# Psychometric Validation and Meaningful Within-Patient Change (MWPC) Thresholds for the Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ) Using Data from the EVOKE-01 Study

PCR54

Mira Patel[1], Matthew Radford[1], Sabeen Mekan[1], Shien Guo[2], Christopher G Pelligra[3]

[1]Gilead Sciences, Inc, Foster City, CA, USA; [2]Evidera, Inc, Waltham, MA, USA; [3]Evidera, Inc, Atlanta, GA, USA



## Conclusions

- Both the shortness of breath (SOB) domain and total score (TS) of the Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ) demonstrated robust psychometric properties in a sample of patients with metastatic NSCLC.
- The meaningful within-patient change (MWPC) thresholds for improvement and deterioration established for the two measures will help define “responders” and/or guide interpretation of treatment effects when these measures are used to define clinical trial endpoints.
- The results support that the NSCLC-SAQ SOB domain and TS are fit-for-purpose for assessing treatment effects in metastatic NSCLC.

## Plain Language Summary

- The NSCLC-SAQ is a questionnaire that is completed by patients with lung cancer to report the severity of their symptoms of cough, pain, SOB, fatigue, and poor appetite.
- This analysis assessed the measurement properties (reliability, validity, and responsiveness) of the NSCLC-SAQ SOB domain and TS and determined the thresholds of improvement or worsening on these measures that can be considered meaningful when assessing lung cancer symptoms in patients enrolled in the EVOKE-01 trial.
- The results showed good measurement properties of the SOB domain and TS for assessing lung cancer symptoms.
- A change of 1 point in the SOB domain and a range of 2–4 points in TS were suggested to indicate meaningful changes in lung cancer symptoms.
- Overall, the findings support that the NSCLC-SAQ SOB domain and TS are suitable for assessing symptom severity in patients with metastatic NSCLC.

## Introduction

- The NSCLC-SAQ is a patient-reported outcome (PRO) instrument that was granted qualification by the FDA for exploratory use to measure overall symptom severity of NSCLC in drug development.[1]
  - The NSCLC-SAQ comprises seven items covering five symptom concepts of NSCLC, including cough, pain, SOB, fatigue, and appetite, with a TS reflecting the sum of all domains.
- The NSCLC-SAQ was used to assess treatment effect on NSCLC symptoms in the phase 3, open-label EVOKE-01 trial (NCT05089734) comparing sacituzumab govitecan with docetaxel in patients with metastatic NSCLC who had progressed on platinum-based chemotherapy and programmed death-(ligand)1 (PD-[L]1) inhibitors.[2]
- This analysis was conducted to provide further evidence to support the psychometric properties and the interpretation of MWPC of the measure.[1]

## Objectives

- To assess the psychometric properties and derive MWPC thresholds for the NSCLC-SAQ SOB domain and TS in the context of EVOKE-01.

## Methods

- Eligible patients in EVOKE-01 had metastatic NSCLC with progression on platinum-based chemotherapy and PD-(L)1 inhibitors; interim blinded datasets including all randomized patients were used in this analysis.
- PROs and the Eastern Cooperative Oncology Group performance status (ECOG PS) were assessed before dosing on day 1 of every 3-week cycle.
  - The NSCLC-SAQ SOB domain uses a 5-point scale ranging from 0 (never) to 4 (always), and the TS ranges from 0 to 20, with a higher score indicating more severe symptomatology.
    - TS was not computed if any domain score was missing.
  - Other PRO measures included: EQ-5D-3L, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30 (EORTC QLQ-C30), and single-item Patient Global Impression of Severity (PGIS) and Patient Global Impression of Change (PGIC).
    - Response options for the PGIS (“Which one response best describes the overall severity of your lung cancer symptoms over the past 7 days”) are: “None,” “Mild,” “Moderate,” “Severe,” and “Very severe.”
    - Response options for the PGIC (“Which one response best describes the overall change in your lung cancer symptoms since the start of this study”) are: “Very much improved,” “Much improved,” “Minimally improved,” “No change,” “Minimally worsened,” “Much worsened,” and “Very much worsened.”

### Reliability

- Test-retest reliability for NSCLC-SAQ SOB domain and TS was assessed in patients who had the same response on the PGIS at weeks 10 and 13.
  - Intraclass correlation coefficients (ICCs) were calculated using a two-way, mixed-effect analysis of variance model.
  - An ICC of ≥0.70 was regarded as acceptable.[3]
- For NSCLC-SAQ TS, internal consistency reliability was assessed by calculating Cronbach’s alpha using pooled data from baseline and week 13, with a Cronbach’s alpha of ≥0.70 indicating acceptable internal consistency reliability.[4]

### Validity

- Convergent validity was assessed by computing the Spearman’s rank correlation coefficient (r) for NSCLC-SAQ SOB domain and TS, using pooled data from baseline and week 13, with each of the following PROs: EQ-5D visual analog scale (EQ-VAS), PGIS, and EORTC QLQ-C30 global health status/quality of life, physical functioning, fatigue, and dyspnea.
  - Correlations were classified as weak (|r|<0.3), moderate (0.30≤|r|<0.70), strong (0.70≤|r|<0.90), or very strong (|r|≥0.9).[5]
  - Moderate-to-strong correlations were expected.
- Known-groups validity was assessed by comparing NSCLC-SAQ SOB domain score and TS using pooled data from baseline and week 13 among patient groups classified by their PGIS scores and ECOG PS.
  - For the SOB domain, known-groups validity was also assessed based on the response (not at all [score=1] to very much [score=4]) to EORTC QLQ-C30 Item 8 (“During the past week: Were you short of breath?”).

### Responsiveness

- Responsiveness was evaluated by first calculating Spearman’s correlation coefficients for changes from baseline (pooled across all post-baseline visits) in NSCLC-SAQ SOB domain score and TS with changes in the PGIS, PGIC, ECOG PS, and EORTC QLQ-C30 Items 8, 29 (“How would you rate your overall health during the past week?”), and 30 (“How would you rate your overall quality of life during the past week?”).
  - Outcomes with a correlation of ≥0.30 were selected as anchors.
  - ANCOVA was used to compare changes in NSCLC-SAQ SOB score and TS between different response groups for the selected anchors.

### MWPC

- MWPC thresholds for improvement and deterioration were estimated using an anchor-based approach, supported by a distribution-based approach.
  - Estimates of ±1 standard error of measurement (SEM), 0.2× standard deviation (SD), 0.5×SD, and 0.8×SD of baseline scores were used to support the selection of MWPC thresholds.

## Results

- Data from 556 randomized patients were included in the analyses.
- The mean age of patients was 64 years; most patients were male (68%) and White (76%).
- At baseline, most patients had an ECOG PS of 1 (63%) and had experienced progressive or stable disease with their last anti-PD-(L)1 therapy (68%).
- Completion rates (percentages of patients expected to complete an assessment at a given visit) for the NSCLC-SAQ SOB domain and TS ranged from 76% to 100% across visits.

### Test-retest reliability and internal consistency

- Test-retest reliability was acceptable for the SOB domain (ICC 0.80 [95% CI 0.71–0.87]) and TS (0.81 [95% CI 0.72–0.87]).
- Internal consistency reliability was demonstrated for NSCLC-SAQ TS (standardized Cronbach’s alpha coefficient 0.74).

### Convergent validity

- Convergent validity for the NSCLC-SAQ SOB domain and TS was consistent with expectations, showing moderate or strong correlations (|r|>0.30) with external PROs measuring similar concepts (**Table 1**).

**Table 1. Correlations of NSCLC-SAQ SOB domain and TS with other measures**

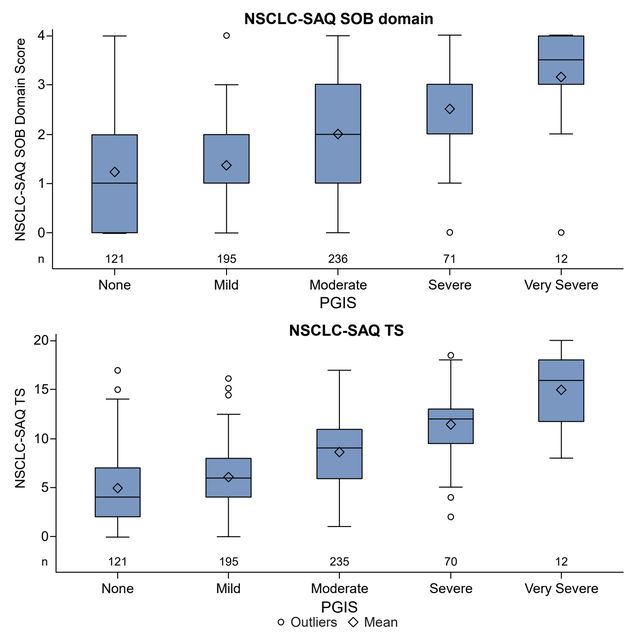
| | NSCLC-SAQ SOB domain | | NSCLC-SAQ TS | |
|---|---|---|---|---|
| | n | Correlation (r)[a] | n | Correlation (r)[a] |
| PGIS | 635 | 0.41 | 633 | 0.55 |
| EQ-VAS | 644 | −0.35 | 642 | −0.50 |
| EORTC QLQ-C30 | | | | |
| GHS/QoL | 652 | −0.35 | 650 | −0.56 |
| Physical functioning | 658 | −0.51 | 656 | −0.61 |
| Fatigue | 659 | 0.49 | 657 | 0.69 |
| Dyspnea | 658 | 0.73 | 656 | 0.56 |

P<0.001 for all correlations. Analyses were conducted using pooled data from baseline and week 13. EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30; EQ-VAS, EQ-5D visual analog scale; GHS, global health status; NSCLC-SAQ, Non-small Cell Lung Cancer Symptom Assessment Questionnaire; PGIS, Patient Global Impression of Severity; QoL, quality of life; SOB, shortness of breath; TS, total score. [a]Spearman’s rank correlation.

### Known-groups validity

- Both NSCLC-SAQ SOB domain and TS were able to distinguish between severity groups according to PGIS score and ECOG PS.
  - SOB domain score and TS were higher in patients with higher PGIS scores (**Figure 1**) and ECOG PS (not shown).
  - Known-groups validity for the SOB domain was also demonstrated based on EORTC QLQ-C30 Item 8.

**Figure 1. Known-groups validity based on PGIS score**




Analyses were conducted using pooled data from baseline and week 13. All P-values from F-test and Kruskal-Wallis test were <0.001. NSCLC-SAQ, Non-small Cell Lung Cancer Symptom Assessment Questionnaire; PGIS, Patient Global Impression of Severity; SOB, shortness of breath; TS, total score.

### Responsiveness

- External anchors with adequate correlation for responsiveness (|r|≥0.30) (**Table 2**) and thus chosen as anchors for anchor-based analyses of MWPC thresholds included:
  - EORTC QLQ-C30 Item 8 for NSCLC-SAQ SOB domain
  - EORTC QLQ-C30 Item 8, PGIS, and PGIC for NSCLC-SAQ TS

**Table 2. Correlations of changes from baseline in NSCLC-SAQ SOB domain and TS with changes on other measures**

| Anchors | NSCLC-SAQ SOB domain | | NSCLC-SAQ TS | |
|---|---|---|---|---|
| | n | Correlation (r)[a] | n | Correlation (r)[a] |
| EORTC QLQ-C30 Item 8 | 1358 | **0.48** | 1342 | **0.39** |
| PGIS | 1286 | 0.16 | 1270 | **0.31** |
| PGIC[b] | 1343 | 0.19 | 1327 | **0.31** |

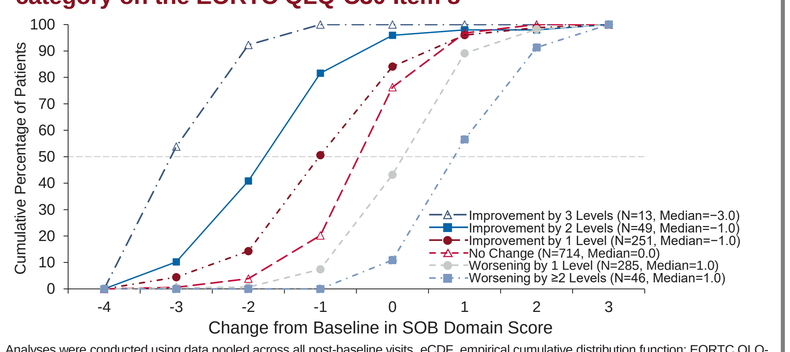
Coefficients ≥0.30 are bolded. Analyses were conducted using data pooled across all post-baseline visits. EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30; NSCLC-SAQ, Non-small Cell Lung Cancer Symptom Assessment Questionnaire; PGIC, Patient Global Impression of Change; PGIS, Patient Global Impression of Severity; SOB, shortness of breath; TS, total score. [a]Spearman’s rank correlation. [b]Absolute values were used as the PGIC measures change directly.

- Although EORTC QLQ-C30 Item 29 was moderately correlated with NSCLC-SAQ TS (r=0.35), it was excluded from anchor-based analyses because:
  - The PGIS and PGIC were FDA-preferred anchors.[6]
  - EORTC QLQ-C30 Item 8 had an easier-to-interpret rating scale and a slightly higher correlation.
- In responsiveness ANCOVA analyses, NSCLC-SAQ SOB domain score and TS decreased (improved) more in patients with a higher level of improvement classified by the selected anchors.

### Triangulation and proposed MWPC thresholds

- Clear separations were observed for changes from baseline in SOB domain score (**Figure 2**) and TS (not shown) by response category on the EORTC QLQ-C30 Item 8.

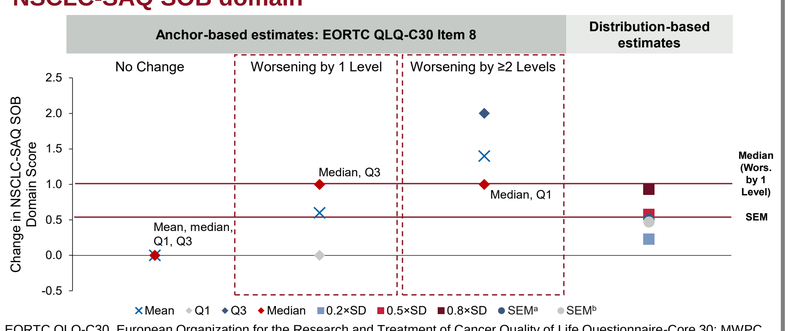
**Figure 2. eCDF curves for SOB domain score changes by response category on the EORTC QLQ-C30 Item 8**




Analyses were conducted using data pooled across all post-baseline visits. eCDF, empirical cumulative distribution function; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire–Core 30; SOB, shortness of breath.

- An MWPC threshold of ≥+1 point was proposed for deterioration in the NSCLC-SAQ SOB domain because (**Figure 3**):
  - Median change from baseline was +1.0 for groups with a worsening of one or two levels on EORTC QLQ-C30 Item 8.
  - Minimum possible change in score is ±1 for an individual.
  - A 1-point worsening exceeded the SEM estimates (0.47–0.51), as well as 0.5×SD (0.58) and 0.8×SD (0.93).
- An MWPC threshold of ≤−1 point was proposed for improvement in the NSCLC-SAQ SOB domain for similar reasons.

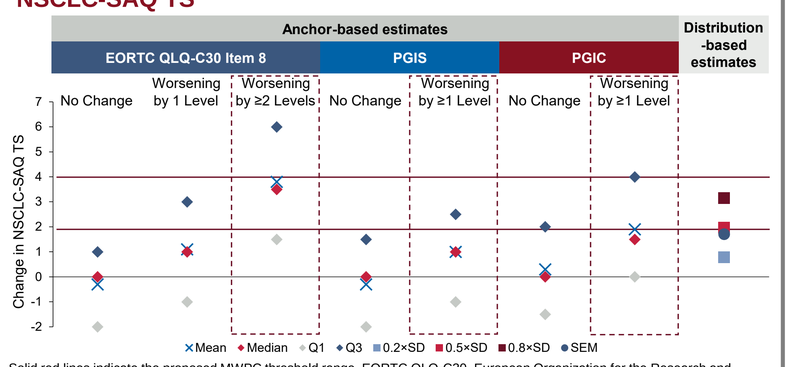
**Figure 3.Triangulation of MWPC thresholds for deterioration – NSCLC-SAQ SOB domain**




EORTC QLQ-C30, European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; MWPC, meaningful within-patient change; NSCLC-SAQ, Non-small Cell Lung Cancer Symptom Assessment Questionnaire; PGIS, Patient Global Impression of Severity; Q1, first quartile; Q3, third quartile; SD, standard deviation; SEM, standard error of measurement; SOB, shortness of breath; Wors., worsening. [a]SEM estimated based on patients stable on the PGIS between weeks 10 and 13. [b]SEM estimated based on patients stable on EORTC QLQ-C30 Item 8 between weeks 10 and 13.

- An MWPC threshold range of +2 to +4 points was proposed for deterioration in NSCLC-SAQ TS because (**Figure 4**):
  - SEM was 1.71, suggesting a threshold of ≥+2.0.
  - Median change for the group with a worsening of ≥2 levels on EORTC QLQ-C30 Item 8 was 3.5.
  - Median changes for groups with a worsening of ≥1 levels on the PGIS and PGIC were 1.0 and 1.5, respectively.
  - A change of −2.0 and −4.0 exceeded 0.5×SD (1.97) and 0.8×SD (3.15), respectively.
- An MWPC threshold range of −4 to −2 was proposed for improvement in NSCLC-SAQ TS for similar reasons.

**Figure 4.Triangulation of MWPC thresholds for deterioration – NSCLC-SAQ TS**




Solid red lines indicate the proposed MWPC threshold range. EORTC QLQ-C30, European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; MWPC, meaningful within-patient change; NSCLC-SAQ, Non-small Cell Lung Cancer Symptom Assessment Questionnaire; PGIC, Patient Global Impression of Change; PGIS, Patient Global Impression of Severity; Q1, 1st quartile; Q3, 3rd quartile; SD, standard deviation; SEM, standard error of measurement; TS, total score.

**References:** **1.** FDA. Qualification of non-small cell lung cancer symptom assessment questionnaire (NSCLC-SAQ): a patient-reported outcome instrument (DDT COA #00009). April 4, 2018. **2.** Paz-Ares LG, et al. *J Clin Oncol*. 2024; 42:2860-2872. **3.** Cappelleri JC. Patient-reported outcomes: measurement, implementation and interpretation. CRC Press/Taylor & Francis Group; 2014. **4.** Aaronson N, et al. *Qual Life Res*. 2002;11:193-205. **5.** Hinkle DE, et al. Applied statistics for the behavioral sciences. Vol 663. Houghton Mifflin Boston; 2003. **6.** FDA. Patient-focused drug development: incorporating clinical outcome assessments into endpoints for regulatory decision-making (FDA-2023-D-0026). April, 2023.

**Acknowledgments:** We extend our thanks to the patients, their families, and all participating investigators. This study was funded by Gilead Sciences, Inc. Medical writing assistance was provided by Hui Zhang, PhD, of PPD, a Thermo Fisher business, in accordance with Good Publication Practice (GPP 2022) guidelines and was funded by Gilead Sciences, Inc. Graphic design support were provided by Kawthar Nakayima and Karissa Calara of Evidera, a business unit of PPD, part of Thermo Fisher Scientific.

**Disclosures:** MP, MR, and SM are employees of Gilead Sciences, Inc. SG and CGP are employees of Evidera, which received funding from Gilead for the conduct of the study.

**Correspondence:** Mira J Patel, Mira.Patel28@gilead.com