Rationale: It has been suggested that patients with chronic obstructive pulmonary disease (COPD) experience considerable daily respiratory symptom fluctuation. A standardized measure is needed to quantify and understand the implications of day-to-day symptom variability.

Objectives: To compare standard deviation with other statistical measures of symptom variability and identify characteristics of individuals with higher symptom variability.

Methods: Individuals in the SubPopulations and InteRmediate Outcome Measures In COPD Study (SPIROMICS) Exacerbations sub-study completed an Evaluating Respiratory Symptoms in COPD (E-RS) daily questionnaire. We calculated within-subject standard deviation (WS-SD) for each patient at week 0 and correlated this with measurements obtained 4 weeks later using Pearson’s r and Bland Altman plots. Median WS-SD value dichotomized participants into higher versus lower variability groups. Association between WS-SD and exacerbation risk during 4 follow-up weeks was explored.

Measurements and Main Results: Diary completion rates were sufficient in 140 (68%) of 205 sub-study participants. Reproducibility (r) of the WS-SD metric from baseline to week 4 was 0.32. Higher variability participants had higher St George’s Respiratory Questionnaire (SGRQ) scores (47.3±20.3 versus 39.6±21.5, p=.04) than lower variability participants. Exploratory analyses found no relationship between symptom variability and health care resource utilization-defined exacerbations.

Conclusions: WS-SD of the E-RS can be used as a measure of symptom variability in studies of patients with COPD. Patients with higher variability have worse health-related quality of life. WS-SD should be further validated as a measure to understand the implications of symptom variability.

Abbreviations: SubPopulations and InteRmediate Outcome Measures In COPD Study, SPIROMICS; Evaluating Respiratory Symptoms in COPD, E-RS; within-subject standard deviation, WS-SD; St George’s Respiratory Questionnaire, SGRQ; patient-reported outcome, PRO; Exacerbations in Chronic Pulmonary Disease Tool, EXACT; square root of the wall area of a hypothetical airway with an internal perimeter of 10 mm, Pi10; health care resource utilization, HCRU; interquartile range, IQR; modified Medical Research Council, mMRC; Pittsburgh Sleep Quality Index, PSQI; hospital anxiety and depression scale, HADS; parametric response mapping, PRM; Global initiative for chronic Obstructive Lung Disease, GOLD; confidence interval, CI; correlation, corr; Body mass index-airway Obstruction-Dyspnea-Exercise capacity, BODE.
COPD Symptom Variability and the E-RS COPD Instrument

Keywords:
chronic obstructive pulmonary disease; patient-reported outcomes; exacerbations; EXACT; symptom variation

Introduction

Patients with chronic obstructive pulmonary disease (COPD) experience respiratory symptoms including breathlessness, cough, and sputum production that necessitate comprehensive assessment.¹ This assessment, and hence, therapeutic decision-making, has not adequately incorporated symptom variability. Prior studies have examined the impact of morning symptoms and have correlated these to more frequent exacerbations and greater impact on health-related quality of life.²⁻⁷ Weaknesses of studies assessing symptom variability during the stable state include cross-sectional design and surveys dependent on participant recall of symptoms.²⁻¹¹ Two prospective studies that examined symptom variability were limited by small sample size and a short follow-up period.⁹⁻¹² The more recent prospective study evaluated 2669 participants over 1 week using the Night-time and the Early Morning Symptoms of COPD instruments and found more severe disease was associated with fluctuation in symptom number and intensity.¹² A metric to define and compare individual symptom variability has not been developed.

Given correlation between morning symptom differences, quality of life, and exacerbation risk, the development of a symptom variability metric using prospectively collected data over a longer time in a highly phenotyped sample could contribute to comprehensive patient assessment. The SubPopulations and InteRmediate Outcome Measures In COPD Study (SPIROMICS) cohort provides a unique opportunity to evaluate the variability of respiratory symptoms prospectively. In the Exacerbation

For personal use only. Permission required for all other uses.

1. Pulmonary and Critical Care Medicine, Weill Cornell Medicine, New York, New York, United States
2. Department of Medicine, Weill Cornell Medicine, New York, New York, United States
3. Division of Biostatistics, Department of Population Health Sciences, Weill Cornell Medical College, New York, New York, United States
4. Evidera, Bethesda, Maryland, United States
5. Division of Pulmonary and Critical Care Medicine, University of Michigan Health System, Ann Arbor, Michigan, United States
6. Department of Medicine, National Jewish Medical and Research Center, Denver, Colorado, United States
7. Pulmonary and Critical Care, University of California San Francisco, San Francisco, California, United States
8. University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States
9. Pulmonary and Critical Care Medicine, Temple University Hospital, Philadelphia, Pennsylvania, United States
10. Medical Service, VA Ann Arbor Healthcare System, Ann Arbor, Michigan, United States
11. Pulmonary, Allergy and Critical Care, University of Alabama at Birmingham, Birmingham, Alabama, United States
12. Division of Pulmonary and Critical Care Medicine, Johns Hopkins University, Baltimore Maryland, United States
13. Division of Pulmonary and Critical Care Medicine, Department of Veterans Affairs Medical Center, University of Utah, Salt Lake City, Utah, United States
14. Section on Pulmonary, Critical Care, Allergy and Immunological Diseases, Wake Forest School of Medicine Medical Center, Winston-Salem, North Carolina, United States
15. National Heart and Lung Institute, Imperial College London, United Kingdom

Address correspondence to:
Jamura K. Krishnan, MD
Division of Pulmonary and Critical Care Medicine
Weill Cornell Medicine
1305 York Avenue, Y-1047, Box 96
New York, NY 10021
Phone: 646-962-2333
Email: jkk9002@med.cornell.edu

For personal use only. Permission required for all other uses.
sub-study, a subset of SPIROMICS participants completed the Evaluating Respiratory Symptoms in COPD (E-RSTM:COPD or E-RS),13,14 an 11-question patient-reported outcome (PRO) diary administered as part of the 14-item Exacerbation in Chronic Pulmonary Disease Tool (EXACT). The E-RS has been shown to be a reliable, valid, and responsive measure of respiratory symptom severity in stable COPD and is sensitive to the treatment effects of drug therapies.14-17 Measuring fluctuation in E-RS scores provides an opportunity to characterize day-to-day symptom variability as an additional dimension of symptom burden. A standardized measurement of symptom variability could in turn be tested for therapeutic responsiveness in future studies. Standard deviation is a common statistical measure to quantify variation and dispersion of biomedical data.18 We hypothesized that the within-subject standard deviation (WS-SD) of E-RS scores can serve as a potential metric to quantify variability of respiratory symptoms in patients with COPD.

Our objectives were as follows: (1) to evaluate the performance of the WS-SD variability metric compared to other statistical measures of variation, (2) to identify differences in baseline characteristics between individuals with higher versus lower WS-SD, and (3) explore the extent to which baseline symptom variability predicted future exacerbations. Preliminary results have been reported in an abstract.19

Materials and Methods

Study Cohort

SPIROMICS is a prospective cohort study that enrolled 2981 participants across 4 strata (never-smokers, smokers without COPD, mild/moderate COPD, and severe COPD) with the goals of identifying new COPD subgroups and markers of disease progression. The SPIROMICS protocol details full inclusion and exclusion criteria and assessments performed at baseline.13 In addition to these assessments, the main SPIROMICS study obtained additional imaging phenotypic markers including the square root of the wall area of a hypothetical airway with an internal perimeter of 10 mm (Pi10) and parametric response mapping.20

Data for these analyses were derived from the subgroup of participants with COPD participating in the SPIROMICS Exacerbation sub-study; this was incremental to the previously published SPIROMICS-wide exacerbation data.13 To be representative of patients at a baseline stable state, sub-study eligibility criteria included being free of a health care resource utilization (HCRU)-defined exacerbation in the 30 days before enrollment. Exclusion criteria included a primary diagnosis of asthma or visual or cognitive impairment preventing completion of the daily diary using a personal digital assistant. Institutional review boards at each center approved SPIROMICS, and all participants provided informed consent.

Symptom and Exacerbation Assessment

Sub-study participants completed the E-RS as part of their daily study diary. The E-RS total score ranges from 1 to 40 with higher scores indicating more severe symptoms, overall.21 Consistent with E-RS scoring guidelines, participants with at least 4 days of diary data per week and 80% completion rate of daily diary data for the first 5 weeks were included in the analysis; participants who did not meet this diary completion rate were excluded.

Exacerbations were defined by HCRU, i.e., an increase in respiratory symptoms that required a health care visit and were treated with antibiotics, systemic corticosteroids, or both.20 In addition, symptom-defined EXACT events were identified using the 14-item EXACT score, which ranges from 0-100 with higher scores indicating a worse health state. EXACT (symptom-defined) events, which are periods of acute sustained symptom worsening, are characterized by an increase from a stable baseline state in an individual's EXACT score of >9 points for 3 consecutive days, or >12 points for 2 consecutive days.22,23

Study Design and Statistical Methods

We compared the performance of the WS-SD of E-RS total scores calculated from the first week (baseline or Week 0) to 5 other metrics for estimating variability: interquartile range (IQR), range, median absolute deviation, variance, and coefficient of variation. The week 0 value for each variability measure was compared with the week 4 value using Pearson’s correlation to determine reproducibility. We constructed Bland-Altman plots for each variability metric also comparing week 0 values to week 4 values. Our main analysis focused on the first 5 weeks because prior to sub-study enrollment, participants were free of HCRU exacerbation for at least 30 days. We were interested in symptom variability during a time period...
closest to this stable state. A schematic of our study and timing of assessments is presented in Figure 1. Finally, we examined the change in WS-SD over each of 13 weeks and used paired t-tests to compare changes.

The week 0 median value of the E-RS total score WS-SD was used to dichotomize individuals into higher versus lower variability groups. Week 0 values were chosen as these were closest to the patient’s known 30-day exacerbation-free (stable) period prior to inclusion in the sub-study and reflect symptom variability at a stable state. Student t test or Wilcoxon rank sum test was used for continuous variables and \( \chi^2 \) test or Fisher exact test was used for categorical variables for comparison of baseline patient characteristics between variability groups.

We explored differences in the number of participants experiencing HCRU-defined exacerbations over 5, 13, and 52 weeks in the higher versus lower variability group using Fisher exact test to compare proportions. We used multivariate logistic regression to explore the association between higher variability with an EXACT (symptom-defined) event during the 4-week post-baseline period. Statistical analysis was performed using R software version 3.6.3.

**Results**

**Characteristics of Included Patients**

The E-RS was completed daily by 205 participants; data from 140 (68%) met criteria for this analysis with sufficient diary completion rates in the first 5 weeks of follow-up (Figure 2). Included participants had a mean age of 63.3±8.4 years and 53% were women. Mean E-RS total score at baseline, representing respiratory symptom severity overall, was 11.4±6.6. Further participant characterization is found in Table 1. Differences between the included and excluded participants were fewer EXACT events in the excluded group (anticipated because lower completion rates would result in fewer detected events), lower modified Medical Research Council (mMRC) Dyspnea Scale scores, and lower St George’s Respiratory Questionnaire (SGRQ) scores (Table E1 in the online supplement). Our extended analysis of E-RS scores through week 13 included a smaller sample size of 130 patients, as fewer participants met diary completion criteria (Figure 2).

**Performance of Within Subject-Standard Deviation as a Symptom Variability Metric**

Reproducibility of the 6 different variability metrics between Week 0 and Week 4 were relatively low, with Pearson-correlation coefficients (r), ranging from 0.18–0.49 (Figure 3). By contrast, mean E-RS total scores for each participant were more constant \[ r=0.78; Cohen’s \text{ kappa}=0.66 \ (95\% \text{ CI: 0.53–0.78}) \] (Figure 4). Coefficient of variation \( r=0.49, 95\% \text{ confidence interval [CI]} 0.35–0.61 \), WS-SD \( r=0.32, 95\% \text{ CI } 0.16–0.46 \) and range \( r=0.34, 95\% \text{ CI } 0.19–0.48 \) had the highest Pearson correlation coefficients. Bland Altman plots were constructed for each of our variation measures and are shown in Figure 5.

**Figure 1. Schematic Illustrating Our Study Design**

<table>
<thead>
<tr>
<th>SPIROMICS enrollment</th>
<th>Exacerbation substudy enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 0</strong></td>
<td><strong>Week 1</strong></td>
</tr>
<tr>
<td>Baseline participant questionnaires and measurements collected</td>
<td>Baseline WS-SD calculated using E-RS</td>
</tr>
<tr>
<td><strong>Week 2</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Week 3</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Week 4</strong></td>
<td>Week 4 WS-SD calculated and compared to Week 0 (Baseline)</td>
</tr>
</tbody>
</table>

Baseline measurements and questionnaires were completed at SPIROMICS enrollment as part of the SPIROMICS protocol. A subgroup of participants entered the Exacerbation sub-study and completed the EXACT daily. E-RS scores are derived from the EXACT. Week 0 WS-SD is calculated and compared to week 4 values in our study.

SPIROMICS=SubPopulations and InterMEDIATE Outcome Measures in COPD Study, EXACT=Exacerbations of Chronic Obstructive Pulmonary Disease Tool, E-RS=Evaluating Respiratory Symptoms in COPD tool, WS-SD=within-subject standard deviation.
The median baseline (week 0) E-RS total WS-SD for all patients was 2.96. This value was used to dichotomize patients into lower variability versus higher variability groups for further study. There was poor agreement between Week 0 and Week 4 classifications, with a Cohen’s kappa of 0.27 (95% CI [0.12–0.42]). Between weeks 0 and 4, 36% of participants changed categories (Table 2) with the most common move from high variability at week 0 to low variability at week 4. Sensitivity analysis using a baseline period of 2 weeks rather than 1 demonstrated similar results.

We investigated overall trends in E-RS scores over the full 13-week study period, using mean E-RS total scores as a measure of overall symptom burden, and WS-SD as a measure of symptom variability. Mean WS-SD decreased by 11% (from 2.97 to 2.63) between Weeks 0 and 4 ($p=0.002$) (Figure E1 and E2 in the online supplement). This decrease attained statistical significance by Week 3 and was maintained throughout follow-up. When comparing Week 0 to Week 12, the mean WS-SD decreased 22%, to 2.23 ($p<0.001$) (Figure E1 and E2 in the online supplement). A large portion of the decrease occurred by Week 3, and almost no changes were detected between Weeks 6-12 (Table E2 in the online supplement).

Symptom Variability, Phenotypic Differences, and Exacerbations

Patients in the high variability group had higher (worse) SGRQ scores ($p=0.04$) and higher (more severe symptoms) E-RS scores ($p=0.001$) during the baseline week (Table 3). Other phenotypic characteristics such as 6-minute walk distance, Body mass index-airway Obstruction-Dyspnea-Exercise tolerance (BODE) index, imaging characteristics, and comorbid anxiety and depression did not separate low versus high variability groups.

Additionally, we examined HCRU exacerbations in the participants included in our analysis. Only 1 participant experienced an HCRU exacerbation in the first 5 weeks of the study, 7 experienced an HCRU exacerbation in 13 weeks, and 18 in 52 weeks (Table 4). Participants were again divided into lower and higher variability based on an overall week 0 median WS-SD of 2.96. There were no differences in HCRU exacerbations between the low and high variability group at 13 weeks (5.7% versus 4.3%, $p$-value 1.0) and 52 weeks (10.0% versus 15.7%, $p$-value 0.45). In a sensitivity analysis, we also found that the mean baseline WS-SD was not different between those participants who experienced an
Table 1. Characteristics of Included Participants at Baseline

<table>
<thead>
<tr>
<th>Characteristica</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>140</td>
</tr>
<tr>
<td>Age (years)</td>
<td>63.3±8.4</td>
</tr>
<tr>
<td>Female Gender</td>
<td>74 [52.9%]</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.4±5.5</td>
</tr>
<tr>
<td>6-minute walk distance (m)</td>
<td>362±123</td>
</tr>
<tr>
<td>FEV₁% predicted</td>
<td>61.5±21.9</td>
</tr>
<tr>
<td>BODE Index</td>
<td>2.3±2.2</td>
</tr>
<tr>
<td>mMRC Dyspnea Scale</td>
<td>1.4±1.1</td>
</tr>
<tr>
<td>PSQ Total Score</td>
<td>6.9±3.8</td>
</tr>
<tr>
<td>SGRQ Total Score</td>
<td>42.9±19.9</td>
</tr>
<tr>
<td>HADS Anxiety Score</td>
<td>6.2±4.0</td>
</tr>
<tr>
<td>HADS Depression Score</td>
<td>4.9±3.7</td>
</tr>
<tr>
<td>Current Smoker</td>
<td>47 [33.8%]</td>
</tr>
<tr>
<td>PRM Emphysema</td>
<td>6 [1,13]</td>
</tr>
<tr>
<td>PRM Functional Small Airways Disease</td>
<td>25 [15,36]</td>
</tr>
<tr>
<td>P10</td>
<td>3.69 [3,6,6,3,75]</td>
</tr>
</tbody>
</table>

GOLD Stage

- Stage 1: 11 [7.9%]
- Stage 2: 31 [22.3%]
- Stage 3: 67 [48.2%]
- Stage 4: 30 [21.6%]

EXACT Mean: 35.2±11.0

E-RS Mean: 11.4±6.6

E-RS Subscales

- E-RS breathlessness: 5.6±3.8
- E-RS cough/sputum: 3.2±1.9
- E-RS chest: 2.6±2.0

EXACT Event Within 4 weeks of Follow-up: 45 [32.1%]

aData are presented as n[%], mean±SD or median [IQR].

BMIf=body mass index; FEV₁=forced expiratory volume in 1 second; BODE=Body mass index airflow Obstruction-Dyspnea-Exercise capacity; mMRC=modified Medical Research Council; PSQ=Pittsburgh Sleep Quality Index; SGRQ=St George’s Respiratory Questionnaire; HADS=Hospital anxiety and depression scale; PRM=parametric response mapping; P10=square root of wall area of a hypothetical airway with an internal perimeter of 10 mm; GOLD=Global initiative for chronic Obstructive Lung Disease; EXACT=Exacerbations of Chronic Pulmonary Disease Tool; E-RS= Evaluating Respiratory Symptoms in COPD Tool

HCRU versus those who did not (3.6 versus 3.0, p-value 0.44).

Higher baseline WS-SD was associated with an increased likelihood of an EXACT (symptom-defined) event (O =2.52; p=0.02) during the subsequent 4 weeks. EXACT events from 2 participants are illustrated in Figure E3 in the online supplement. Even after adjusting for symptom severity (using mean baseline E-RS total score) and other potential confounders (sex, dyspnea score, 6-minute walk distance, and forced expiratory volume in 1 second [FEV₁]), participants classified as having high baseline symptom variability had 2.75 greater odds of experiencing an EXACT event during the 4 weeks than individuals with low symptom variability (OR=2.75 [1.13–6.86], p=0.02).

**Discussion**

We used a novel approach for characterizing symptom variability in patients with COPD by using prospectively collected E-RS data from participants in the SPIROMICS Exacerbations sub-study. Our analyses demonstrate that: (1) the WS-SD of the E-RS total score performed similarly to other variability measures in that baseline values are only weakly associated with variability at week 4, (2) mean WS-SD scores decreased significantly throughout a 12-week follow-up period, primarily during the first half of follow-up, and (3) participants with higher baseline variability have increased symptom severity, although they do not demonstrate differences in other phenotypic markers compared to low variability participants. These findings provide novel insights into the patient experience in COPD.

Day-to-day symptom variability is itself variable over time, as evidenced by relatively low correlations between weeks 0 and 4 in 6 different E-RS score variability metrics. Coefficient of variation has the highest correlation between week 0 and week 4 values and could be used as a metric by which to compare the magnitude of symptom variability to variation in other measures. The Bland-Altman plots for all variability metrics have a similar shape, with better agreement among patients with overall lower variability compared to participants with larger differences between week 0 and week 4 values. The changes to symptom variation within patients over time calls into question one component of the traditional HCRU exacerbation definition, namely what constitutes a given patient’s typical day-to-day symptom variation and what aspect of worsening informs a patient’s decision to seek care. Importantly, the fluctuation in symptoms we show may have important implications when evaluating a response to therapy, especially in the absence of a control group as in n-of-1 experiments. The median E-RS total score WS-SD was 2.96 in our study. To
Figure 3. Panels A-F Each Represent a Different Measure of Variability Considered

A. Standard deviation

Corr = 0.32; 95% CI (0.16, 0.46)

B. Coefficient of variation

Corr = 0.49; 95% CI (0.35, 0.61)

C. Range

Corr = 0.34; 95% CI (0.19, 0.48)

D. Variance

Corr = 0.25; 95% CI (0.09, 0.40)

E. Median absolute deviation

Corr = 0.27; 95% CI (0.11, 0.42)

F. Interquartile range

Corr = 0.18; 95% CI (0.02, 0.34)

The x-axis is the week 0 value for each measure and the y-axis is the week 4 value. The Pearson’s correlation coefficient and 95% confidence interval for each variability measure is shown on the corresponding graph.

Corr=Pearson’s correlation coefficient; CI=confidence interval; EXACT=Exacerbations of Chronic Pulmonary Disease Tool
contextualize these findings, thresholds for symptomatic improvement using the E-RS tool were developed with clinical trial data examining subgroups of patients who demonstrated improvement using established measures including the SGRQ and the Breathlessness, Cough and Sputum Scale as anchors. An overall decline in E-RS score by 2 points has been proposed as the threshold to define symptomatic improvement.\(^\text{15}\) With a median WS-SD of 2.96 in our study, we show that patients can exhibit day-to-day variation higher than proposed thresholds of symptomatic change meaningful to patients. This makes the longitudinal analysis of E-RS daily scores even more important, with persistent improvements in severity of 2 points or more needed to demonstrate improvement over time, taking into consideration day-to-day variability.

Interestingly, throughout this 5-week study period, E-RS total scores remained stable over time, while WS-SD decreased significantly, suggesting symptom severity overall is stable, while day-to-day variability is not. The stability of mean E-RS scores as a measure of symptom severity has been demonstrated in the past.\(^\text{24}\) Participants who were in the high variability group at week 0 were not always in the same variability group at week 4. These findings should be interpreted in the context of the fact that participants were only known to be stable and exacerbation free in the weeks prior to study entry. It is possible that an intervening change in therapy during a routine clinic visit or an HCRU event resulted in less variability with no effect on overall severity, though this is unlikely. It is also possibly that participant complacency or “settling on an average” occurs over time, to facilitate diary completion. The Hawthorne effect, where the knowledge that one is being observed leads to a change in behavior, could result in increased introspection and impact variability as the study progresses. Periodic reminders to patients to complete the diary thoughtfully and accurately could address this potential issue. It is noteworthy that even with the flattening of variability scores over time, mean symptom severity levels were stable and EXACT events were detected during the latter weeks of the study, with baseline variability a predictor of these events.

Our analysis suggests that individual symptom variability may be associated with increased morbidity. Patients with higher variability also had higher mean E-RS scores, i.e., more severe respiratory symptoms. Elevated SGRQ, a marker of exacerbation risk,\(^\text{25}\) was also found in the higher symptom variability group. Notably, we were unable to distinguish individuals with higher versus lower symptom variability using other phenotypic markers, suggesting that symptom variability may be a construct that is different from other physiological or imaging metrics.

While the small number of HCRU exacerbations limited our ability to determine an association between symptom variability and exacerbation using the traditional definition, we were able to explore an association between baseline day-to-day variation in symptoms and risk of developing periods of sustained symptom worsening (EXACT events). Should these findings be confirmed in larger samples, symptom variability could be a promising component of comprehensive patient assessment.

Our cohort had a higher rate of EXACT events compared to what has been previously reported in the literature.\(^\text{26,27}\) We hypothesize that the principal reason for this is due to differences in the patient population. The majority of longitudinal EXACT data have been
Figure 5. Bland Altman Plot

The Y axis is the difference between week 0 versus week 4 for each variability metric. The x axis is the mean of the week 0 value and the week 4 value. For example, panel A is the Bland Altman plot for standard deviation (WS-SD). The Y axis is the difference between the week 0 and week 4 standard deviation. The X axis is the mean of the week 0 and week 4 standard deviations in E-RS scores for each participant.

WS-SD=within-subject standard deviation; E-RS=Evaluating Respiratory Symptoms in COPD tool
reported in the context of placebo-controlled trials in patients meeting strict inclusion and exclusion criteria. Our patient population reflects a different opportunistic cohort with limited inclusion/exclusion criteria, perhaps closer to real-world data. Differences in approach to data collection also may influence the number of exact events. For example, in the FLAME randomized control trial, participants completed an electronic daily diary in addition to the daily exact. Diary responses in turn triggered the participant to respond to a clinical center, which may have led to more HCRU events relative to the number of unreported exact events.

Our study has limitations, including a relatively small number of participants and, for WS-SD stability, a short study period of 5 weeks. Additionally, we did not have data regarding changes to individual treatment between week 0 and week 4 which may have affected symptom variability at week 4. A larger dataset over an extended period may allow for better characterization of longitudinal trends in variability. Requiring a high completion rate of the E-RS/EXACT diary may have resulted in selection bias in this natural history study, with only the most adherent participants included. We also did not have daily criteria measures to assist with the evaluation of stability, such as global assessments, alternative symptom diaries, or activity monitors.

In summary, ostensibly stable COPD individuals experience wide temporal variability in respiratory symptoms. Using common statistical techniques such as WS-SD to quantify day-to-day variability can lead to further insight into the association between symptom fluctuation, symptom severity, and future exacerbations.

### Table 2. Classification of Participants at Week 0 and at Week 4 into Variability Categories Using Week 0 Median Within-Subject Standard Deviation as Cutoff

<table>
<thead>
<tr>
<th>Baseline (week 0)</th>
<th>Lower Variability WS-SD&lt;2.96 (n)</th>
<th>Higher Variability WS-SD≥2.96 (n)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Variability WS-SD&lt;2.96</td>
<td>56</td>
<td>14</td>
<td>70</td>
</tr>
<tr>
<td>High Variability WS-SD≥2.96</td>
<td>37</td>
<td>33</td>
<td>70</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>47</td>
<td>140</td>
</tr>
</tbody>
</table>

aMedian WS-SD=2.96

WS-SD=within-subject standard deviation

### Table 3. Comparison of Baseline Characteristics When Participants are Dichotomized Into Low Variability Versus High Variability Using the Baseline Week 0 Median Within-Subject Standard Deviation

<table>
<thead>
<tr>
<th>Characteristicb</th>
<th>Lower Variability WS-SD&lt;2.96</th>
<th>Higher Variability WS-SD≥2.96</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>70</td>
<td>70</td>
<td>0.13</td>
</tr>
<tr>
<td>Age, years</td>
<td>64.4±7.8</td>
<td>62.3±8.9</td>
<td>0.61</td>
</tr>
<tr>
<td>Female sex,</td>
<td>35 [50%]</td>
<td>39 [56%]</td>
<td>0.97</td>
</tr>
<tr>
<td>BMI</td>
<td>27.4±5.0</td>
<td>27.5±6.0</td>
<td>0.42</td>
</tr>
<tr>
<td>6-Minute Walk Distance, m</td>
<td>353.2±121.5</td>
<td>370.6±123.9</td>
<td>0.20</td>
</tr>
<tr>
<td>FEV1/% predicted</td>
<td>59.1±22.2</td>
<td>63.9±21.5</td>
<td>0.77</td>
</tr>
<tr>
<td>BODE Index</td>
<td>2.4±2.3</td>
<td>2.2±2.0</td>
<td>0.04</td>
</tr>
<tr>
<td>SGRQ Total Score</td>
<td>39.6±21.5</td>
<td>47.3±20.3</td>
<td>0.52</td>
</tr>
<tr>
<td>HADS Anxiety Score</td>
<td>5.9±4.0</td>
<td>6.4±4.0</td>
<td>0.49</td>
</tr>
<tr>
<td>HADS Depression score</td>
<td>4.7±3.5</td>
<td>5.1±3.6</td>
<td>0.13</td>
</tr>
<tr>
<td>PRM Emphysema (median [IQR])</td>
<td>7 [1,16]</td>
<td>5 [1,11]</td>
<td>0.30</td>
</tr>
<tr>
<td>PRM Functional Small Airways Disease (median [IQR])</td>
<td>28 [16,34]</td>
<td>24 [12,48]</td>
<td>0.20</td>
</tr>
<tr>
<td>Pi10</td>
<td>3.69 [3.66,3.75]</td>
<td>3.72 [3.67,3.75]</td>
<td>0.001</td>
</tr>
<tr>
<td>Average Baseline EXACT (week 0, mean±SD)</td>
<td>31.9±11.9</td>
<td>38.5±8.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average Baseline E-RS (week 0, mean±SD)</td>
<td>9.5±6.9</td>
<td>13.3±5.8</td>
<td>0.001</td>
</tr>
</tbody>
</table>

aMedian=2.96

bData are presented as n[%], mean±SD or median [IQR]

WS-SD=within-subject standard deviation; BMI=body mass index; FEV1=forced expiratory volume in 1 second; BODE=Body mass index-airflow Obstruction-Dyspnea-Exercise capacity; SGRQ=St George’s Respiratory Questionnaire; HADS=Hospital anxiety and depression scale; PRM=parametric response mapping; IQR=interquartile range; Pi10=square root of wall area of a hypothetical airway with an internal perimeter of 10 mm; EXACT=Exacerbations of Chronic Pulmonary DiseaseTool; E-RS=Evaluating Respiratory Symptoms in COPD Tool

Determining the utility of the E-RS variability metric to quantify day-to-day symptom variability in COPD will require study of more individuals over a longer follow-up period and using additional modes of symptom measurement.

### Acknowledgements

**Author Contributions:** KMA, JKK, CO, and FJM had full access to all of the data in the study and take
responsibility for the integrity of the data and accuracy of the analysis. KMA, JKK, CO, MKH, JAW, and FJM contributed to the conception and design of the study. KMA, JKK, CO, IE, KVB, MKH, and FJM contributed to data analysis and interpretation. EEC, SAC, DJC, GJC, JLC, MTD, MKH, NNI, RP, RPB, SPP, PGW, and FJM contributed to the acquisition of the data. KMA and JKK contributed to the drafting of the manuscript. KMA, JKK, CO, NKL, RPB, SAC, DJC, GJC, JLC, MTD, MKH, NNI, ASI, RP, SPP, JAW, PGW, KVB, and FJM contributed to revisions for critically important intellectual content. All of the authors approved this version of the manuscript to be published.

Availability of data: More information about the study and how to access SPIROMICS data is at www.spiromics.org.

The authors thank the SPIROMICS participants and participating physicians, investigators, and staff for making this research possible. We would like to acknowledge the following current and former investigators of the SPIROMICS sites and reading centers: Neil E. Alexis, MD; Wayne H. Anderson, PhD; Mehrdad Arjomandi, MD; Igor Barjaktarevic, MD, PhD; R. Graham Barr, MD, DrPH; Lori A. Bateman, MSc; Surya P. Bhatt, MD; Eugene R. Blecker, MD; Richard C. Boucher, MD; Russell P. Bowler, MD, PhD; Stephanie A. Christenson, MD; Alejandro P. Comellas, MD; Christopher B. Cooper, MD, PhD; David J. Couper, PhD; Gerard J. Criner, MD; Ronald G. Crystal, MD; Jeffrey L. Curtis, MD; Claire M. Doerschuk, MD; Mark T. Dransfield, MD; Brad Drummond, MD; Christine M. Freeman, PhD; Craig Galban, PhD; Meilan K. Han, MD, MS; Nadia N. Hansel, MD, MPH; Annette T. Hastie, PhD; Eric A. Hoffman, PhD; Yvonne Huang, MD; Robert J. Kaner, MD; Richard E. Kanner, MD; Eric C. Kleerup, MD; Jerry A. Krishnan, MD, PhD; Lisa M. LaVange, PhD; Stephen C. Lazarus, MD; Fernando J. Martinez, MD, MS; Deborah A. Meyers, PhD; Wendy C. Moore, MD; John D. Newell, Jr, MD; Robert Paine, III, MD; Laura Paulin, MD, MHS; Stephen P. Peters, MD, PhD; Cheryl Pirozzi, MD, NIRUPAMA PATHA, MD, MHS; Elizabeth C. Oelsner, MD, MPH; Wanda K. O'Neal, PhD; Victor E. Ortega, MD, PhD; Sanjeev Raman, MBBS, MD; Stephen I. Rennard, MD; Donald P. Tashkin, MD; J. Michael Wells, MD; Robert A. Wise, MD; and Prescott G. Woodruff, MD, MPH. The project officers from the Lung Division of the National Heart, Lung, and Blood Institute were Lisa Postow, PhD and Lisa Viviano, BSN.

Table 4. Comparison of the Number of Health Care Resource Utilization Events When Participants are Dichotomized into Low Variability versus High Variability Using the Baseline Week 0 Median Within-Subject Standard Deviation*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Lower Variability WS-SD&lt;2.96</th>
<th>Higher Variability WS-SD≥2.96</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>HCRU event in 5 weeks</td>
<td>1 [1.4%]</td>
<td>0 [0.0%]</td>
</tr>
<tr>
<td>HCRU event in 13 weeks</td>
<td>4 [5.71%]</td>
<td>3 [4.3%]</td>
</tr>
<tr>
<td>HCRU event in 52 weeks</td>
<td>7 [10.0%]</td>
<td>11 [15.7%]</td>
</tr>
</tbody>
</table>

*WS-SD=within-subject standard deviation; HCRU=health care resource utilization.

Declaration of Interest

JKK reports grants from NIH T32 HL134629, during the conduct of the study; non-financial support from GlaxoSmithKline, outside the submitted work. NKL reports other from Evidera, outside the submitted work. MKH reports personal fees from GlaxoSmithKline, AstraZeneca, Boehringer Ingelheim, Cipla, Chiesi, Novartis, Pulmonx, Teva, Verona, Merck, Mylan, Sanofi, DevPro, Aerogen, Polarian, Regeneron, United Therapeutics, UpToDate, Medscape, and Integrity. She has received either in-kind research support or funds paid to the institution from the NIH, Novartis, Sunovion, Nuvaria, Sanofi, AstraZeneca, Boehringer Ingelheim, Gala Therapeutics, Bodesis, the COPD Foundation, and the American Lung Association. She has participated in Data Safety Monitoring Boards for Novartis and Medtronic with funds paid to the institution. She has received stock options from Meissa Vaccines. RP reports grants from the NHLBI, grants from the COPD Foundation during the conduct of this study; grants from Department of Veterans Affairs, outside the submitted work. SAC reports personal fees from AstraZeneca, personal fees from GlaxoSmithKline, personal fees from Amgen, personal fees from Glenmark, personal fees from Sunovion, personal fees from Genentech, personal fees from UpToDate, outside the submitted work. DJC reports grants from the NHLBI, grants from the COPD Foundation, during the conduct of the study. GJC reports grants and personal fees from GlaxoSmithKline, grants...
and personal fees from Boehringer Ingelheim, grants and personal fees from Chiesi, grants and personal fees from Merco, personal fees from Verona, grants and personal fees from AstraZeneca, grants and personal fees from Pulmonox, grants and personal fees from Pneumirx, personal fees from BTG, grants and personal fees from Olympus, grants and personal fees from Broncus, personal fees from EOLO, personal fees from NGM, grants and personal fees from Lungpacer, grants and personal fees from Nuivaira, grants and personal fees from ResMed, grants and personal fees from Respironics, grants and personal fees from Fisher Paykel, grants and personal fees from Patara, grants from Galapagos, outside the submitted work. JLC reports a grant from the NIH/NHLBI, during the conduct of the study; grants from the Department of Veterans Affairs and the Department of Defense, personal fees and non-financial support from AstraZeneca, and personal fees from Novartis, outside the submitted work. MTD reports grants from the NIH, during the conduct of the study; personal fees and other from Boehringer Ingelheim, personal fees and other from GlaxoSmithKline, other from Novartis, personal fees and other from AstraZeneca, other from Yungjin, personal fees and other from PneumRx/ BTG, non-financial support and other from Pulmonox, from Boston Scientific, personal fees from Quark Pharmaceuticals, personal fees from Mereo, grants from American Lung Association, grants from the NIH, grants from the Department of Veterans Affairs, other from Gala, other from Nuivaira, grants from the Department of Defense, outside the submitted work. NNH reports grants from the NIH, personal fees from Nuvaira, personal fees and non-financial support from the Department of Veterans Affairs, outside the submitted work. CO, IE, and RPB report have no conflicts to disclose.
References


