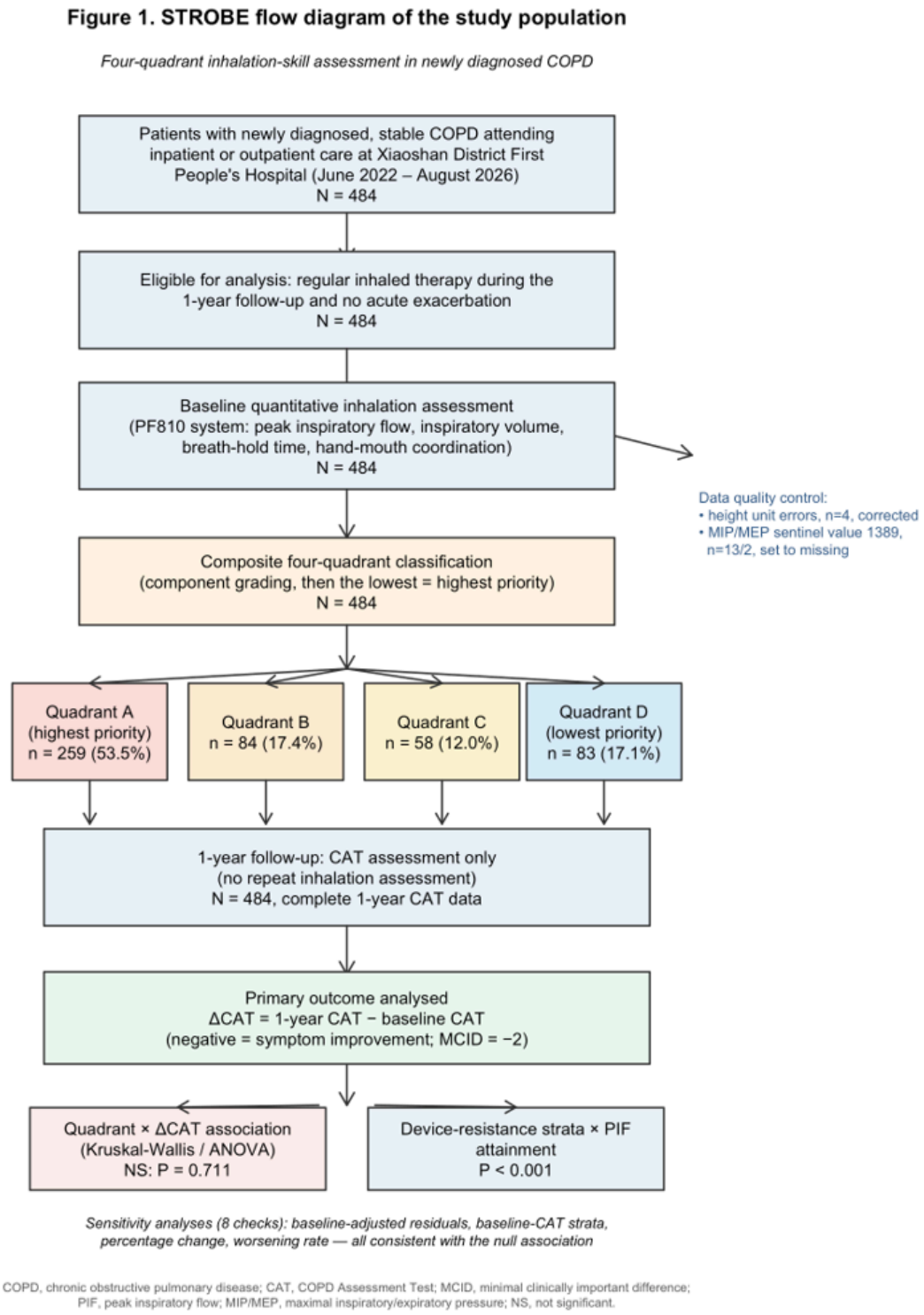


Figure 1

Overview of the study population and analyses.
The cohort comprised 484 patients with newly diagnosed, stable COPD who maintained regular inhaled therapy without an acute exacerbation during one-year follow-up. Baseline PF810 assessment assigned 259, 84, 58 and 83 patients to quadrants A–D, respectively. All had paired CAT scores. CAT, COPD

Assessment Test; MCID, minimal clinically important difference; MIP/MEP, maximal inspiratory/expiratory pressure; NS, not statistically significant; PIF, peak inspiratory flow.

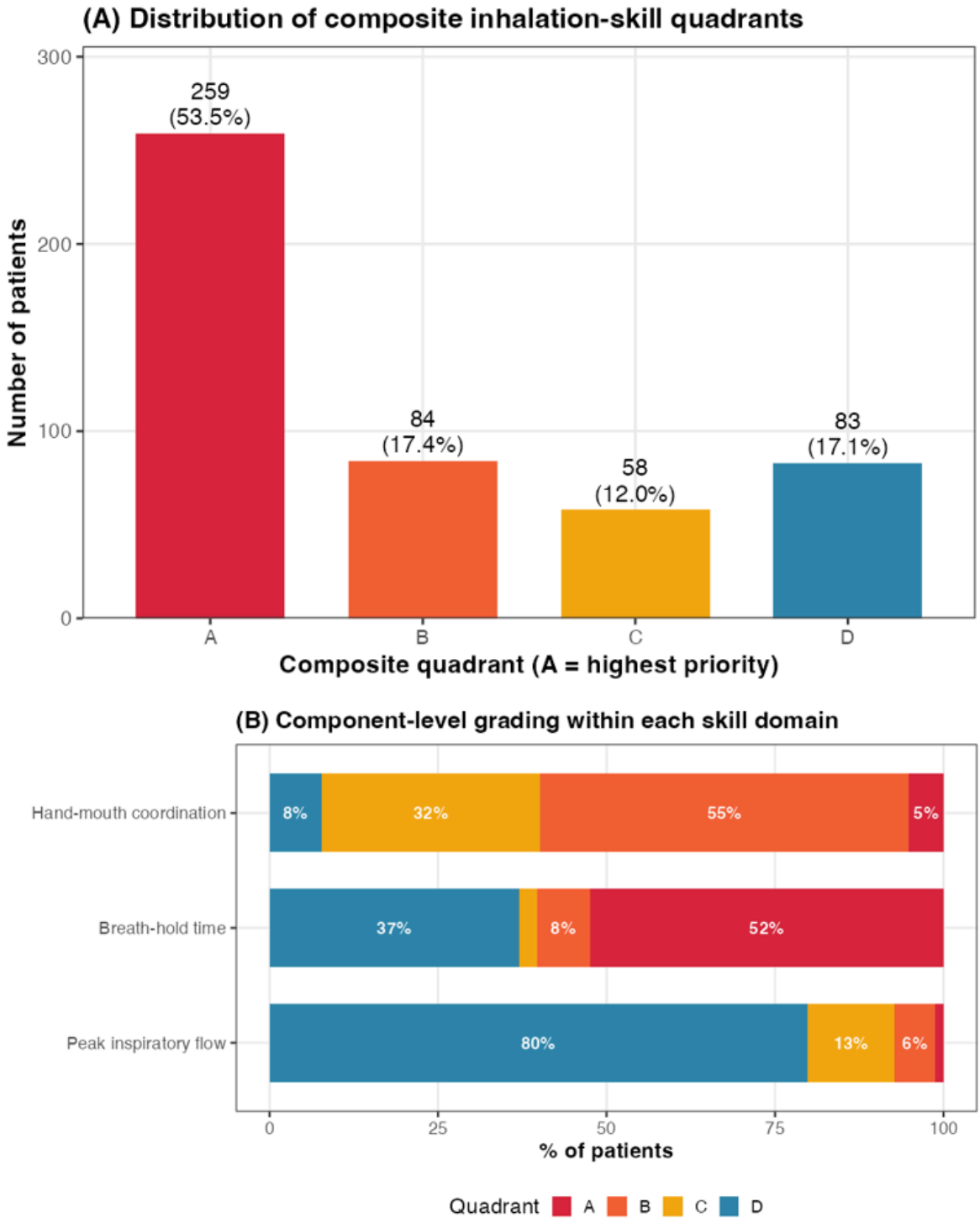


Figure 2

Composite and component grade distributions.
(A) Distribution across the four composite quadrants (N = 484). Each patient's composite grade is determined by the highest-priority contributing component grade; A denotes the highest assessment

priority. (B) Component grades expressed as percentages of gradable patients. PIF and breath-hold duration were gradable in 484 patients, and coordination in 234 patients with recorded impedance grade 0 and a non-missing assessment.

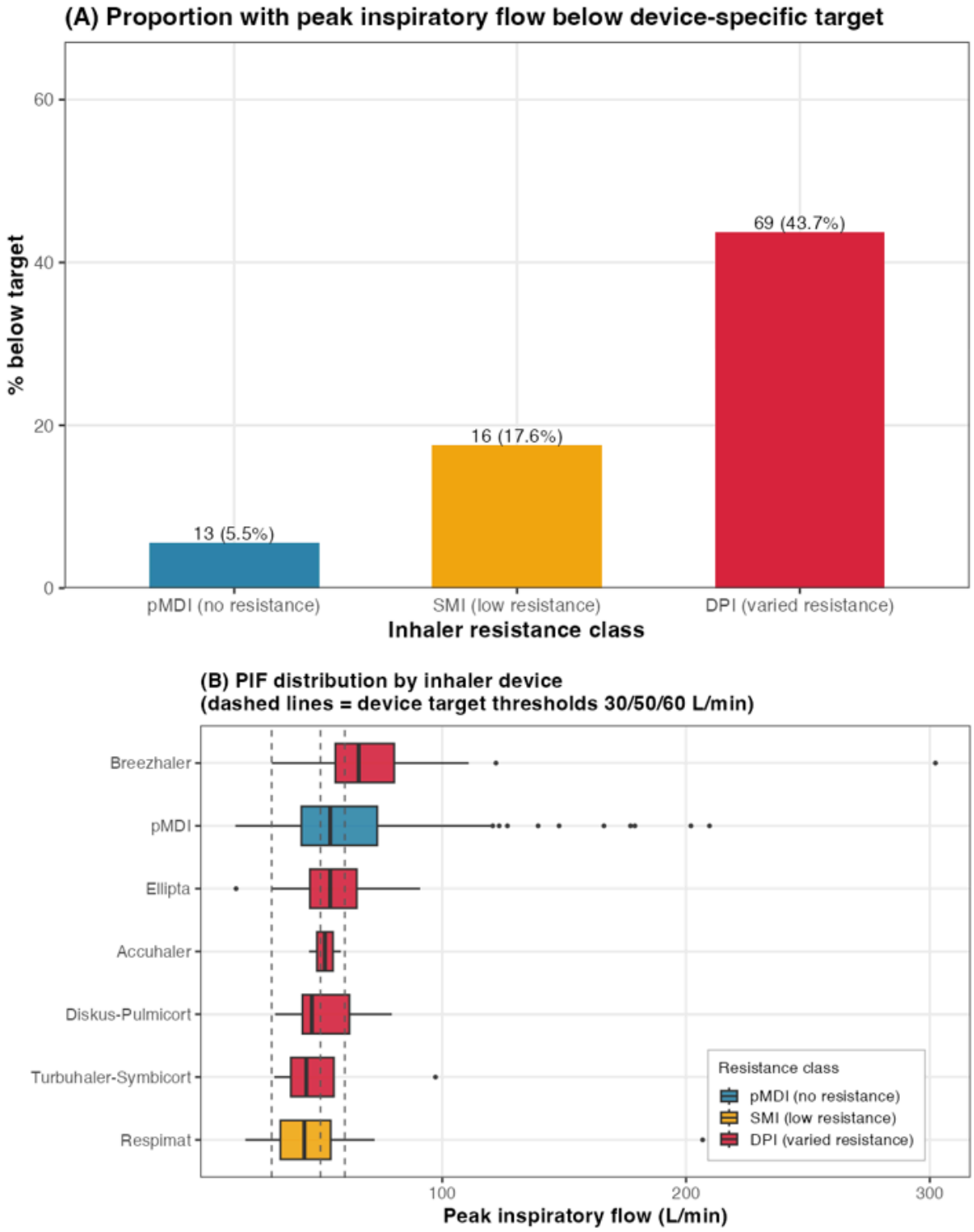


Figure 3

Attainment of study-defined PIF thresholds by inhaler class.
(A) Proportions below threshold: pMDI, 13/235 (5.5%); SMI, 16/91 (17.6%); and DPI, 69/158 (43.7%; chi-square $P < 0.001$). (B) Box-and-whisker plots of PIF by recorded device category, ordered by median; dashed lines mark thresholds of 30, 50 and 60 L/min. Boxes show the interquartile range, central lines show medians, and dots show outliers. Colours identify inhaler classes. Thresholds and measurement conditions differ across devices, so the figure does not isolate a resistance effect or establish optimal inhalation. DPI, dry-powder inhaler; pMDI, pressurised metered-dose inhaler; PIF, peak inspiratory flow; SMI, soft-mist inhaler.

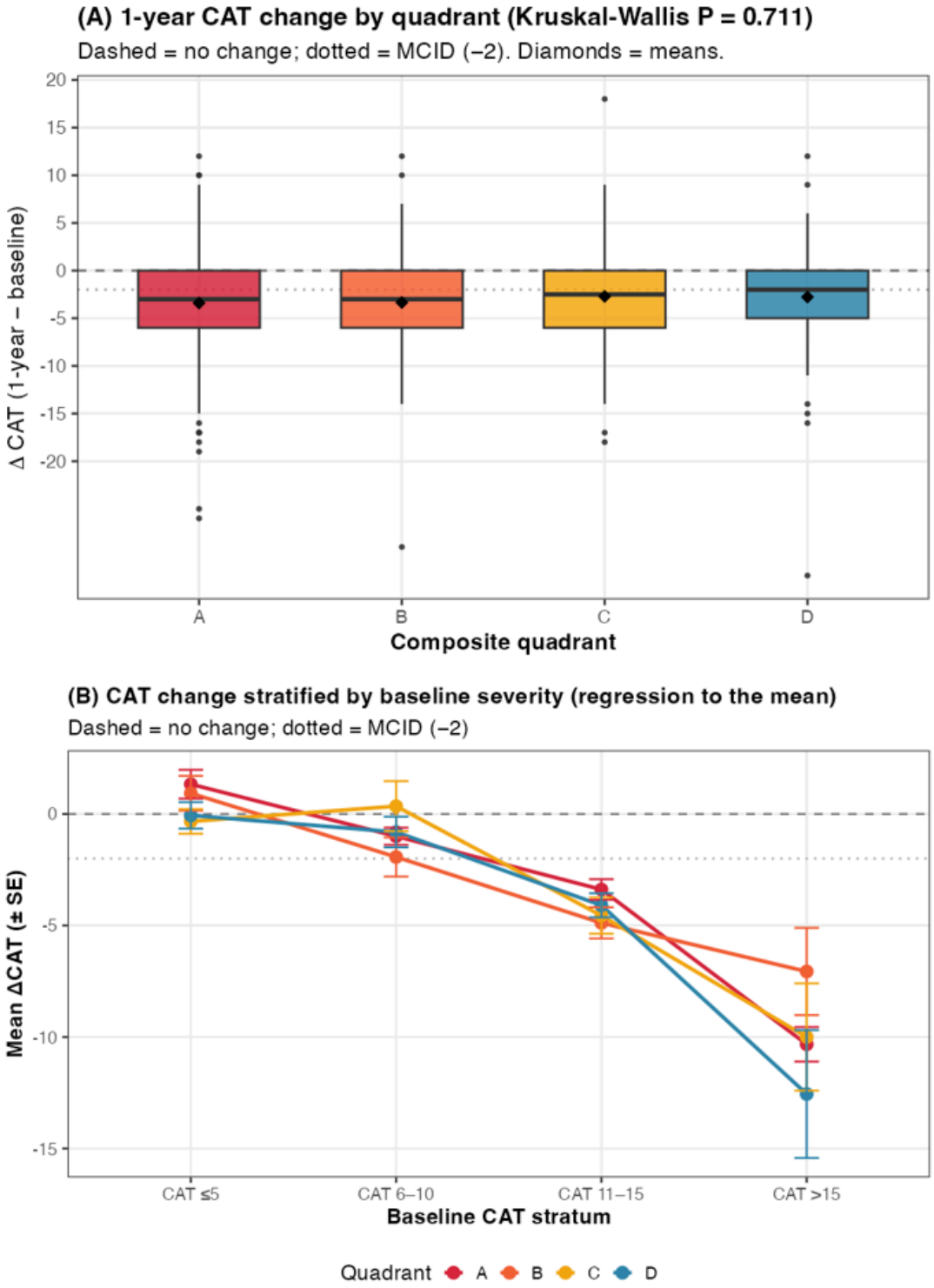


Figure 4

One-year CAT change by baseline quadrant and CAT stratum.
(A) Box-and-whisker plots of ΔCAT (one-year minus baseline CAT) by quadrant. Negative values indicate improvement. Dashed and dotted lines show no change and a 2-point reduction, respectively; diamonds show group means and dots show outliers (Kruskal–Wallis P = 0.711). (B) Mean ΔCAT with standard

errors by baseline CAT stratum and quadrant. Larger reductions accompanied higher baseline scores. This descriptive pattern does not establish a causal relationship or quantify regression to the mean.