Chronic Obstructive Pulmonary Diseases:
Journal of the COPD Foundation®

Original Research

Lung Function Variability in Alpha-1 Antitrypsin Deficiency: Implications for Clinical Trials

Charlie Strange, MD[1,2] Kristen E. Holm, PhD, MPH[2,3] Robert A. Sandhaus, MD, PhD[2,3] David M. Mannino, MD[4] Radmila Choate, PhD, MPH[2,5]

Abstract

Background: Alpha-1 antitrypsin deficiency (AATD) has considerable interindividual variability in lung disease progression. Pulmonary impairment is often measured by forced expiratory volume in 1 second (FEV_1) and diffusing capacity of the lung for carbon monoxide (DLCO). This study aimed to quantify within-person variability in pulmonary function test (PFT) measures and identify baseline factors associated with variability.

Methods: Adult participants with AATD enrolled in the National Heart, Lung, and Blood Institute Registry (1989–1992) were eligible for inclusion in this analysis if they had 3 serial PFTs and severe genetic deficiency. PFT variability was evaluated using linear mixed-effects models with maximum likelihood estimation. Intraparticipant variability was assessed through residual analyses. Baseline predictors, including age, sex, augmentation therapy, liver disease, and FEV_1 percentage predicted, were examined as fixed effects.

Results: The analytic FEV_1 cohort (n=832) had a mean age of 46.3±10.1 years, 44.4% were female, and 99% were White. Over a mean of 4.4±1.4 years of follow-up, participants completed an average of 5.7±2.0 spirometry tests. FEV_1 varied by more than 10% of baseline in 82% of participants and by more than 20% in 55%. DLCO varied by more than 10% in 91% and by more than 20% in 43% of participants.

Conclusions: Substantial individual variability in FEV_1 and DLCO was observed in this large AATD cohort. These findings suggest that FEV_1 and DLCO are unreliable as clinical trial or stopping rule endpoints in AATD.

1. Division of Pulmonary, Critical Care, Allergy, and Sleep Medicine, Medical University of South Carolina, Charleston, South Carolina, United States
2. AlphaNet, Inc., Coral Gables, Florida, United States
3. Department of Medicine, National Jewish Health, Denver, Colorado, United States
4. College of Medicine, University of Kentucky, Lexington, Kentucky, United States
5. College of Public Health, University of Kentucky, Lexington, Kentucky, United States

Abbreviations:

AATD=alpha-1 antitrypsin deficiency; **AAT**=alpha-1 antitrypsin; **AIC**=Akaike information criterion; **BD**=bronchodilator; **BIC**=Bayesian information criterion; **BMI**=body mass index; **COPD**=chronic obstructive pulmonary disease; **CT**=computed tomography; **DLCO**=diffusing capacity of the lung for carbon monoxide; **FEV_1**=forced expiratory volume in 1 second; **FEV_1 %pred**=forced expiratory volume in 1 second percentage predicted; **FVC**=forced vital capacity; **LMM**=linear mixed-effects models; **MCID**=meaningful clinically important difference; **NHLBI**=National Heart, Lung, and Blood Institute; **PFT**=pulmonary function test; **SD**=standard deviation

Funding Support:

This work was a collaborative research project between Takeda Pharmaceutical Company and AlphaNet, Inc, where all statistical work was performed.

Citation:

Strange C, Holm KE, Sandhaus RA, Mannino DM, Choate R. Lung function variability in alpha-1 antitrypsin deficiency: implications for clinical trials. *Chronic Obstr Pulm Dis.* 2026;13(5):395-405. doi: https://doi.org/10.15326/jcopdf.2026.0790

Publication Dates:

Date of Acceptance: August 4, 2026
Published Online Date: August 10, 2026

Address correspondence to:

Charlie Strange, MD
Division of Pulmonary, Critical Care, Allergy, and Sleep Medicine
Medical University of South Carolina
Charleston, SC
Phone: (843) 792-3174
Email: strangec@musc.edu

Keywords:

alpha-1 antitrypsin deficiency; emphysema; forced expiratory volume in 1 second; clinical trial outcomes; pulmonary function test

This article has an online supplement.

Introduction

Alpha-1 antitrypsin deficiency (AATD) is a genetic cause of chronic obstructive pulmonary disease (COPD) resulting from mutations in the *SERPINA1* gene.[1] Although pulmonary emphysema is the predominant COPD endotype observed in AATD, airway hyperresponsiveness, chronic bronchitis, bronchiectasis, and normal lung function can occur among individuals with the ZZ genotype, the most common severe deficiency genotype.[2]

Despite a common genetic etiology, there is substantial heterogeneity in disease expression among individuals with AATD. This variability reflects the influence of environmental exposures on lung function, more frequent exacerbations compared with individuals with COPD without *SERPINA1* variants, and differences in therapeutic management.[3] Interpretation of lung function is further complicated by the frequent presence of hyperinflation due to emphysema, as spirometric variables and quantitative measures of emphysema are often poorly correlated, and these relationships can change over time.[4]

Variability in pulmonary function test (PFT) measurements is well recognized[5] and has been described in different patient populations, including those without lung disease and those with COPD or asthma. However, few studies have examined PFT variability specifically in AATD, particularly in individuals with manifestations of lung or liver disease.[6] Historically, various thresholds have been proposed to define clinically meaningful change in lung function over time (for example, absolute change of 100mL, percentage change of 10%, or short-term changes of 12% and 200mL for forced expiratory volume in 1 second [FEV_1]).[7-11] However, more recent guidelines acknowledge that PFT changes over time depend on multiple factors, including age, sex, baseline lung function, and underlying disease, which limits the generalizability of these thresholds and emphasizes the importance of individualized approaches for accurate interpretation.[12]

Since the 1980s, clinical trial designs in AATD populations have included FEV_1 and diffusing capacity of the lung for carbon monoxide (DLCO) as key outcome measures.[13,14] However, few studies have demonstrated that available therapies significantly alter these spirometric biomarkers.[15] In part, this may reflect the poor correlation between FEV_1, a measurement of airway flow, and emphysema, an alveolar disease. Moreover, DLCO is known to exhibit marked variability, which has limited its use in most clinical trial designs. More recently, changes in FEV_1 and DLCO have been proposed as stopping rules for some clinical trials in AATD. Stopping rules are designed to prevent study participant harm; however, if they are too rigid, then the validity of trial outcomes can be affected if participants do not get their active or placebo medications.

To address this gap, we aimed to characterize PFT variability and identify participant-related factors associated with PFT changes over time within a large longitudinal cohort of individuals with AATD. The overarching goal of the study was to generate robust evidence on the natural variability of lung function in AATD, with and without liver disease, to improve understanding of the disease trajectory and disease burden, and better distinguish clinically meaningful changes in PFTs over time from PFT variability.

Methods

This was a retrospective analysis of data from the "Registry of Patients with Severe Deficiency of Alpha 1-Antitrypsin," a longitudinal epidemiological study initiated by the National Heart, Lung, and Blood Institute (NHLBI) and the U.S. Food and Drug Administration.[16,17] The registry protocol was reviewed and approved by the appropriate institutional review board at each of the 37 participating clinical centers. The registry followed individuals with AATD for up to 7 years, with annual or semiannual spirometry measurements. Between March 1989 and October 1992, a total of 1129 individuals of ≥18 years of age were enrolled. Of these, 1026 had a serum alpha-1 antitrypsin (AAT) level ≤11μM and 103 had an SZ, ZZ, or ZNull genotype confirmed by DNA gene-probe analysis. Follow-up continued through April 1996. Among all enrolled participants, 927 had at least 2 postbronchodilator FEV_1 and forced vital capacity (FVC) measurements obtained at least 1 year apart.

Important to this study, rigorous spirometry quality control procedures were implemented across the 37 participating centers in the United States and Canada to ensure high-quality, reproducible spirometry results with daily calibration and biologic controls. High reproducibility rates were achieved for both prebronchodilator (95.0%) and postbronchodilator (95.7%) FEV_1 measurements at baseline.[18] Although interpretation strategies for spirometry have changed over the past 30 years, the recommended testing maneuvers have remained largely consistent. Therefore, this dataset was deemed suitable for evaluating longitudinal variability in spirometry and DLCO over time in individuals with AATD. Decisions about treatment with intravenous AAT were made by the participants' physicians, not by the registry. The smoking status of the registry participants was self-reported.

The inclusion criteria for this analysis were a serum AAT level ≤11μM or an SZ, ZZ, or ZNull genotype confirmed

by DNA gene-probe analysis. To generate reliable estimates of lung function slopes, participants were required to have at least 3 postbronchodilator FEV_1 measurements and 3 or more DLCO measurements within each 18 months of follow-up. Participants were censored from analysis at the time of lung transplantation. The final analytic sample included 832 individuals with 3 or more spirometry assessments and 586 with 3 or more DLCO measurements (Figure 1).

Predicted and percentage predicted values for FEV_1 and FVC were calculated using the Global Lung Function Initiative race-neutral reference equations, implemented through the R package rspiro (R Foundation; Vienna, Austria). Bronchodilator responsiveness was assessed using 2 approaches: the "old" criteria, defined as an increase of at least 12% from the prebronchodilator value and an absolute change of at least 200mL, and the "new" method, defined as an increase of at least 15% of the predicted value.

Statistical Analysis

Baseline demographic and clinical characteristics were summarized using descriptive statistics. Categorical variables were described using frequency and percentage within each category; continuous variables were described using mean and standard deviation. Frequencies and percentages of missing and non-missing observations were reported. Baseline characteristics were summarized for the entire NHLBI cohort, the analytic datasets, and relevant subgroup analyses.

PFT variability was evaluated using linear mixed-effects models (LMM), which are robust to unbalanced data and missing observations under maximum likelihood estimation. Model fit was evaluated using the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Intraparticipant variability was quantified through residual analysis, defined as the difference between observed PFT values and model-predicted values (Figure 2). Residual distributions were examined, and participant-level analyses were performed to characterize intraparticipant variability. Time to event analyses were performed by Kaplan- Meier methodology.

To evaluate the associations between PFT variability and baseline characteristics, baseline variables of interest (age, sex, augmentation therapy use, presence of liver disease, and FEV_1 percentage predicted [FEV_1 %pred]) were included as fixed effects in the LMMs. Interaction terms were incorporated as appropriate, and AIC and BIC were used to evaluate the best model fit. A sensitivity analysis was performed in the subset of participants who reported never smoking.

Analyses were conducted using SAS (SAS Institute; Cary, North Carolina) version 9.4.

Results

Baseline Characteristics of the Study Population

Of the overall NHLBI dataset (n=1126), 832 participants were included in the FEV_1/FVC analytic dataset based on the study inclusion criteria (Figure 1). Baseline characteristics of the participants included in the spirometry analysis were similar to those of the overall registry cohort (Table 1). In the spirometry cohort, the mean age was 46.3±10.1 years, 44.4% were female, and 99.0% were White. The majority (76.3%) had a history of smoking. In terms of alcohol use, 16.1% were former drinkers and 63.8% were current drinkers. Liver disease was present in 6.4% of the participants, whereas 83% had lung disease; 2 participants had undergone liver transplant at baseline. Augmentation therapy was used at baseline by 67.4% of the cohort. Participants completed an average of 5.7±2.0 spirometries over 4.4±1.4 years that were typically performed at 6-month intervals. Baseline characteristics of the DLCO analytic cohort (n=568) were similar.

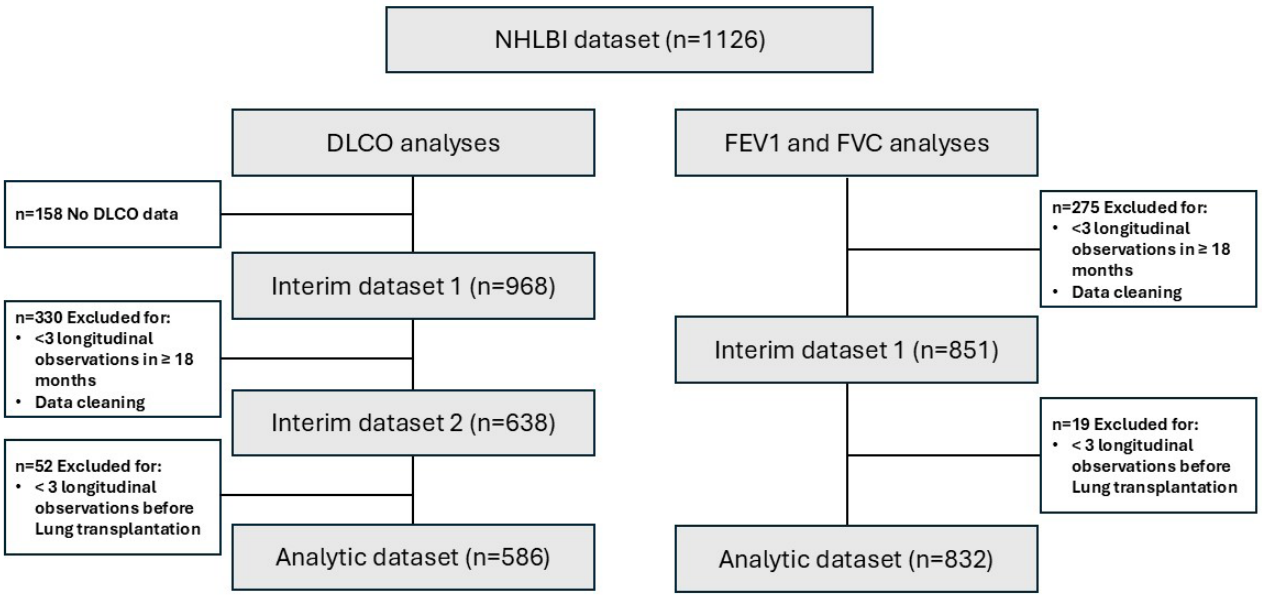
Baseline characteristics of the FEV_1/FVC analytic dataset were evaluated by several parameters: baseline FEV_1 %pred (<50% and ≥ 50%), use of augmentation therapy, presence of liver disease, and bronchodilator responsiveness (defined as ≥12% from prebronchodilator values and with ≥200ml change, and by a >10% change of the predicted FEV_1 value) (Supplemental Tables 1.1–1.4 in the online supplement).

Forced Expiratory Volume in 1 Second Variability

In 82.1% of participants (683/832), FEV_1 had a 10% or more change from the baseline value on at least one occasion. The frequency distribution of the largest percentage change over time is shown in Figure 3. Among all participants whose largest observed change was a decline, the mean largest decline in FEV_1 from baseline was 23.5%, while those with their largest FEV_1 change as an increase had a mean largest change of 19.4%. The distribution of the number of events with a ≥10% or ≥20% change in FEV_1 from baseline is shown in Figure 4. Overall, most participants demonstrated FEV_1 variability ≥10% from baseline on multiple occasions. When applying varying thresholds of FEV_1 percentage change from baseline (Supplemental Table 1.5 in the online supplement), 44.8% of the cohort still had changes that exceeded a 20% threshold. The frequency distribution of the number of events surpassing the ≥20% threshold is also shown in Figure 4.

Among all study participants in the spirometry analytic cohort, the mean annual decline in FEV_1 was 55ml, consistent with that reported in the larger NHLBI cohort. We observed a steeper FEV_1 decline (-63.33ml/year) in males compared to females (-44.59ml/year) (p<0.0001). In addition,

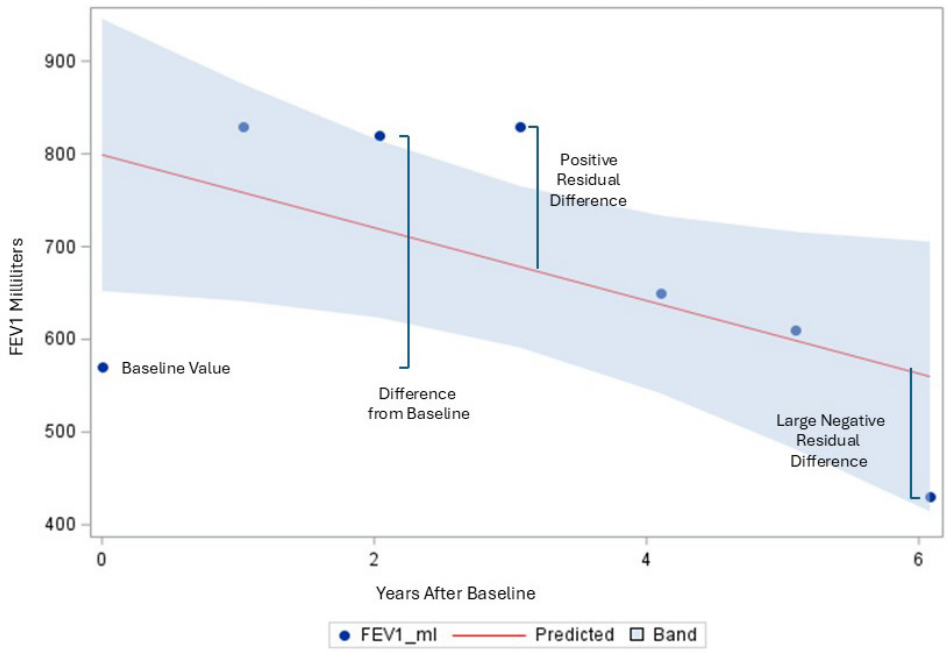
Figure 1. Study Flow for Forced Expiratory Volume in 1 Second, Forced Vital Capacity, and Diffusing Capacity of the Lung for Carbon Monoxide Analytic Cohorts in the National Heart, Lung, and Blood Institute Dataset



A minimum of 3 values were required for inclusion in the current study

NHLBI=National Heart, Lung, and Blood Institute; DLCO=diffusing capacity of the lung for carbon monoxide; FEV_1=forced expiratory volume in 1 second; FVC=forced vital capacity

Figure 2. Example of Intra-Person Variability in Forced Expiratory Volume in 1 Second Based on a Study Participant Example



A representative patient had a low baseline FEV_1 that improved over the following 3 visits. The difference from baseline and the residual differences from the line of regression give differential insight into stopping rules for clinical trials.

FEV_1=forced expiratory volume in 1 second

younger individuals aged <39 years had a steeper decline (-69.21ml/year) compared to those aged >52 (-45.39ml/year) (p<0.0001). FEV_1 slope was not different between smokers and nonsmokers, likely because most individuals had stopped smoking during the NHLBI registry. Detailed results are presented in the supplemental materials (Supplemental Figures 1.1–1.6 in the online supplement).

Residual-based analysis provided a more precise evaluation

Table 1. Baseline Characteristics of the Study Population in the Overall Cohort, and Analytic Subsets For Spirometry and Diffusing Capacity of the Lung for Carbon Monoxide Analyses

Baseline Characteristics	Overall Dataset (n=1126)	FEV_1 and FVC Analytic Sample (n=832)	DLCO Analytic Sample (n=586)
Age, years, mean (± SD)	46.6 (10.5)	46.3 (10.1)	46.0 (10.2)
Sex, female, n (%)	500 (44.4)	369 (44.4)	261 (44.5)
Race, White, n (%)	1117 (99.2)	824 (99.0)	579 (98.8)
BMI, mean (± SD)	25.8 (5.4)	26.3 (5.3)	26.3 (5.2)
Current Employment Status, n (%)	n=1123	n=829	n=585
Employed Full-Time	477 (42.5)	396 (47.8)	282 (48.2)
Employed Part-Time	121 (10.8)	99 (11.9)	77 (13.2)
Unemployed (Medical Reasons)	226 (20.1)	139 (16.8)	91 (15.6)
Unemployed (Other Reasons)	38 (3.4)	23 (2.8)	19 (3.3)
Retired (Medical)	111 (9.9)	70 (8.4)	49 (8.4)
Retired (Age)	53 (4.7)	36 (4.3)	25 (4.3)
Full-Time Homemaker	70 (6.2)	46 (5.6)	30 (5.1)
Other	27 (2.4)	20 (2.4)	12 (2.1)
Marital Status, n (%)	n=1122	n=828	n=585
Never Married	84 (7.5)	60 (7.3)	47 (8.0)
Widowed	26 (2.3)	14 (1.7)	10 (1.7)
Married	843 (75.1)	637 (76.9)	439 (75.0)
Separated	22 (2.0)	13 (1.6)	12 (2.1)
Divorced	147 (13.1)	104 (12.6)	77 (13.2)
Education, n (%)			
College Graduate With Advanced Training	167 (14.8)	131 (15.8)	99 (16.9)
College Graduate	160 (14.2)	123 (14.8)	89 (15.2)
At Least 1 Year of College	301 (26.7)	236 (28.4)	165 (28.2)
High School Graduate	389 (34.6)	269 (32.3)	186 (31.7)
Completed 10–11 Years of School	64 (5.7)	47 (5.7)	33 (5.6)
Unknown/Other	45 (4.0)	26 (3.1)	14 (2.4)
Smoking History, n (%)	n=1122	n=828	n=584
Yes	857 (76.4)	632 (76.3)	437 (74.8)
Former Smoker	728/857 (84.9)	544/632 (86.1)	382/437 (87.4)
Current Smoker	114/857 (13.3)	76/632 (12.0)	47/437 (10.8)
Pack Years	n=798	n=589	n=411
Mean (SD)	23.1 (16.5)	22.7 (15.0)	22.4 (15.0)
Alcohol Use, n (%)	n=1121	n=828	n=584
Never	254 (22.7)	167 (20.2)	106 (18.2)
Former Drinker	195 (17.4)	133 (16.1)	89 (15.2)
Current Drinker	672 (60.0)	528 (63.8)	389 (66.6)
Liver Disease, n (%)	n=1122	n=828	n=584
Yes	84 (7.5)	53 (6.4)	41 (7.0)
Cirrhosis	29/84 (34.5)	13/53 (24.5)	13/41 (31.7)
Neonatal Jaundice	25/84 (29.8)	19/53 (35.9)	14/41 (34.2)
Hepatitis	42/84 (50.0)	27/53 (50.9)	20 (48.8)
Lung Disease, n (%)	n=1121	n=827	n=583
Yes	925 (82.5)	686 (83.0)	482 (82.7)
Emphysema	820/925 (88.7)	601/ 686 (87.6)	419/482 (86.9)
Reversible Airways Disease	396/922 (43.0)	314/685 (45.8)	220/482 (45.6)
Bronchitis	463/925 (50.1)	346/686 (50.4)	243/482 (50.4)
Liver/Lung Transplant, yes, n (%)	n=6 lung; n=2 liver at baseline	n=2 liver at baseline	n=0 at baseline
Augmentation Therapy, always/sometimes, n (%)	674 (59.9)	561 (67.4)	410 (69.9)
FEV_1 %pred			
<50%		515 (61.9)	343 (58.5)
≥50%		317 (38.1)	243 (41.5)
Bronchodilator Responsiveness, yes, n (%)		n=814	n=575
Old Method (≥12% from pre-BD and ≥200ml change)		249 (30.6)	194 (33.7)
New Method (≥15% from predicted)		17 (2.1)	11 (1.9)
Number of Follow-Up Visits, mean (SD)		5.7 (2.0)	5.4 (1.9)
Number of Follow-Up Years, mean (SD)		4.4 (1.4)	4.3 (1.4)

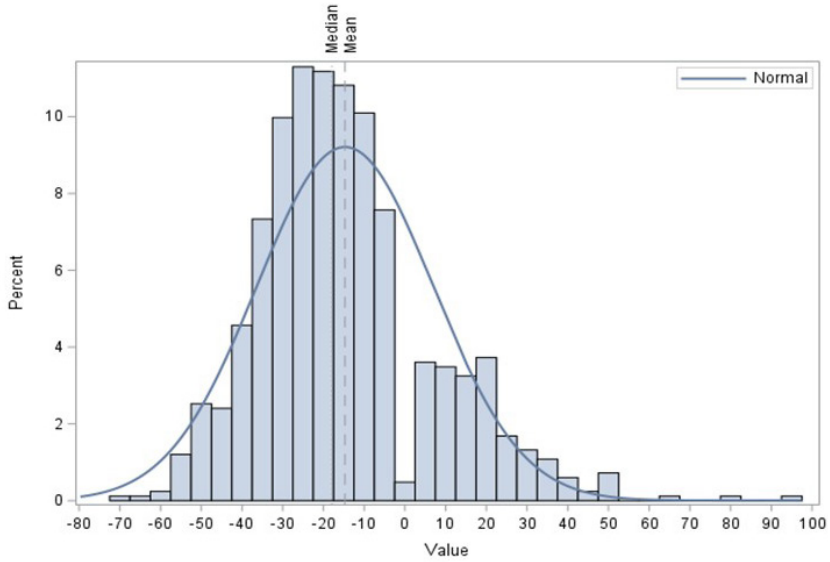
FEV_1=forced expiratory volume in 1 second; FVC=forced vital capacity; DLCO=diffusing capacity of the lung for carbon monoxide; SD=standard deviation; BMI=body mass index; FEV_1 %pred=forced expiratory volume in 1 second percentage predicted; BD=bronchodilator

of longitudinal differences in FEV_1. The baseline lung function value heavily influenced subsequent variability. In the example given in Figure 2, a low baseline FEV_1 value heavily influenced the chances of a downstream event. This supports the practice in clinical trials of performing spirometry twice at baseline to establish a more robust reference point for slope estimation.

Forced Vital Capacity Variability

Based on the largest difference value from baseline per patient, 75.6% of the cohort had one or more events of ≥10% change in FVC% from the baseline. Nearly half (47.24%) of the cohort had a decline that was 10% or more from

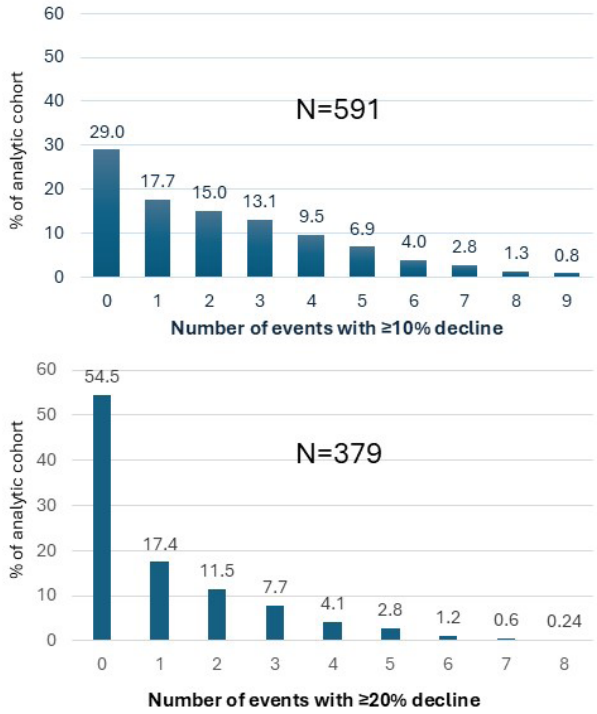
Figure 3. Distribution of the Largest Percentage Change in Forced Expiratory Volume in 1 Second from Baseline[a]



[a] n=832

FEV_1=forced expiratory volume in 1 second

Figure 4. Number of Events With a (A) 10% and (B) 20% Change in Forced Expiratory Volume in 1 Second from Baseline in the Analytic Sample of n=832



the baseline. The median value of the largest difference per patient in FVC% was -8.87%. The average annual decline in FVC in the overall cohort was 58ml. We did not observe differences in FVC slopes by sex, age, augmentation therapy, or liver disease at baseline (data not shown).

Diffusing Capacity of the Lung for Carbon Monoxide Variability

Among the 586 individuals in the DLCO analytic cohort, the mean baseline DLCO was 17.3±7.5mL/min/mmHg. The majority (91.3%) of the cohort experienced a change from baseline of ≥10%, 58.4% of the cohort experienced

that change as a ≥10% decrease from baseline, and 43.2% experienced a ≥20% decrease. The results of the DLCO analysis are included in the Supplemental materials (Supplemental Figures 2.1–2.7 in the online supplement).

Kaplan-Meier Analyses

Time to event analyses in Figure 5 show that half of participants reached a 10% decline in FEV_1 at 3.0 years and a 10% decline in DLCO at 2.9 years. Median time to 20% decline occurred at 5.1 years for FEV_1 and 5.2 years for DLCO (Table 2). The number of participants who met these thresholds appears linear over the 5 years that the majority of participants were being evaluated.

Discussion

Variability in pulmonary function testing remains a major challenge in clinical trials across all obstructive pulmonary diseases. Consequently, marked time and dedication applied to minimizing testing variability have been used, including standardizing spirometry equipment to the same machine and the same operator for study durability. However, these strategies have had limited success, as FEV_1 and DLCO are intrinsically affected by participant effort, emphysema, mucus, airway tone, medications, and environmental factors. As a result, large sample sizes are typically required for studies in COPD and asthma to overcome this inherent variability. Further, the FEV_1 is an intrinsic target of bronchodilator medications, but not medications that have no bronchodilator activity.

In AATD, FEV_1 and DLCO are insensitive biomarkers of disease progression. Emphysema on computed tomography (CT) imaging clearly precedes any decline in FEV_1 in the natural history of disease. As emphysema develops, lung volumes often increase. The subsequent increase in FVC usually elevates FEV_1 in mild disease before FEV_1 begins to fall from airways disease and bronchiolar collapse. Lung function also changes during exacerbations, and recovery periods in CT and spirometry may take 6 weeks or longer to show. Thus, the variability in FEV_1 does not reproducibly reflect the progression of emphysema. Current state-of-the-art monitoring of PFTs in COPD suggests that measuring more frequently than every 6–12 months does not reveal changes that reflect disease progression. This underscores the need for other, more specific biomarkers of emphysema progression.

In addition, our analysis of intraparticipant variability revealed that sex and age significantly influenced pulmonary function decline. Males exhibited a steeper decline in FEV_1 and DLCO compared to females, while younger individuals (<39 years) experienced a more pronounced decline in FEV_1 than those over 52 years. However, these findings should be interpreted with caution, as the changes are small and may reflect smoking habits common in the 1990s. Further, these observations may be the effect of higher losses if lung function is preserved, since there is more opportunity to lose volume from higher baseline values.

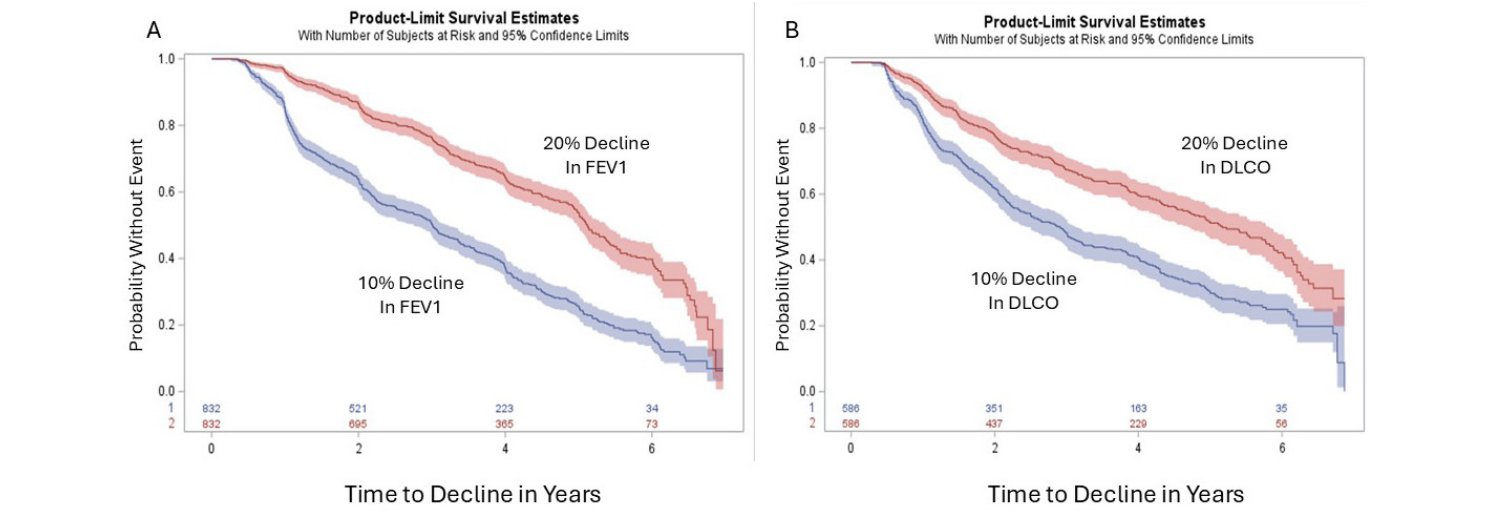
Increases of 10% in either measure are almost as frequent as decreases, highlighting the magnitude of natural variability even under standardized conditions. These findings align with prior observations in AATD cohorts and in COPD, where repeated spirometry frequently fails to capture clinically meaningful decline.[7] For example, studies have shown that spirometric follow-up over 5 years identifies only about half of patients with rapid FEV_1 decline, and more than half of those with CT-based emphysema progression show no parallel FEV_1 change.[19,20] Similarly, studies show that both FEV_1 and DLCO exhibit day-to-day variability, with intra-individual variability reaching up to 10% depending on device and test conditions.[21,22] No technique can be universally applied to improve this natural variability seen in the majority of AATD trials. Alternative biomarkers such as CT densitometry and enhancements to CT scan metrics of airway and alveolar function, which directly quantify lung tissue density and correlate more closely with emphysema progression, may offer greater sensitivity and reproducibility.[23]

The high intrinsic variability of PFTs in AATD reflects several factors: the effort-dependent nature of spirometry, the heterogeneity of lung involvement, and the impact of exacerbations (which occur on average twice yearly in AATD patients),[24] and differences in airway versus alveolar pathology. Furthermore, reliance on single measurements amplifies these effects, and at least 3 serial PFTs are needed to estimate a reliable slope of decline. Further, the difficulty finding a meaningful clinically important difference (MCID) in usual COPD that requires triangulation with meaningful clinical events is even more difficult in AATD. This study would suggest that an AATD MCID for FEV_1 or DLCO may be an abstract concept.

Historically, AATD clinical trials have used FEV_1 and DLCO as preferred clinical outcome measures. However, few trials have been successful when these outcomes are used. As an example, the original NHLBI registry showed no difference in FEV_1 slope associated with use of augmentation therapy in the entire cohort,[25] despite that fact that mortality differences were present.[26]

The current analysis demonstrates that neither FEV_1 nor DLCO is suitable as a stopping rule in AATD clinical trials, due to the high intrinsic variability.[27] There are many reasons that stopping rules can fail to meet the goal of participant safety protection. The most important aspects of this study show a very high intrinsic variability. Further, the natural history of lung function decline suggests that over time all participants would be stopped from receiving the drug or placebo. Because the average annual decline in FEV_1

Figure 5. Time to Events of 10% and 20% Decline in (A) Forced Expiratory Volume in 1 Second and (B) Diffusing Capacity of the Lung for Carbon Monoxide



FEV_1=forced expiratory volume in 1 second; DLCO=diffusing capacity of the lung for carbon monoxide

Table 2. Time to 10% and 20% Decline in Forced Expiratory Volume in 1 Second and Diffusing Capacity of the Lung for Carbon Monoxide by Kaplan-Meier Analysis

Percentage	Point Estimate	95% Confidence Interval		
		Transform	Lower	Upper
		Time to 10% Decline in Years		
		Quartile Estimates- FEV_1		
75	5.01000	LOGLOG	4.64000	5.26000
50	3.01000	LOGLOG	2.75000	3.23000
25	1.19000	LOGLOG	1.13000	1.37000
		Quartile Estimates- DLCO		
75	5.78000	LOGLOG	5.08000	6.71000
50	2.91000	LOGLOG	2.44000	3.13000
25	1.20000	LOGLOG	1.09000	1.48000
		Time to 20% Decline in Years		
		Quartile Estimates- FEV_1		
75	6.60000	LOGLOG	6.48000	6.87000
50	5.12000	LOGLOG	4.98000	5.37000
25	3.05000	LOGLOG	2.85000	3.24000
		Quartile Estimates- DLCO		
75	.	LOGLOG	6.44000	.
50	5.17000	LOGLOG	4.69000	5.83000
25	2.16000	LOGLOG	1.96000	2.65000

FEV_1=forced expiratory volume in 1 second; DLCO=diffusing capacity of the lung for carbon monoxide

in this cohort was approximately 55ml, the frequency of reaching stopping thresholds of >10% decline from baseline FEV_1 measures due to natural history progression is high. A more meaningful threshold of at least 20% would better correlate with clinical disease progression, but any threshold also depends on trial duration. Conceivably, a yearly discount in expected values from baseline lung function would be needed to prevent cessation of participation in all study participants over time. Therefore, adjusting expected lung function over time is a more rational scientific methodology for clinical trial stopping rules compared to an established 10% stopping rule from baseline that lives throughout the life of a clinical trial.

Our study has some limitations. Clinical care for the NHLBI cohort participants was not centralized and may have led to systematic differences in outcomes depending on the clinical centers and the managing physicians. Other unobserved factors, such as intensity of care received and the daily behaviors of participants, could have confounded

the relationship between pulmonary function change and baseline characteristics. Data on several variables included in this study are self-reported and subject to recall and reporting bias. The NHLBI cohort did not query COPD exacerbations in real time, and some of the observed variability events may reflect exacerbations that would be recognized in a more robust clinical trial design today. Although follow-up visits were delayed up to 6 weeks after exacerbations in the NHLBI study, to attempt to mitigate this, residual effects likely remained. In addition, excluding participants with less than 3 lung function measurements that are at least 6 months apart may have introduced potential selection bias. Thus, the results of this study may not be generalizable to all individuals with AATD. Participants who returned for serial testing (832 for spirometry and 586 for DLCO) represented a subset of the full NHLBI cohort and may differ from others in the AATD populations at large. In general, this population had more lung disease than many of the liver-focused studies today. Additionally, a floor effect exists for both FEV_1 and DLCO in which PFTs do not fall as quickly with advanced lung disease.

Since these individuals were included in the NHLBI cohort, the variability is likely to be even higher in those with higher baseline lung function. Our findings demonstrate that higher baseline lung function is associated with substantial variability in FEV_1 and DLCO. We were surprised to see that FEV_1 variability is nearly identical to DLCO variability in time course and intensity. As such, neither FEV_1 nor DLCO are fit for purpose to scientifically serve as effective stopping rules for AATD studies.

In summary, the NHLBI registry represents a herculean effort to enroll and follow the largest number of AATD-affected individuals in a prospective study that has ever been done in this rare disease. The data are robust and remain relevant today, since few new therapies have emerged that may change the natural history of AATD. As gene therapy and RNA-based treatments advance, trial designs should incorporate more robust and biologically relevant endpoints, such as the serum AAT concentration, rather than depending on spirometric or diffusing capacity changes.

Acknowledgements

Author contributions: CS, KEH, RAS, DMM, and RC participated in study design, analysis, interpretation, and approved the final manuscript.

Other acknowledgements: The authors thank the alpha-1 antitrypsin deficiency-affected individuals who participated in the NHLBI registry from 1989–1992.

Declaration of Interest

CS is a medical director for AlphaNet and Pulmanage. He has grants paid to the Medical University of South Carolina from Beam, Biomarin, CSA Medical, Grifols, Nuvaira, Renovion, Takeda, Tessera and Zion for COPD. He consults for AstraZeneca, Intuitive Surgical, and Sanofi. KEH and RC receive consulting and research income from AlphaNet. RAS is a medical director for AlphaNet. DMM is a consultant to AstraZeneca, GlaxoSmithKline, Lilly, Regeneron, Sanofi, Genentech, Chiesi, and Up-to-Date. He is the Chief Medical Officer for the COPD Foundation and a medical expert on behalf of people suing the tobacco and vaping industries.

References

1. 1. Stoller JK, Aboussouan LS. A review of α1-antitrypsin deficiency. *Am J Respir Crit Care Med.* 2012;185(3):246-259. https://doi.org/10.1164/rccm.201108-1428CI

2. Strange C. Airway disease in alpha-1 antitrypsin deficiency. *COPD.* 2013;10(sup1):68-73. https://doi.org/10.3109/15412555.2013.764404

3. Turner AM, Ficker JH, Vianello A, et al. Advancing the understanding and treatment of lung pathologies associated with alpha 1 antitrypsin deficiency. *Ther Adv Respir Dis.* 2025;19. https://doi.org/10.1177/17534666251318841

4. Miravitlles M, Dirksen A, Ferrarotti I, et al. European Respiratory Society statement: diagnosis and treatment of pulmonary disease in α1-antitrypsin deficiency. *Eur Respir J.* 2017;50(5):1700610. https://doi.org/10.1183/13993003.00610-2017

5. Berry G. Longitudinal observations. Their usefulness and limitations with special reference to the forced expiratory volume. *Bull Physiopathol Respir (Nancy).* 1974;10(5):643-655. https://pubmed.ncbi.nlm.nih.gov/4441757/

6. DeMeo D, Campbell E, Brantly M, et al. Heritability of lung function in severe alpha-1 antitrypsin deficiency. *Hum Hered.* 2008;67(1):38-45. https://doi.org/10.1159/000164397

7. Donohue JF. Minimal clinically important differences in COPD lung function. *COPD.* 2005;2(1):111-124. https://doi.org/10.1081/COPD-200053377

8. Drummond MB, Schwartz PF, Duggan WT, et al. Intersession variability in single-breath diffusing capacity in diabetics without overt lung disease. *Am J Respir Crit Care Med.* 2008;178(3):225-232. https://doi.org/10.1164/rccm.200801-090OC

9. Hathaway EH, Tashkin D, Simmons MS. Intraindividual variability in serial measurements of DLCO and alveolar volume over one year in eight healthy subjects using three independent measuring systems. *Am Rev Respir Dis.* 1989;140(6):1818-1822. https://doi.org/10.1164/ajrccm/140.6.1818

10. Pellegrino R, Viegi G, Brusasco V, et al. Interpretative strategies for lung function tests. *Eur Respir J.* 2005;26(5):948-968. https://doi.org/10.1183/09031936.05.00035205

11. Wise RA, Connett J, Kurnow K, et al. Selection of spirometric measurements in a clinical trial, the Lung Health Study. *Am J Respir Crit Care Med.* 1995;151(3):675-681. https://doi.org/10.1164/ajrccm.151.3.7881655

12. Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J.* 2022;60(1):2101499. https://doi.org/10.1183/13993003.01499-2021

13. Rahaghi FF, Miravitlles M. Long-term clinical outcomes following treatment with alpha 1-proteinase inhibitor for COPD associated with alpha-1 antitrypsin deficiency: a look at the evidence. *Respir Res.* 2017;18:105. https://doi.org/10.1186/s12931-017-0574-1

14. Santos G, Turner AM. Alpha-1 antitrypsin deficiency: an update on clinical aspects of diagnosis and management. *Fac Rev.* 2020;9:1. https://doi.org/10.12703/b/9-1

15. Edgar RG, Patel M, Bayliss S, Crossley D, Sapey E, Turner AM. Treatment of lung disease in alpha-1 antitrypsin deficiency: a systematic review. *Int J Chron Obstruct Pulmon Dis.* 2017;12:1295-1308. https://doi.org/10.2147/COPD.S130440

16. McElvaney NG, Stoller JK, Buist AS, et al. Baseline characteristics of enrollees in the National Heart, Lung and Blood Institute Registry of alpha1-antitrypsin deficiency. *Chest.* 1997;111(2):394-403. https://doi.org/10.1378/chest.111.2.394

17. The Alpha 1-Antitrypsin Deficiency Registry Study Group. A registry of patients with severe deficiency of alpha 1-antitrypsin. Design and methods. *Chest.* 1994;106(4):1223-1232. https://doi.org/10.1378/chest.106.4.1223

18. Stoller JK, McCarthy K, Buist AS, et al. Quality control of spirometry testing in the registry for patients with severe alpha-1 antitrypsin deficiency. *Chest.* 1997;111(4):899-909. https://doi.org/10.1378/chest.111.4.899

19. Lundquist A, Lindberg A, Eriksson-Ström J, Blomberg A, Backman H. Number of follow-up years needed to identify a rapid decline in FEV_1. *Am J Respir Crit Care Med.* 2024;209(1):119-120. https://doi.org/10.1164/rccm.202309-1664LE

20. Green CE, Parr DG, Edgar RG, Stockley RA, Turner AM. Lung density associates with survival in alpha 1 antitrypsin deficient patients. *Respir Med.* 2016;112:81-87. https://doi.org/10.1016/j.rmed.2016.01.007

21. Jensen RL, Teeter JG, England RD, White HJ, Pickering EH, Crapo RO. Instrument accuracy and reproducibility in measurements of pulmonary function. *Chest.* 2007;132(2):388-395. https://doi.org/10.1378/chest.06-1998

22. Madsen F, Ulrik CS, Dirksen A, et al. Patient-administered sequential spirometry in healthy volunteers and patients with α1-antitrypsin deficiency. *Respir Med.* 1996;90:131-138. https://doi.org/10.1016/S0954-6111(96)90154-7

23. Crossley D, Renton M, Khan M, Low EV, Turner AM. CT densitometry in emphysema: a systematic review of its clinical utility. *Int J Chron Obstruct Pulmon Dis.* 2018;13:547-563. https://doi.org/10.2147/COPD.S143066

24. Choate R, Sandhaus RA, Holm KE, Mannino DM, Strange C. Patient-reported pulmonary symptoms, exacerbations, and management in a cohort of patients with alpha-1 antitrypsin deficiency. *Chronic Obstr Pulm Dis.* 2022;9(4):549-561. https://doi.org/10.15326/jcopdf.2022.0317

25. The Alpha-1-Antitrypsin Deficiency Registry Study Group. Survival and FEV_1 decline in individuals with severe deficiency of alpha1-antitrypsin. *Am J Respir Crit Care Med.* 1998;158(1):49-59. https://doi.org/10.1164/ajrccm.158.1.9712017

26. Stoller JK, Tomashefski J Jr, Crystal RG, et al. Mortality in individuals with severe deficiency of alpha1-antitrypsin: findings from the National Heart, Lung, and Blood Institute Registry. *Chest.* 2005;127(4):1196-1204. https://doi.org/10.1016/S0012-3692(15)34467-6

27. Turner AM, Whittkopf P, von Wilamowitz-Moellendrorff C, et.al. Pulmonary function decline in alpha-1 antitrypsin deficiency: a systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis.* 2026;21:552227. https://doi.org/10.2147/COPD.S552227