

# The accuracy with which the 5 times sit-to-stand test, versus gait speed, can identify poor exercise tolerance in patients with COPD

## A cross-sectional study

Roberto Bernabeu-Mora, MD<sup>a,b,\*</sup>, Francesc Medina-Mirapeix, MD<sup>b</sup>, Eduardo Llamazares-Herrán, MS<sup>c</sup>, Silvana Loana de Oliveira-Sousa, MD<sup>d</sup>, M<sup>a</sup> Piedad Sánchez-Martinez, MS<sup>b</sup>, Pilar Escolar-Reina, MD<sup>b</sup>

### Abstract

Identifying those patients who underperform in the 6-minute walk test (6MWT <350 m), and the reasons for their poor performance, is a major concern in the management of chronic obstructive pulmonary disease.

To explore the accuracy and relevance of the 4-m gait-speed (4MGS) test, and the 5-repetition sit-to-stand (5STS) test, as diagnostic markers, and clinical determinants, of poor performance in the 6MWT.

We recruited 137 patients with stable chronic obstructive pulmonary disease to participate in our cross-sectional study. Patients completed the 4MGS and 5STS tests, with quantitative (in seconds) and qualitative ordinal data collected; the latter were categorized using a scale of 0 to 4. The following potential covariates and clinical determinants of poor 6MWT were collated: age, quadriceps muscle-strength (QMS), health status, dyspnea, depression, and airflow limitation. Area under the receiver-operating characteristic curve data (AUC) was used to assess accuracy, with logistic regression used to explore relevance as clinical determinants.

The AUCs generated using the 4MGS and 5STS tests were comparable, at 0.719 (95% confidence interval [CI] 0.629–0.809) and 0.711 (95% CI 0.613–0.809), respectively. With ordinal data, the 5STS test was most accurate (AUC of 0.732; 95% CI 0.645–0.819); the 4MGS test showed poor discriminatory power (AUC <0.7), although accuracy improved (0.726, 95% CI 0.637–0.816) when covariates were included. Unlike the 4MGS test, the 5STS test provided a significant clinical determinant of a poor 6MWT (odds ratio 1.23, 95% CI 1.05–1.44).

The 5STS test reliably predicts a poor 6MWT, especially when using ordinal data. Used alone, the 4MGS test is reliable when measured with continuous data.

**Abbreviations:** 4MGS test = 4-m gait-speed test, 5STS test = 5-repetition sit-to-stand test, 6MWT = 6-minute walk test, CAT = Chronic Obstructive Pulmonary Disease Assessment Test, COPD = chronic obstructive pulmonary disease, FEV<sub>1</sub> = forced expiratory volume in 1 second, MRC = Medical Research Council, GOLD = Global Initiative for Chronic Obstructive Lung Disease, HAD = Hospital Anxiety and Depression Scale, QMS = quadriceps muscle strength, LR+ = positive likelihood ratio, LR– = negative likelihood ratio, ROC = receiver-operating characteristic, AUC = area under the curve, SPPB = Short Physical Performance Battery.

**Keywords:** 4-m gait speed, 5-repetition sit-to-stand, 6-minute walk test, chronic obstructive pulmonary disease, Hospital Anxiety and Depression Scale, quadriceps muscle strength, Short Physical Performance Battery

Editor: Anish Thachangattuthodi.

Author contributions: The study conception and design was developed by RB-M, FM-M, and SLO-S. Statistical analysis was conducted by RB-M, FM-M, ELI-H, SLO-S, MPS-M, and PE-R. The manuscript was drafted and revised by RB-M, FM-M, ELI-H, and PE-R.

Ethics approval: All study participants gave their informed consent, and the study protocol was approved by the institutional review board of the hospital (approval number: EST-35/13).

Funding: The study was supported by AstraZeneca pharmaceutical Spain, S.A., Menarini and Pfizer.

The authors report no conflicts of interest.

<sup>a</sup> Division of Pneumology, Hospital Morales Meseguer, <sup>b</sup> Department of Physical Therapy, University of Murcia, Murcia, <sup>c</sup> Division of Rehabilitation, Hospital Ramón y Cajal, Madrid, <sup>d</sup> Department of Physical Therapy, Catholic University, Murcia, Spain.

\* Correspondence: Roberto Bernabeu-Mora, Division of Pneumology, Hospital Morales Meseguer, Department of Physical Therapy, University of Murcia, Avda Marqués de los Velez s/n. 30008. Murcia, Spain (e-mail: rbernabeumora@hotmail.com).

Copyright © 2016 the Author(s). Published by Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medicine (2016) 95:35(e4740)

Received: 15 January 2016 / Received in final form: 3 August 2016 / Accepted: 5 August 2016

<http://dx.doi.org/10.1097/MD.0000000000004740>

## 1. Introduction

Exercise intolerance is an important indicator of the severity of chronic obstructive pulmonary disease (COPD)<sup>[1]</sup>; this disability erodes the ability to perform everyday tasks, and substantially diminishes quality of life.<sup>[2,3]</sup> For this reason, our ability to quantify exercise capacity using tests such as the 6-minute walk test (6MWT) is a necessary and integral part of the routine clinical assessment of COPD to timely identify the functional decline, and to provide advice and preventive interventions.<sup>[4]</sup>

The 6-minute walk test (6MWT) is widely regarded to be a reliable measure of exercise tolerance in patients with COPD,<sup>[5]</sup> and a good predictor of morbidity and mortality.<sup>[6]</sup> Indeed, walking less than 350 m in this test has been associated with a significant increase in mortality.<sup>[3]</sup> Given the prognostic significance of performing poorly on the 6MWT, some authors have gone on to identify degree of airflow limitation, dyspnea, and the presence of depressive symptoms as 3 significant clinical determinants of a poor performance in this test (<350 m).<sup>[7]</sup>

Despite the relevance of the 6MWT, it is not typically measured by pulmonologists in the clinic due to time constraints, equipment, and obvious space requirements.<sup>[8]</sup> Instead, there is currently the tendency to identify tests and measures that are faster, and need less space, but still reliably measure exercise capacity.<sup>[9,10]</sup> In this sense, the 4-m gait-speed (4MGS) test, or 5-repetition sit-to-stand (5STS) test are plausible substitutes for the 6MWT.<sup>[8,10]</sup> Additionally, the 4MGS has been recently recommended as a potentially useful marker to classify subjects with poor 6MWT performance.<sup>[11]</sup>

Factors other than exercise tolerance can influence scores for these 2 tests in COPD patients (i.e., test scores tend to decline with increasing age).<sup>[10,11]</sup> However, to date, it remains unknown as to which factors or covariates associated with the 4MGS test affects accuracy to classify subjects between a normal and poor 6MWT performance. Moreover, the capacity of the 5STS test to accurately report, and act as a clinical determinant of exercise intolerance, also remains largely untested.

The primary aim of this study was to identify whether, and to what extent, both the 4MGS and 5STS tests could be used as diagnostic markers, and clinical determinants, of a poor 6MWT performance. Our specific objectives were: to explore whether covariates associated with the 5STS and 4MGS tests could substantially affect their accuracy; their ability of these tests to be used as stand-alone markers of performance; and to explore their relative efficacy as clinical determinants of a poor 6MWT performance, after adjustment for other clinical determinants. We hypothesized that both tests, recorded using continuous data (seconds), or ordinal data (0–4) as used in the short physical performance battery (SPPB),<sup>[12]</sup> would show good and comparable discriminatory accuracy, and that this accuracy would improve after correction for significant covariates.

## 2. Methods

### 2.1. Study design and participants

A controlled, cross-sectional study was devised. Patients with stable COPD were prospectively recruited from an outpatient pulmonary service at Meseguer Hospital in Murcia, Spain, during 2015. All study participants provided written informed consent, and the study protocol was approved by the institutional review board of the hospital called “Comité Ético de Investigación Clínica del Hospital General Universitario José María Morales Meseguer” (approval number: EST-35/13).

Inclusion criteria included COPD patients with a forced expiratory volume in 1 second (FEV<sub>1</sub>)/forced vital capacity ratio of <70% of the predicted value, and an age of between 40 and 80 years. The diagnosis of COPD, and its severity, was based of the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines.<sup>[13]</sup> Patients with an unstable cardiac condition within 4 months of the start of the study were excluded, as were those with cognitive deterioration, or a limitation in walking. During a 1-year period, a consecutive sample of eligible patients was identified from patient health examinations. A pulmonary physician assessed the eligibility criteria for recruitment.

### 2.2. Measures

**2.2.1. Six-minute walk test.** The 6MWT was performed indoors, along a flat, straight, 30-m walking course, supervised by 2 well-trained nurses (with a mean of 19 years of experience), according to American Thoracic Society guidelines.<sup>[14]</sup> Patients were instructed and encouraged to walk as far as possible in 6 minutes, using standard incentive phrases every minute. Patients were permitted to stop and rest during the test, but were instructed to resume walking as soon as they felt able to do so.

**2.2.2. Physical performance tests and potential covariates.** Participants performed the 4MGS test and the 5STS test according to the National Institute on Aging protocol for the SPPB.<sup>[15]</sup> The 5STS measures the time (in seconds) taken to stand 5 times from a sitting position, as rapidly as possible. The 4MGS test measures the time taken to cross a 4-m line at usual speed. Categorical scores (range 0–4) for both the 4MGS and the 5STS tests were based on timed quartiles, established previously in a large population.<sup>[12]</sup>

A total of 5 variables were selected from our literature search, and measured as potential covariates, based on their association with either test. These variables were age (years), quadriceps muscle strength (QMS), perception of health status, dyspnea, and reporting of symptoms of depression.<sup>[10,11,16]</sup> Most of these variables—the 5STS and the 4MGS tests—were measured by a well-trained pulmonology physician (12 years of experience). QMS was measured first, and then, the 4MGS and the 5STS tests were measured. Afterwards, the 6MWT was performed. Finally, self-reported questions were surveyed using specific questionnaires.

The QMS was measured using a calibrated dynamometer. Based on previous descriptions,<sup>[17]</sup> participants remained seated on a raised plinth, with their assessor placing the participant's knee in a flexed position, at 70°. The dynamometer was placed perpendicular to the anterior tibia, 5 cm above the lateral malleolus, with force applied gradually over a 4-second period to allow maximal muscle fiber recruitment. Two trials per test were performed, as previously described.<sup>[17]</sup>

Perception of health status was assessed using the COPD Assessment Test (CAT).<sup>[18]</sup> Dyspnea was measured by the self-administered modified Medical Research Council (mMRC) dyspnea scale.<sup>[19]</sup> Presence of depressive symptoms were assessed according to the Hospital Anxiety and Depression Scale (HADS).<sup>[20]</sup> This scale ranges from 0 to 21 points. Binary categories (yes/no) were created for this study, using an accepted cut-off of 8 for depression.<sup>[21]</sup> QMS was measured first, and then, the 4MGS and the 5STS tests were measured. Afterwards, the 6MWT was performed. Finally, self-reported questions were surveyed using specific questionnaires.

### 2.2.3. Clinical determinants of a poor 6MWT performance.

We selected 3 significant clinical determinants of a poor 6MWT performance (<350 m) as determined by Spruit et al.<sup>[7]</sup> These

were airflow limitation by GOLD guidelines<sup>[13]</sup>, dyspnea, and symptoms of depression. Airflow was measured using a MasterScope spirometer. To calculate the per cent of predicted pulmonary function, we used predictive equations derived from the Third National Health and Nutrition Examination Survey.<sup>[22]</sup>

### 2.3. Statistical analyses

Descriptive statistics were used to characterize our patient cohort. We used Pearson chi-square test and the independent *t* test to examine between-group differences in criteria with respect to the presence or absence of a poor 6MWT performance.

We constructed receiver-operating characteristic (ROC) curves, with area under the curve (AUC) data used to determine the test accuracy alone, or combined with covariates. Separate analyses were conducted for continuous and ordinal data. According to previous authors,<sup>[23]</sup> an AUC >0.7 was used as the criterion of good discrimination.

To determine the combination score of each test and covariates, we calculate the predicted probability of being able to discriminate between a normal and poor 6MWT test score using logistic regression models (with enter method), where the dependent variable was a poor 6MWT (yes/no), and independent variables were each test and their covariates. To select these covariates, we initially explored among controls (i.e., patients with good 6MWT) the association between all measured covariates, and each marker, by means of linear regression (with enter method).<sup>[24]</sup> Those covariates found to be significant were then selected.

In addition, we selected the best cut-off score for each test, when used alone. This was the value at which sensitivity +

specificity – 1 was maximized. Using these cut-off scores, we calculated the sensitivity (Se), specificity (Sp), positive (LR+), and negative (LR–) likelihood ratios. We also calculated Se, Sp, and LRs for the combined scores, by examining data fit with regression models and predicted probabilities.

Finally, multivariate logistic regression analyses were used to assess the relative importance of the 2 tests (4MGS and 5STS) as clinical determinants of a poor 6MWT performance. Separate models were devised for each test in isolation, or combined with covariates. Each model was adjusted according to those clinical determinants of a poor 6MWT, as identified in the literature,<sup>[7]</sup> which showed significant between-group differences ( $P < 0.10$ ) in previous bivariate analyses. “Goodness-of-fit” and regression diagnostics for each model were assessed using methods as described elsewhere.<sup>[24]</sup>

As a rule of thumb, we tested 15 subjects per predictor; this was an estimate of the sample size needed to generate reliable data.<sup>[25]</sup> We recruited a minimum of 120 participants, assuming a maximum of 8 predictors. All analyses were performed with the SPSS statistical software program (SPSS version 19.0; IBM SPSS, Chicago, IL).

## 3. Results

### 3.1. Characteristics of patients

A total of 173 patients with COPD were screened during the study period. Of them, 147 were initially enrolled as potentially eligible. All agreed to participate in the study, but 10 failed to meet our eligibility criteria. Therefore, 137 patients were finally included in our patient cohort. Participants’ characteristics are described in Table 1. Patients were generally male, with moderate

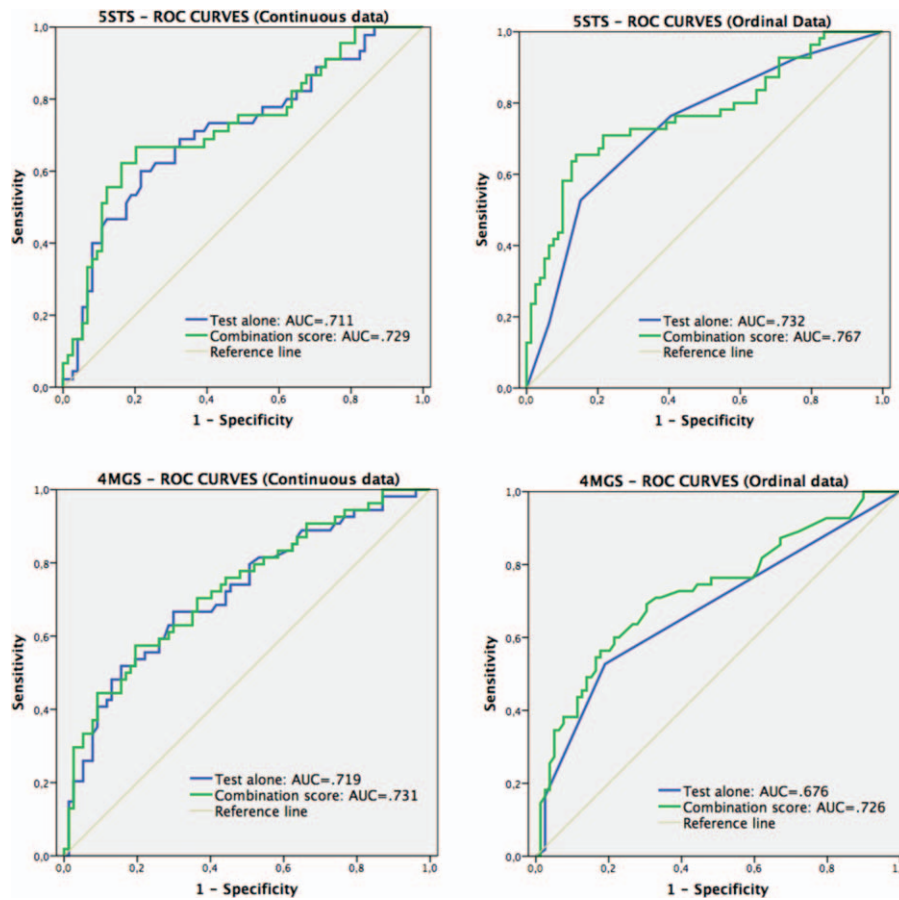
**Table 1**  
Patient characteristics stratified according to poor 6MWT performance.

Characteristics	Number (%) or mean ± SD			P
	Total sample (N = 137)	<350 m walked (n = 55)	≥350 m walked (n = 79)	
Demographic and clinical				
Age, y	66.9 ± 8.3	70.0 ± 7.8	64.7 ± 8.0	0.000*
Male	120 (87.6)	51 (92.7)	69 (84.1)	0.135
FEV <sub>1</sub> , %	50.2 ± 16.5	46.0 ± 16.5	53.0 ± 16.0	0.015*
Airflow limitation				
Mild	5 (3.6)	2 (4.0)	3 (6.0)	0.074†
Moderate	71 (51.8)	23 (32.4)	48 (67.6)	
Severe	50 (36.5)	22 (44)	28 (56)	
Very severe	11 (8.0)	8 (72.7)	3 (27.3)	
mMRC	1.4 ± 0.8	1.6 ± 0.9	1.3 ± 0.8	0.013*
≥2	49 (35.8)	29 (52.7)	20 (24.4)	0.011†
HAD score (depression scale)				
≥8	4.5 ± 3.8	5.4 (3.9)	3.9 (3.6)	0.488*
≥8	26 (19.0)	12 (21.8)	14 (17.1)	
Test measures				
5STS test (in s)				
0: unable	14.2 ± 3.1	15.6 ± 3.2	13.3 ± 2.8	0.000*
1: >16.6	9 (6.6)	7 (87.5)	1 (12.5)	0.000†
2: 16.6–13.7	29 (21.2)	20 (71.4)	8 (28.6)	
3: 13.6–11.2	37 (27.0)	15 (41.7)	21 (58.3)	
4: <11.2	38 (27.7)	9 (23.7)	29 (76.3)	
4MGS test (in s)	4.7 ± 2.7	4.9 ± 1.1	4.6 ± 3.4	0.470*
0: unable	0	0 (0.0)	0 (0.0)	0.000†
1: >8.70	1 (0.7)	0 (0.0)	0 (0.0)	
2: 8.70–6.21	9 (6.6)	9 (100.0)	0 (0.0)	
3: 6.20–4.82	35 (25.5)	20 (60.6)	13 (39.4)	
4: <4.82	92 (67.2)	26 (28.3)	66 (71.7)	
6MWT, m	349.1 ± 84.7	271.3 ± 69.8	403.3 ± 40.1	0.001*

% = percentage predicted, 4MGS test = 4-m gait-speed test, 5STS test = 5-repetition sit-to-stand test, 6MWT = 6-minute walk test, FEV<sub>1</sub> = forced expiratory volume in 1 second, HAD score = Hospital Anxiety and Depression score, mMRC = Modified Medical Research Council.

\*Independent *t* test.

†Pearson chi-square test.



**Figure 1.** Receiver-operating characteristic curves comparing the 5STS and 4MGS tests alone (using continuous and ordinal data), with the predicted probability that combines each test and its covariates, in determining poor exercise tolerance. 4MGS test=4-m gait-speed test, 5STS test=5-repetition sit-to-stand test.

to severe COPD. Of our cohort, 35.8% scored grade 2, or higher, on the mMRC scale, with 19% reporting symptoms of depression (HADS  $\geq 8$ ). Three values were missing in our 6MWT dataset (2.18%).

A total of 55 patients (41.05%) had a 6MWT  $< 350$  m (mean  $271.3 \pm 69.8$ ), with the remaining 79 patients (58.95%) demonstrating a 6MWT of  $\geq 350$  m (mean  $403.3 \pm 40.1$ ). Participants with poor 6MWT performance were older, with a higher proportion of mMRC scores of 2, or higher (Table 1). Additionally, patients with poor 6MWT test results had lower mean FEV<sub>1</sub> scores. No significant differences were found with the inclusion of depression as a covariate.

### 3.2. Discriminatory accuracy of the tests

The ROC curves (Fig. 1) illustrated that the AUC for the 5STS test alone (in seconds) was 0.711 (95% confidence interval [CI] 0.613–0.809), and, when combined with significant covariates (age, CAT score, and QMS), increased slightly to 0.729 (95% CI 0.632–0.825). Using ordinal data, the AUC increased slightly from 0.732 (95% CI 0.645–0.819) for the test alone, to 0.767 (95% CI 0.681–0.852) for the “combination” score. The cut-off scores yielding the most accurate discrimination of a poor 6MWT using the test alone were 2 for ordinal measurements, and 13 seconds for continuous data. The values of Se, Sp, and LR, for these cut-offs, and combination scores, are shown in the Table 2.

**Table 2**

**Accuracy of 5STS\* and 4MGS† scores, and of their predicted probabilities that combine each test and its covariates.**

Measure	5STS test				4MGS test			
	Continuous data		Ordinal data		Continuous data		Ordinal data	
	Test alone	Combination score	Test alone	Combination score	Test alone	Combination score	Test alone	Combination score
Sensitivity	75.6	55.6	76.4	63.6	79.6	50	52.7	50.9
Specificity	45.9	86.5	49.6	86.1	49.4	83.1	81.0	83.5
LR+	1.4	4.1	1.9	4.6	1.6	2.9	2.8	3.1
1/LR–	1.9	1.9	2.1	2.4	2.4	1.7	1.7	1.7

4MGS test=4-m gait-speed test, 5STS test=5-repetition sit-to-stand test, LR+=positive likelihood ratio, LR–=negative likelihood ratio.

\* The cut-off scores were 2 for ordinal measurements and 13seconds for continuous data.

† The cut-off scores were 3 for ordinal measurements and 4seconds for continuous data.

**Table 3****Adjusted multivariate logistic regression analysis of determinants of poor exercise tolerance.**

Clinical determinants	Model 1	Model 2
Age	1.05 (1.00, 1.12)	1.07 (1.03, 1.13)*
Dyspnea (mMRC)	1.95 (0.78, 4.85)	2.11 (0.88, 5.04)
Airflow limitation	1.48 (0.75, 2.92)	1.92 (1.00, 3.69)*
4MGS test	—	1.52 (0.97, 2.40)
5STS test	1.23 (1.05, 1.44)*	—

The symbol '—' indicates not included in the model.

4MGS test=4-m gait-speed test, 5STS test=5-repetition sit-to-stand test, mMRC=modified Medical Research Council.

\* $P < 0.05$ .

The AUC for the 4MGS test, when used alone (in seconds), was 0.719 (95% CI 0.629–0.809), and in combination, 0.731 (95% CI 0.643–0.819). Test accuracy was worse when using ordinal data, with an increase of accuracy >5% when using the combination score (using the covariates of age and symptoms of depression). The solo test yielded an AUC of 0.676 (65% CI 0.580–0.771), that is, poor, which increased to 0.726 (95% CI 0.637–0.816) in the combined score. The values of Se, Sp, and LRs are showed in Table 2 for cut-off scores of 3, and 4 seconds for ordinal and continuous data, respectively.

### 3.3. The value of each test as a clinical determinant

Results of multivariate models for determinants of a poor 6MWT performance are shown in Table 3. Model 1 shows that the 5STS test was a strong clinical determinant of a poor 6MWT, even after adjustment for other relevant clinical determinants. Using this model, only the 5STS test remained significant, recording an odds ratio (OR) of 1.23 (95% CI 1.05–1.44). This indicates that for every 1-second increase in the 5STS test result, a 23% increase (1–1.23) in the odds of recording a poor 6MWT performance is expected, as controlled for by other factors in the model. Model 2 shows that the 4MGS test was, by itself, not a significant predictor of performance. For this test, only the covariates of age and airflow limitation were retained.

## 4. Discussion

In this study, we examined whether the 5STS and 4MGS tests might be useful as markers and/or clinical determinants of poor exercise performance. We now report 2 important findings. First, continuous scores for the 4MGS and 5STS tests, and especially, ordinal scores for the 5STS test, were accurate tests with which to predict a poor 6MWT performance. Second, the 5STS test is a significant clinical determinant of poor performance (<350 m). To our knowledge, this is the first study to compare the accuracy with which these 2 tests can predict a poor 6MWT.

Our findings that ordinal scores of 2 or less collected using the 5STS test, and a time over 13 seconds or more, can usefully discriminate poor 6MWT performance, are novel in this field. Prior research has primarily used values based on anthropometric indicators<sup>[26]</sup> and gait speed.<sup>[8,11]</sup> In relation to the 4MGS test, our cut-off value of 4 seconds (or 1 m/s) is consistent with previous studies.<sup>[27]</sup>

The 4MGS test conducted using ordinal data could not accurately predict performance (AUC <0.7), although accuracy was strongly and positively affected by the inclusion of age and depression covariates. When, as in this case, the accuracy of a test depends on its covariates, it is commonplace to adjust for these

variables in the statistical analyses.<sup>[24]</sup> However, to our knowledge, this study is the first to assess the influence of covariates on test accuracy using the gait-speed test for COPD patients. Although adjustment for covariates is routine in other medical areas, this issue is still not well-appreciated for functional tests in COPD.

Our data showing that the accuracy of the 5STS test was unaffected by covariates, suggest that this test can be used alone, or adjusted. This finding has implications for situations in which one may choose to use the test alone, or use the combination score. When choosing between the 2 options, the duration of testing may be a factor. In settings where there are time constraints, a stand-alone test may be more practicable given that combination scores would necessitate QMS and CAT scores. On the contrary, discriminating between true and false cases may be a greater priority. It is important to note that our study showed that using the test alone, with ordinal data, resulted in higher Se and LR– values than the combination score, whereas the combination score generated higher Sp and LR+ values. Given these findings, the clinician should establish, before testing, the consequences (e.g., additional costs for care and rehabilitation) of diagnosing a poor 6MWT, for a patient with good exercise capacity, rather than a good performance for a patient with a poor 6MWT (where failure to initiate early intervention may be critical factor). So, for clinicians with an imperative to avoid false-negatives, the test alone may be more convenient.

The results of this study also showed that the 5STS test was a significant clinical determinant of poor 6MWT, unlike the 4MGS test. Our finding for the 4MGS test was unexpected given a previous study,<sup>[11]</sup> which came to the opposite conclusion. A possible reason for the higher relevance of the 5STS test could be the extra respiratory demand required in this sit-to-stand test.<sup>[10]</sup> Surprisingly, 3 clinical determinants identified by Spruit et al<sup>[7]</sup> (age, airflow limitation, and dyspnea) were found to be insignificant in our model for the 5STS test. The closer approximation of this test to the 6MWT may explain the reduced influence of these pulmonary and nonpulmonary factors.

A valuable aspect of this study was its inclusion of covariates to determine the accuracy of testing. Although adjustment for covariates is common in therapeutic and etiologic studies, the issue of covariate effects has received little attention in the development of markers for screening or diagnosis.<sup>[24]</sup> Our study now addresses this issue, and in doing so, has refined our thinking as to how to use functional tests in COPD patients. Our use of 2 kinds of data (ordinal and continuous) to explore the discriminatory accuracy of the 5STS and 4MGS tests is an additional strength in our approach, and the consistency of our results, using both kinds of data, reinforces our findings.

Our study had several limitations. First, our sample was recruited from a public Spanish hospital; our data may not apply equally to other settings. Therefore, until further research is conducted using a broader patient cohort, these results should be interpreted with caution. Second, we implemented a cross-sectional design in our study; our data must now be confirmed using additional prospective studies. Third, although the identified covariates for the 2 tests were relevant in terms of discriminatory accuracy, it may be speculated that other factors, not included in this study (e.g., disability, physical activity, body mass index), could also impact accuracy.

In summary, the present study shows that 5STS scores can be used as a robust clinical determinant of a poor 6MWT, and a marker with which to classify poor performance for COPD patients, especially when using ordinal data. The 4MGS test score

alone is also a good marker when measured using continuous data. Covariate effects on accuracy were only relevant for the 4MGS test, when measured using ordinal data. This information could be useful for the selection and interpretation of these tests, so that appropriate care and/or preventative interventions can be provided. Further research should focus on whether changes in test data over time (using either test) can increase their usefulness.

## Acknowledgments

The authors wish to thank their patients, and the personnel of the hospital unit, for their cooperation during the course of this study.

## References

- [1] Agusti A, MacNee W. The COPD control panel: towards personalised medicine in COPD. *Thorax* 2013;68:687–90.
- [2] Bossenbroek L, deGreef MH, Wempe JB, et al. Daily physical activity in patients with chronic obstructive pulmonary disease: a systematic review. *COPD* 2011;8:306–19.
- [3] Cote CG, Casanova C, Marin JM, et al. Validation and comparison of reference equations for the 6-min walk distance test. *Eur Respir J* 2008;31:571–8.
- [4] Vestbo J, Hurd SS, Agusti AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: gold executive summary. *Am J Respir Crit Care Med* 2013;187:347–65.
- [5] Casanova C, Cote CG, Marin JM, et al. The 6-min walking distance: long-term follow up in patients with COPD. *Eur Respir J* 2007;29:535–40.
- [6] Cote CG, Pinto-Plata V, Kasprzyk K, et al. The 6-min walk distance, peak oxygen uptake, and mortality in COPD. *Chest* 2007;132:1778–85.
- [7] Spruit MA, Watkins ML, Edwards LD, et al. Determinants of poor 6-min walking distance in patients with COPD: The ECLIPSE cohort. *Respir Med* 2010;104:849–57.
- [8] Karpman C, Benzo RP. Gait speed as a measure of functional status in COPD patients. *Int J Chron Obstruct Pulmon Dis* 2014;9:1315–20.
- [9] Pessoa BV, Arcuri JF, Labadessa IG, et al. Validity of the six-minute step test of free cadence in patients with chronic obstructive pulmonary disease. *Braz J Phys Ther* 2014;18:228–36.
- [10] Ozalevli S, Ozden A, Itilb O, et al. Comparison of the sit-to-stand test with 6 min walk test in patients with chronic obstructive pulmonary disease. *Respir Med* 2007;101:286–93.
- [11] Karpman C, Depew ZS, Lebrasseur NK, et al. Determinants of gait speed in chronic obstructive lung disease. *Chest* 2014;146:104–10.
- [12] Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994;49:85–94.
- [13] Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2015. Available at: <http://www.goldcopd.org/>. Accessed November, 2015.
- [14] ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;166:111–7.
- [15] Guralnik JM. Assessing Physical Performance in the Older Patient [CD-ROM]. Bethesda, MD: National Institutes of Aging. <http://www.grc.nia.nih.gov/branches/leps/sppb/>. Accessed November, 2015.
- [16] Patel MS, Mohan D, Andersson YM, et al. Phenotypic characteristics associated with reduced short physical performance battery score in COPD. *Chest* 2014;145:1016–24.
- [17] O'Shea SD, Taylor NF, Paratz J. Measuring muscle strength for people with chronic obstructive pulmonary disease: retest reliability of hand-held dynamometry. *Arch Phys Med Rehabil* 2007;88:32–6.
- [18] Pinto LM, Gupta N, Tan W, et al. Derivation of normative data for the COPD assessment test (CAT). *Respir Res* 2014;15:68–76.
- [19] Bestall JC, Paul EA, Garrod R, et al. Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54:581–6.
- [20] Dowson C, Laing R, Barraclough R, et al. The use of the Hospital Anxiety and Depression Scale (HADS) in patients with chronic obstructive pulmonary disease: a pilot study. *N Z Med J* 2001;114:447–9.
- [21] Bjelland I, Dahl AA, Haug TT, et al. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002;52:69–77.
- [22] Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159:179–87.
- [23] Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. Hoboken, NJ, USA: John Wiley & Sons; 2000.
- [24] Janes H, Pepe MS. Adjusting for covariates in studies of diagnostic, screening, or prognostic markers: an old concept in a new setting. *Am J Epidemiol* 2008;168:89–97.
- [25] Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 6th ed. Boston: Pearson; 2013.
- [26] Ho SC, Hsu MF, Kuo HP, et al. The relationship between anthropometric indicators and walking distance in patients with chronic obstructive pulmonary disease. *Intern J COPD* 2015;10:1857–62.
- [27] Cesari M, Kritchevsky SB, Penninx BWHJ, et al. Prognostic value of usual gait speed in well-functioning older people: results from the Health, Aging and Body Composition Study. *J Am Geriatr Soc* 2005;53:1675–80.