Thorac Res Pract. 2024; 25(6): 215-220 DOI: 10.5152/ThoracResPract.2024.23135

Original Article

The Validity and Reliability of the Turkish version of Modified Medical Research Council Dyspnea Scale in Systemic Sclerosis Patients with Interstitial Lung Disease

Elif Gur Kabul[1], Pervin Demir[2], Berna Cagla Balkisli[3], Firdevs Ulutas[4], Sinem Yenil[5], Bilge Basakci Calik[5], Veli Cobankara[4]

[1]Uşak University Faculty of Health Sciences, Physiotherapy and Rehabilitation, Uşak, Türkiye
[2]Department of Biostatistics and Medical Informatics, Ankara Yıldırım Beyazıt University, Ankara, Türkiye
[3]Okan University Faculty of Health Sciences, Physiotherapy and Rehabilitation, İstanbul, Türkiye
[4]Department of Rheumatology, Pamukkale University Faculty of Medicine, Denizli, Türkiye
[5]Pamukkale University Faculty of Physiotherapy and Rehabilitation, Denizli, Türkiye

Cite this article as: Gur Kabul E, Demir P, Cagla Balkisli B, et al. The validity and reliability of the turkish version of modified medical research council dyspnea scale in systemic sclerosis patients with interstitial lung disease. *Thorac Res Pract.* 2024;25(6):215-220

Abstract

OBJECTIVE: The aim was to investigate the validity and reliability of the Turkish version of the Modified Medical Research Council (mMRC) Dyspnea Scale in Systemic Sclerosis (SSc) patients with Interstitial Lung Disease.

MATERIAL AND METHODS: Thirty patients diagnosed with SSc according to the 2013 EULAR/ACR criteria were included. After recording the demographic data of the patients, dyspnea was evaluated with the Visual Analogue Scale (VAS), exercise capacity with the 6 Minute Walk Distance (6MWD), fatigue level with the Fatigue Severity Scale (FSS), disease activity with the Medsger Disease Severity Scale, skin involvement with the Modified Rodnan Skin Score, and dyspnea level with the mMRC Dyspnea Scale. The mMRC Dyspnea Scale was administered to the patients with SSc who did not receive any treatment for test-retest reliability at 1-week intervals.

RESULTS: The observed scale range in mMRC (TR) was 0-4, and twelve out of the thirty patients (40%) were classified as having "moderate dyspnea." mMRC (TR) showed a significant moderate positive correlation with VAS dyspnea (rho: 0.718), a low negative correlation with 6MWD (rho: –0.445), and a low positive correlation with FSS (rho: 0.385). The weighted kappa statistic, used as an agreement scale for ordinal responses, was found to be 0.587 (indicating moderate agreement).

CONCLUSION: The Turkish version of the mMRC Dyspnea Scale demonstrates validity and reliability in SSc patients with interstitial lung disease.

KEYWORDS: Dyspnea, fatigue, sclerosis, skin
Received: *December 26, 2023* ***Revision Requested:*** *April 5, 2024* ***Last Revision Received:*** *July 31, 2024*
Accepted: *August 15 2024* ***Publication Date:*** *October 7, 2024*

INTRODUCTION

Systemic sclerosis (SSc) is a rare autoimmune disorder characterized by vasculopathy and fibrosis affecting both the skin and internal organs. Following skin involvement, SSc significantly impacts the lungs, heart, and digestive system. The lung manifestations of SSc typically present as interstitial lung disease (ILD) or pulmonary arterial hypertension, which is the primary cause of mortality in these patients.[1]

Despite the absence of approved treatments for SSc, significant progress is being made in the development of therapies for SSc-related interstitial lung disease (SSc-ILD). The number of studies examining the most effective methods for evaluating therapeutic responses in SSc-ILD is increasing every year, and these studies are providing valuable contributions to the literature.[2]

In addition to conventional assessments such as computed tomography scans and pulmonary function tests, the evaluation of patients' treatment response through patient-reported outcome measures (PROMs) is gaining recognition and significance. Patient-reported outcome measures aim to capture the beneficial and adverse impacts of changes in patients' health status on their daily activities, work, social engagements, and family interactions.[2]

Dyspnea is a subjective experience that can vary depending on individual and cultural contexts. To minimize misconceptions caused by this subjectivity, various scales are used.[3] The Modified Medical Research Council (mMRC) Dyspnea Scale is commonly employed due to its ease of application and understanding.[4] When validating surveys in languages other than the original, special attention is advised to maintain the integrity of the questions, as cultural factors and subjective interpretations can influence their meanings.[5] The Global Initiative for Chronic Obstructive Lung Disease recommends the mMRC Dyspnea Scale as the primary tool for assessing the severity of dyspnea.[4] It is a reliable measure

Corresponding author: Elif Gur Kabul, e-mail: elifgur1988@hotmail.com

positively correlated with lung function measurements and is considered suitable for evaluating symptoms in routine clinical practice.[6] The MRC, validated in idiopathic pulmonary fibrosis but not SSc-ILD, demonstrated responsiveness, reproducibility, and construct validity and independently predicted anxiety and depression in ILD.[2]

The aim of this study was to investigate the validity and reliability of the Turkish version of the mMRC Dyspnea Scale in individuals with SSc-ILD.

MATERIAL AND METHODS

Study Design

The study was planned as cross-sectional to assess the validity and reliability of the Turkish version of the Modified Medical Research Council (mMRC) Dyspnea Scale in patients diagnosed with systemic sclerosis (SSc) and interstitial lung disease (ILD). It was conducted in the rheumatology clinic and physiotherapy unit of Pamukkale University hospital, with approval from the ethics committee of Pamukkale University (approval no: E-60116787-020-132736; date: 11.02.2021). Prior to participation, all patients provided both written and verbal consent, in accordance with the Declaration of Helsinki guidelines. The patients diagnosed with SSc were prospectively registered.

Patients

Thirty SSc patients (27 females, 3 males, mean age 51.50 ± 12.14 years) were included in the study. Inclusion criteria were defined as a diagnosis of SSc-ILD by a rheumatologist in accordance with the 2013 EULAR/ACR criteria, aged at least 18 years, speaking and understanding Turkish, and voluntary participation in the study. Exclusion criteria were the presence of another autoimmune, neurological, orthopedic, psychiatric, or respiratory disease or a history of orthopedic surgery within the last year.

Variables

The patients were assessed in the outpatient clinic by the same investigator. Demographic data were recorded, and the following variables were evaluated: dyspnea using the Visual Analogue Scale (VAS), exercise capacity using the 6 minute walk distance (6MWD), fatigue level using the Fatigue Severity Scale (FSS), disease activity using the Medsger Disease Severity Scale (MDSS), skin thickness using the Modified Rodnan Skin Score (MRSS), and dyspnea levels using the mMRC Dyspnea Scale. The mMRC Dyspnea Scale was administered to the patients at 1-week intervals to assess test-retest reliability, with no changes made to pharmacological treatments during this period. Evaluations were conducted within approximately 30 minutes.

Main Points

- The weighted kappa statistic, used as an agreement scale for ordinal responses, was found to be 0.587.
- The mMRC Dyspnea Scale (TR) showed low to moderate correlation with dyspnea, exercise capacity, and fatigue.
- The mMRC Dyspnea Scale (TR) was found to be a valid and reliable scale for assessing dyspnea as an outcome measurement in SSc-ILD.

Translation and Cultural Adaptation of mMRC Dyspnea Scale

During the process of translation and cultural adaptation of the Modified Medical Research Council (mMRC) Dyspnea Scale, previously recommended procedures were followed in 5 stages.[7,8] The mMRC Dyspnea Scale (TR) is provided in Appendix 1.

Visual Analogue Scale

In this study, VAS was used to evaluate the "dyspnea" (0 = "I have no problems" to 10= "I always have problems"). The distance from 0 to the point marked by the patient was measured and defined as the score.[9]

Six-Minute Walk Distance

The 6 minute walk distance was used to assess exercise capacity, conducted within a designated 30-meter corridor following established testing protocols. Before the assessment, the patient's heart rate, peripheral oxygen saturation, and blood pressure at rest were recorded. Patients were instructed to walk the longest distance possible for 6 minutes, starting from a marked area under investigator supervision in a closed hospital corridor. The patient's walking distance was then recorded in meters.[10]

Fatigue Severity Scale

The FSS is a questionnaire comprising 9 items aimed at evaluating the intensity of fatigue and its effect on daily activities and quality of life. Each item is rated on a 7-point scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Scores range from a minimum of 1 to a maximum of 7, with higher scores indicating more severe fatigue. The scores from each item are summed, and the total value is divided by 9. A FSS score of 4 or above is indicative of severe fatigue.[11]

Medsger's Disease Severity Scale

Medsger's Disease Severity Scale, developed by Medsger et al, assesses the severity of the disease across 9 organ systems: general health, peripheral vascular, skin, joint/tendon, muscle, gastrointestinal (GI) tract, lungs, heart, and kidneys. Each organ system is assigned a score from 0 to 4, representing the absence of involvement to end-stage severity. The total score is obtained by summing the scores of all 9 organ systems for each variable, resulting in a scale ranging from 0 (indicating lower severity) to 36 (indicating higher severity).[12]

Modified Rodnan Skin Score

Skin thickening in 17 different parts of the body is scored between 0 and 3 (0 = normal, 1 = slight thickening, 2 = moderate thickening, 3 = severe thickening) and summed (total score 0-51) by compressing the skin between the fingers.[13]

Modified Medical Research Council Dyspnea Scale

The scale provides information about the levels of perception of dyspnea, consisting of 5 stages. Patients were instructed to select the most appropriate activity level that induces breathlessness. Scoring ranges from 0 to 4: 0 points indicate no dyspnea (I only get breathless with strenuous exercise); 1 point indicates mild dyspnea (I get short of breath when hurrying on the level or walking up a slight hill); 2 points indicates moderate dyspnea (I walk slower

than people of the same age on the level because of breathlessness or have to stop for breath when walking at my own pace on the level); 3 points indicates severe dyspnea (I stop for breath after walking about 100 yards or after a few minutes on the level) and 4 points indicates very severe dyspnea (I am too breathless to leave the house or I am breathless when dressing).[14]

Study Size

The required sample size was calculated to be at least 25 patients (5 patients per response category) based on literature recommendations.[15] The estimated minimum sample size was determined to be 29 in order to detect at least a 0.50 (moderate) correlation coefficient, and for the weighted kappa statistics, the sample size was determined to be 15 in order to detect at least a 0.40 (moderate) agreement, based on an alpha of 0.05 and a power of 80%.[16] In this study, the results of 30 patients were evaluated.

Statistical Methods

The statistical analyses and calculations were performed using the IBM Statistical Package for the Social Sciences version 21.0 software for Windows (IBM SPSS Corp.; Armonk, NY, USA), Microsoft Excel version 2010 (Microsoft, Redmond, State of Washington, USA), and "DescTools" package[17] in R software[18] (company, city, country). The 2-sided significance level was accepted as $P < 0.05$. The normality of continuous variables was assessed using the Shapiro–Wilk test. Quantitative data were summarized as mean ± SD or median (minimum; maximum), while qualitative data were presented as frequency (percentage). The Spearman rho correlation coefficient was calculated between mMRC (TR) and other measures to assess convergent-construct validity, which refers to the degree of theoretical relationship between 2 measures of mMRC (TR). The coefficient interpretation followed the classification proposed by Hinkle et al[19] The test-retest reliability of the mMRC (TR) was established using the weighted Kappa coefficient for the ordinal scale. Standards for the strength of agreement for kappa values were as follows: 0 or lower = poor, 0.01-0.20 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, and 0.81-1.0 = almost perfect.[20] The standard error of measurement (SEM) and minimal detectable change (MDC95) based on the test-retest reliability were calculated for agreement using the following equations: $SEM = SD^* \sqrt{1-r}$, (SD is the standard deviation of the measures, r is kappa as the test-retest reliability coefficient) and $MDC_{95} = SEM^*1.96^* \sqrt{2}$.[21]

RESULTS

Patients

The mean age of the patients was 51.50 ± 12.14 years. Twenty-seven of the thirty patients were female. The dominant side for 28 patients was the right. Most of the patients were not working ($n = 23$, 76.7%). The demographic characteristics and clinical measurements of the patients are summarized in Table 1.

The observed scale range was 0-4 in mMRC (TR). Twelve patients (40%) in the total sample were classified as "moderate," and the median mMRC (TR) score was 2.

Table 1. Demographic Characteristics and Clinical Measurements of the Patients

Variables	Mean ± SD	Median (Min-Max)
Age (years)	51.50 ± 12.14	52 (28-68)
BMI (kg/m^2)	25.79 ± 4.97	26.22 (16.16-35.94)
Disease duration (years)	13.53 ± 9.27	12 (2-38)
ANA	(*n*)	
Nukleolar	8	
Centromere	8	
Homogeneous	10	
Negative	2	
Nuklear DOT	1	
Granular	1	
VAS-Dyspnea	0 response: 14	1.7 (0-9)
6 Minute Walk Distance (meter)	341.93 ± 89.54	352.5 (120-516.3)
Fatigue Severity Scale	4.91 ± 1.62	5.44 (1.11-6.88)
Medsger's Disease Severity Scale	5.67 ± 2.56	5 (2-13)
Modified Rodnan Skin Score	15.47 ± 8.36	12 (5-32)

BMI, body mass index; Max, maximum; Min, minimum;VAS: visual analogue scale.

Validity

mMRC (TR) had moderate positive correlation with VAS dyspnea (rho: 0.718; 95% CI: 0.482;0.857), low negative correlation with 6MWD (rho:-0.445; 95% CI: -0.694; -0.100) and low positive correlation with FSS (rho: 0.385; 95% CI: 0.028;0.654) (Table 2).

Test Re-Test Reliability

The weighted kappa statistic for agreement on the scale for ordinal response was found to be 0.587 (moderate agreement; 95% CI: 0.415, 0.759). The mMRC (TR) and mMRC (TR) re-test scores were similar ($P = 0.449$).

DISCUSSION

The mMRC Dyspnea Scale (TR) was found to be a valid and reliable scale for use as an outcome measurement to assess dyspnea in SSc-ILD.

Table 2. Spearman Correlation Analysis Result

Variables	Rho* (Lower; Upper)	*P*
VAS – Dyspnea	0.718 (0.482; 0.857)	<.001
6 Minute Walk Distance (meter)	–0.445 (–0.694; –0.100)	.014
Fatigue Severity Scale	0.385 (0.028; 0.654)	.036

*Spearman's rho correlation coefficients and 95% CI. If the confidence interval includes 0 (zero), that means that no significant correlation. Interpreting the size of absolute rho: <0.30: negligible; <0.50: low; <0.70: moderate; <0.90: high; ≥0.90: very high correlation.
VAS, visual analogue scale.

Interstitial lung disease alone has a 33% impact on mortality in SSc.[22] However, dyspnea (initially on exertion and eventually at rest) and non-productive cough are the most common symptoms for respiratory problems.[23] Therefore, dyspnea assessment can be considered necessary and important. In the literature, Functional Assessment of Chronic Illness Therapy-Dyspnea,[24] Turkish Breathlessness Beliefs Questionnaire,[25] Saint George's Respiratory Questionnaire (SGRQ),[26] were reported to be valid and reliable to assess respiratory symptoms in SSc. Respiratory-specific health-related quality of life (HRQOL) tools, such as the SGRQ, reflect the burden of respiratory symptoms, while dyspnea-specific assessment instruments have been shown to be useful in assessing dyspnea symptom response to therapy.[27] The MRC scale does not consider the range of motion achievable by a movement nor does it define the strength of resistance that can be overcome by a movement.[28] Additionally, the use of mMRC is recommended for the evaluation of dyspnea.[29] Compared to other questionnaires evaluating respiratory symptoms, mMRC was chosen for its validity and reliability. It is dyspnea-specific, short, and related to daily life. It is also frequently used in practice because it is both easy to administer and has prognostic value in measuring the severity of dyspnea.[30]

The original English version of the mMRC dyspnea scale was developed by Bestall et al[4] Kovelis et al observed that the mMRC Dyspnea Scale is valid for Brazilian patients with chronic obstructive pulmonary disease (COPD) (ICC = 0.83).[6] Additionally, Ribeiro et al conducted the cultural adaptation and validation of the modified mMRC scale in Portuguese COPD patients for the measurement of breathlessness (ICC = 0.912).[31]

Saketkoo et al examined patient-reported outcome measures for SSc-ILD and recommended the mMRC for its strong utility in stratifying severity.[2] However, to the best of our knowledge, there has been no validity and reliability report in the literature of mMRC to assess the dyspnea in SSc-ILD. This study is the first and only study on this subject.

The distance walked and the level of desaturation during the 6MWD can be measured practically, inexpensively, and easily at a submaximal level, and its usage is recommended by the American Thoracic Society.[10,32] Originally developed to study patients with heart failure and pulmonary disease, the 6MWD has been increasingly used as an outcome measure in clinical SSc trials.[33] Six-minute walk distance is highly reproducible in SSc-ILD patients.[34] In the literature, studies evaluating the validity and reliability of the mMRC scale and other scales measuring dyspnea have used non-scale parameters such as spirometric tests, blood gas tensions, a shuttle walking test, and the 6MWD, which are not scored as a Likert scale.[4,31,35,36] Consequently, in our study where we assessed the Turkish validity and reliability of the mMRC scale in SSc patients, we chose the 6MWD as one of the parameters used for the validity of the scale. Dyspnea is a multifaceted symptom influenced by various factors, not solely limited to the degree of respiratory functional impairment in patients.[37] In addition to respiratory reserve, cardiovascular and peripheral muscle conditions can also impact it. The 6MWD does not reflect the same physiological process in every disease; it can be influenced by dyspnea in patients with pulmonary involvement and by lower extremity pain in SSc patients without ILD or PH.[38] A study assessing the reproducibility of the 6MWD in a cohort of SSc patients demonstrated that it is both feasible and reliable. The test measurements correlated reasonably but variably with functional and morphological measures of disease severity.[39] In a systematic review,[40] 6MWD results were found to average 388 meters (362-415 meters), while in our study, the average was 341.93 meters (120-516.3 meters). However, we believe that the low negative correlation observed between the mMRC (TR) scale and 6MWD may be specifically influenced by peripheral muscle conditions.

The VAS-breathing scale is the simplest validated measure of dyspnea in SSc-ILD. This scale provides a direct measurement and is associated with forced vital capacity (FVC), high-resolution computed tomography, and health-related quality of life (HRQOL) measurement. Additionally, VAS-breathing was found to be valid for assessing SSc-ILD in the Scleroderma Lung Study.[41] mMRC (TR) had a moderate correlation with VAS dyspnea in the present study. Wallace et al used the VAS breathing scale, which ranged from 0-150. A higher score means a worse degree to which dyspnea affects daily activities in patients. They stated that the average VAS breathing score was 20.4 and that the VAS for breathing scale had a high correlation with the SGRQ scales (r = 0.46-0.61).[26]

Fatigue is one of the common symptoms in SSc-ILD.[20] According to a study by Yakut et al[42] 80% of SSc patients were experiencing fatigue. Cognitive, physical, psychosocial subscale, and total scores of fatigue were higher than healthy controls. Additionally, there was a significant correlation between dyspnea severity (mMRC) and fatigue (r = 0.621, $P \leq 0.01$). Gök et al revealed that FSS had high reproducibility in SSc.[43] In the present study, FSS was used to assess fatigue severity and the validity of mMRC (TR) in SSc-ILD. mMRC (TR) had a low positive correlation with FSS (r: 0.385; 95% C.I.: 0.028; 0.654). The dyspnea severity may increase with pulmonary dysfunction, which can affect the daily life and functionality of the person, resulting in fatigue.

The reliability of clinicians' ratings is crucial such as the interpretation of examination results and diagnosis. These ratings are often categorized on a nominal or ordinal scale. The kappa coefficient serves as a suitable measure of reliability for such data.[44] The kappa coefficient is a measure of agreement, with 2 known paradoxes including bias and prevalence, which takes into account the agreement by change (expected agreement).[45] In the present study, the weighted kappa statistic for agreement on the mMRC (TR) for ordinal response was found to be moderate. mMRC (TR) can achieve the same result with repeated administrations. This means that the mMRC Dyspnea Scale (TR) was found to be a reliable scale for assessing dyspnea as an outcome measurement in SSc-ILD.

Disease severity and skin thickness did not correlate to mMRC (TR) in the present study. This may be attributed to the overall health status of our group. Similar to our findings, Hinchcliff et al[24] reported that patients with high skin thickness scores

may experience dyspnea secondary to restrictive lung disease and the mean severity index score in their sample group was 2.79 (observed range of 0-9), while in the present study it was 5.67 (observed range of 2-13).

This was the first study to evaluate the validity and reliability of the mMRC in SSc-ILD patients, and the clinical heterogeneity of this study population may have contributed to the evaluation of the sensitivity of the scale. However, this study had some limitations. First, the results were not supported by multicenter recruitment of the SSc population. The authors could not collaborate with other hospitals to evaluate patients with SSc in different regions. Therefore, our results were limited to patients in a single region. Second, the male-to-female ratio was skewed, with most of the study sample consisting of females. As a result, the findings predominantly reflect symptoms in females. Third, no objective measurement of dyspnea markers could be used.

In conclusion, the mMRC Dyspnea Scale (TR) can be used as an outcome measurement to assess dyspnea and give an idea about pulmonary influence in SSc-ILD. It can be preferred as a valid and reliable scale for evaluating the disease and/or examining the efficacy of pharmacological and non-pharmacological treatments in SSc, which is often accompanied by pulmonary involvement.

Availability of Data and Materials: The data that support the findings of this study are available on request from the corresponding author.

Ethics Committee Approval: Ethics committee of Pamukkale University (approval no: E-60116787-020-132736; date: 11.02.2021).

Informed Consent: Verbal and written informed consent was obtained from the patients who agreed to take part in the study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – E.G.K., S.Y., B.B.C., F.U., V.C.; Design – E.G.K., S.Y., B.B.C., F.U., V.C.; Supervision – E.G.K., B.B.C., F.U.; Data Collection and/or Processing – E.G.K., F.U.; Analysis and/or Interpretation – P.D.; Literature Search – B.C.B., S.Y.; Writing – E.G.K., B.C.B.; Critical Review – B.B.C., V.C.

Declaration of Interests: The authors have no conflicts of interest to declare.

Funding: This study received no funding.

REFERENCES

1. Tyndall AJ, Bannert B, Vonk M, et al. Causes and risk factors for death in systemic sclerosis:a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis*. 2010;69(10):1809-1815. [CrossRef]
2. Saketkoo LA, Scholand MB, Lammi MR, Russell AM. Patient-reported outcome measures in systemic sclerosis-related interstitial lung disease for clinical practice and clinical trials. *J Scleroderma Relat Disord*. 2020;5(2)(suppl):48-60. [CrossRef]
3. von Leupoldt A, Taube K, Henkhus M, Dahme B, Magnussen H. The Impact of affective states on the perception of dyspnea in patients with chronic obstructive pulmonary disease. *Biol Psychol*. 2010;84(1):129-134. [CrossRef]
4. Bestall JC, Paul EA, Garrod R, Garnham R, Jones PW, Wedzicha JA. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax*. 1999;54(7):581-586. [CrossRef]
5. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-1432. [CrossRef]
6. Kovelis D, Segretti NO, Probst VS, Lareau SC, Brunetto AF, Pitta F. Validation of the modified pulmonary functional status and dyspnea questionnaire and the Medical Research Council scale for use in Brazilian patients with chronic obstructive pulmonary disease. *J Bras Pneumol*. 2008;34(12):1008-1018. [CrossRef]
7. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila. Pa 1976)*. 2000;25(24):3186-3191. [CrossRef]
8. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94-104. [CrossRef]
9. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health*. 1990;13(4):227-236. [CrossRef]
10. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002;166(1):111-117. [CrossRef]
11. Merkies ISJ, Schmitz PIM, Van der Meché FGA, Samijn JPA, Van Doorn PA, Inflammatory Neuropathy Cause and Treatment (INCAT) Group. Connecting impairment,disability and handicap in immune mediated polyneuropathies. *J Neurol Neurosurg Psychiatry*. 2003;74(1):99-104. [CrossRef]
12. Medsger TA, Bombardieri S, Czirjak L, Scorza R, Della Rossa A, Bencivelli W. Assessment of disease severity and prognosis. *Clin Exp Rheumatol*. 2003;21(3)(suppl 29):S42-S46.
13. Clements P, Lachenbruch P, Siebold J, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol*. 1995;22(7):1281-1285.
14. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest*. 1988;93(3):580-586. [CrossRef]
15. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin JB. Sample size used to validate a scale:a review of publications on newly-developed patient reported outcomes measures. *Health Qual Life Outcomes*. 2014;12(12):176. [CrossRef]
16. Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Cohen's Kappa. *Epidemiol Biostat Public Health*. 2017;14(2):e12267. [CrossRef]
17. Signorell A, Aho K, Alfons A, Anderegg N, Aragon T, Arachchige C, et al. DescTools: Tools for descriptive statistics. R package version 0.99.40. https://cran-r-project.org/package=DescTools. 2021.
18. R Core Team. R: A language and environment for statistical computing. *R foundation for statistical computing*, Vienna, Austria. https://www.R-project.org/. 2021.
19. Hinkle DE, Wiersma W, Jurs SG. *Applied Statistics for the Behavioral Sciences*. 5th ed. Boston(MA):Houghton Mifflin Co.; 2003.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174. [CrossRef]
21. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231-240. [CrossRef]
22. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis,1972-2002. *Ann Rheum Dis*. 2007;66(7):940-944. [CrossRef]

23. Perelas A, Silver RM, Arrossi AV, Highland KB. Systemic sclerosis-associated interstitial lung disease. *Lancet Respir Med*. 2020;8(3):304-320. [CrossRef]
24. Hinchcliff M, Beaumont JL, Thavarajah K, et al. Validity of two new patient-reported outcome measures in systemic sclerosis:Patient-Reported Outcomes Measurement. *Arthritis Care Res (Hoboken)*. 2011;63(11):1620-1628. [CrossRef]
25. Ustun O, Bayraktar D, Kurut Aysin I, et al. Assessing dyspnea-related kinesiophobia in patients with systemic sclerosis (SSc):validity and reliability of Turkish Breathlessness Beliefs Questionnaire for SSc. *Clin Rheumatol*. 2022;16.
26. Wallace B, Kafaja S, Furst DE, et al. Reliability, validity and responsiveness to change of the Saint George's Respiratory Questionnaire in early diffuse cutaneous systemic sclerosis. *Rheumatol (Oxf Engl)*. 2015;54(8):1369-1379. [CrossRef]
27. Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med*. 2014;370(22):2071-2082. [CrossRef]
28. Bohannon RW. Manual muscle testing of the limbs: considerations, limitations, and alternatives. *Phys Ther Pract*. 1992;2: 11-21.
29. Singh D, Agusti A, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease: the GOLD Science Committee report 2019. *Eur Respir J*. 2019;53(5):1900164. [CrossRef]
30. Perez T, Burgel PR, Paillasseur JL, et al. Modified Medical Research Council scale vs Baseline Dyspnea Index to evaluate dyspnea in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2015;10(10):1663-1672. [CrossRef]
31. Ribeiro S, Cardoso CS, Valério M, et al. Confirmatory evaluation of the modified Medical Research Council questionnaire for assessment of dyspnea in patients with chronic obstructive pulmonary disease in Portugal. *Acta Med Port*. 2022;35(2):89-93. [CrossRef]
32. Lama VN, Flaherty KR, Toews GB, et al. Prognostic value of desaturation during a 6-minute walk test in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med*. 2003;168(9):1084-1090. [CrossRef]
34. Ahmadi-Simab K, Hellmich B, Gross WL. Bosentan for severe pulmonary arterial hypertension related to systemic sclerosis with interstitial lung disease. *Eur J Clin Investig*. 2006;36(suppl 3):44-48. [CrossRef]
35. Buch MH, Denton CP, Furst DE, et al. Submaximal exercise testing in the assessment of interstitial lung disease secondary to systemic sclerosis: reproducibility and correlations of the 6-min walk test. *Ann Rheum Dis*. 2007;66(2):169-173. [CrossRef]
36. Kim HK, Lee H, Kim SH, et al. Validation of the Korean version of the bronchiectasis health questionnaire. *Tuberc Respir Dis (Seoul)*. 2020;83(3):228-233. [CrossRef]
37. Simsic AA, Yorke J, Regueiro EG, Di Lorenzo VPD, Baddini-Martinez J. Validation of the dyspnoea-12 scale into Portuguese speaking COPD patients. *Clin Respir J*. 2018;12(5):1942-1948. [CrossRef]
38. Parshall MB, Schwartzstein RM, Adams L, et al. An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. *Am J Respir Crit Care Med*. 2012;185(4):435-452. [CrossRef]
39. Garin MC, Highland KB, Silver RM, Strange C. Limitations to the 6-minute walk test in interstitial lung disease and pulmonary hypertension in scleroderma. *J Rheumatol*. 2009;36(2):330-336. [CrossRef]
40. Chatterjee AB, Rissmiller RW, Meade K, et al. Reproducibility of the 6-minute walk test for ambulatory oxygen prescription. *Respiration*. 2010;79(2):121-127. [CrossRef]
41. Vandecasteele E, De Pauw M, De Keyser F, et al. Six-minute walk test in systemic sclerosis:A systematic review and meta-analysis. *Int J Cardiol*. 2016;212(212):265-273. [CrossRef]
40. Tashkin DP, Roth MD, Clements PJ, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLSII): a randomised controlled, double-blind, parallel group trial. *Lancet Respir Med*. 2016;4(9):708-719. [CrossRef].
42. Yakut H, Özalevli S, Birlik AM. Fatigue and its relationship with disease-related factors in patients with systemic sclerosis: A cross-sectional study. *Turk J Med Sci*. 2021;51(2):530-539. [CrossRef].
43. Gök K, Cengiz G, Erol K, Özgöçmen S. The Turkish version of multidimensional assessment of fatigue and fatigue severity scale is reproducible and correlated with other outcome measures in patients with systemic sclerosis. *Arch Rheumatol*. 2016;31(4):329-332.
44. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85(3):257-268. [CrossRef]
45. Altman DG. *Practical Statistics for Medical Research*. CRC Press; Boca Raton; 1990.

Appendix 1 The Turkish version of Modified Medical Research Council Dyspnea Scale

0 Yalnızca yorucu egzersizlerden dolayı nefesim kesiliyor

1 Düz yolda acele ederken veya hafif bir yokuş yukarı yürürken nefesim daralıyor

2 Nefes darlığı nedeniyle düz yolda aynı yaştaki insanlara göre daha yavaş yürüyorum ya da düz yolda kendi hızımda yürürken nefes almak için durmak zorunda kalıyorum

3 Yaklaşık 100 metre yürüdükten sonra veya düz yolda birkaç dakika yürüdükten sonra nefes almak için duruyorum

4 Evden çıkamayacak kadar nefesim kesiliyor veya giyinirken nefes nefese kalıyorum