ERJ Express. **Published on March 21, 2013 as doi: 10.1183/09031936.00131612**

Simple functional performance tests and mortality in COPD

Milo A. Puhan, MD, PhD [1,2], Lara Siebeling[3], MD, Marco Zoller[4], MD, Patrick Muggensturm[2], MD, Gerben ter Riet, MD, PhD [3]

Authors' affiliations:

1 Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, USA

2 Horten Centre for patient-oriented research and knowledge transfer, University of Zurich, Switzerland

3 Department of General Practice, University of Amsterdam, Netherlands

4 Department of General Practice, University of Zurich, Switzerland

Corresponding author:

Milo Puhan, MD, PhD

Johns Hopkins Bloomberg School of Public Health

615 North Wolfe Street, Mail room E6153

Baltimore, MD 21205 USA

phone: 443-287-8777

fax: 410-502-4621

mpuhan@jhsph.edu

Abstract

Exercise tests are important to characterize COPD patients and predict their prognosis but often not available outside of rehabilitation or research settings. The aim was to assess the predictive performance of the sit-to-stand (STS) and handgrip strength tests.

The prospective cohort study in Dutch and Swiss primary care settings included a broad spectrum of patients (n=409) with GOLD stages II to IV. To assess the association of the tests with outcomes, we used Cox proportional hazards (mortality), negative binomial (centrally adjudicated exacerbations) and mixed linear regression models (longitudinal health-related quality of life [HRQL]) while adjusting for age, sex and severity of disease.

The STS test was strongly (adjusted hazard ratio per 5 more repetitions of 0.58 [95% CI 0.40-0.85], p=0.004) and the handgrip strength test moderately strongly (0.84 [95% CI 0.72-1.00], p=0.04) associated with mortality. Both tests were also statistically significantly associated with HRQL but not with exacerbations. The STS test alone was a stronger predictor of 2-year mortality (area under curve 0.78) than body mass index (0.52), FEV1 (0.61), dyspnea (0.63) and handgrip strength (0.62).

The STS test may close an important gap in the evaluation of exercise capacity and prognosis of COPD patients across practice settings.

Study registration

ClinicalTrials.gov (NCT00706602)

Keywords (MeSH terms): Pulmonary Disease, Chronic Obstructive; Exercise capacity; Mortality; Prognosis

Short sentence

The one minute sit-to-stand test predicts mortality in COPD patients and can easily be implemented across practice settings.

Introduction

Numerous studies have shown strong associations of exercise capacity with mortality, exacerbations and health-related quality of life (HRQL) in patients with chronic obstructive pulmonary disease (COPD).[1-5] In fact, exercise capacity is among the strongest predictors of mortality and showed consistently stronger associations than lung function or dyspnea.[2, 4] The BODE index (body mass index, FEV_1, dyspnea and 6-minute walk distance) includes exercise capacity to predict mortality.[3] Multivariable indices inform patients about their prognosis and help them understand what the diagnosis of COPD actually means for them. Also, predictors of outcomes such as exercise capacity help estimating the risk of future outcomes, the absolute effects of treatments for individuals and, thereby, the benefits and harms of treatments (e.g. number-needed-to-treat). Therefore, assessment of prognosis is of great importance not only for patients but also for policy makers, regulatory agencies and clinical guideline developers.

Unfortunately, measurement of exercise capacity has not been implemented in most clinical practice settings relevant to COPD patients. Exercise capacity has probably rarely been tested in the vast majority of COPD patients despite of its great importance. While Six-minute walk, constant work load or incremental exercise tests are commonly used in pulmonary rehabilitation settings, they are not often used in inpatient, or outpatient specialist or primary care settings. An important barrier is that the established Six-minute walk and constant or incremental work load tests require trained staff, space and equipment[6], which is not commonly available in many practice settings.[7]

Over the last ten years, a significant amount of research has been done to explore simpler tests such as sit-to-stand (STS) or step tests to measure exercise capacity or the handgrip strength test to measure upper limb strength. The STS, where the number of repetitions during 30 seconds or one minute are counted, yields reproducible results in elderly people (intraclass correlation coefficients ≥0.84) and the STS test showed high correlations with established tests for exercise capacity such as Six-minute walk distance ($r>0.7$).[8-10] The STS and Six-minute walk tests showed very similar correlations with other validation measures indicating that they may measure similar aspects of functional exercise capacity.[8]

However, we were unable to find studies that directly associated simple exercise tests with clinical outcomes. If STS or handgrip strength tests are to be used in practice to assess health status and associated prognosis, or implemented in prognostic indices, a validation of their predictive properties is needed. Such a direct validation provides stronger evidence for predictive properties than studies, which show high correlations of STS with commonly used exercise tests and indirectly suggests a link of simple exercise tests to clinical outcomes. Therefore, our aim was to assess if and how strongly the STS and handgrip strength tests are associated with mortality, exacerbations and longitudinal HRQL in a diverse population of COPD patients.

Methods

Study design and population

We based our analyses on a prospective multicenter cohort study with COPD patients from primary care in Switzerland and the Netherlands.[11, 12] At inclusion (April 2008 to August 2009), all patients (≥40 years of age) had GOLD stages II–IV and had been free of exacerbations for ≥4 weeks. The only exclusion criteria were life expectancy of <12 months, dementia or psychotic morbidity. All included patients have provided written informed consent. The study has been approved of by all local ethics committees and is registered on ClinicalTrials.gov (NCT00706602).

1-minute STS and handgrip strength test

Using the STS test protocol[10], trained study nurses asked patients to sit down on a chair (height 46-48cm) without arm rests, keep their legs apart with about 90 degrees knee flexion and aligned with their hips, and to hold their hands stationary on their hips. Patients were asked to stand up and to sit down once or twice in order to familiarize them with the task and to assess its feasibility and safety. Study nurses instructed patients about the duration of the test (1 minute) and to do as many repetitions as possible at a self-paced speed allowing for short breaks if needed but without using the arms for support. Study nurses started the test by giving the command “attention, ready, go”. When 15 seconds were left patients were told “You have 15 seconds left until the test is over”.

To measure handgrip strength we used the Jamar® Hydraulic Hand Dynamometer (JA Preston Corporation, Jackson, MI), which is widely used and serves as reference for evaluation of other devices.[13] Patients were seated with the shoulders adducted, elbows flexed to 90˚ and forearms in neutral position. Trained study nurses then instructed patients to squeeze the handle as much as possible and read to the nearest kg where the needle stopped. As recommended we used the best of the six measurements (three times each hand) for the statistical analyses.[13]

Outcome measurements

We focused on three patient-important outcomes, namely, mortality, exacerbations and COPD-specific HRQL. Deaths were captured at the biannual follow-up assessments (at 6, 12, 18 and 24 months) and exact dates of death were ascertained from the treating primary care doctors. We focused on all-cause death since determination of causes of death is prone to misclassification and difficult to perform, even if centrally adjudicated.[14]

Potential exacerbations were captured through patient interviews (biannual assessments) and through review of all patient records two years after enrolment. In both Switzerland and the Netherlands, primary care doctors play a central role in coordinating patient care and reports from other health care providers (e.g. hospitals, pulmonologists) are sent to the primary care doctors. These reports as well as the primary care doctors' case records were available to us. All potential exacerbations were centrally adjudicated by committees that comprised three independent experts (general practitioners with special expertise in pulmonary medicine and pulmonologists) in Switzerland and four in the Netherlands. The experts first independently reviewed the documents and determined the number and dates of exacerbations for each patient. We used an event-based definition for exacerbations with two criteria that had to be fulfilled: (1) unscheduled physician contact in a hospital, private practice or by telephone for worsening of dyspnea, cough, increased sputum production or a change in sputum color and (2) electronic or hand-written documentation of new prescription or a dosage increase of systemic steroids or new prescription of an antibiotic. The experts sent their reviews to the Swiss and Dutch study coordinating centers. At the consensus meetings of the two adjudication committees the members compared their independent reviews on

those patients for whom discordances existed and reached a consensus on the number and dates of exacerbations. More details about the adjudication process will be reported elsewhere.

To assess COPD-specific health-related quality of life, we used the validated German and Dutch self-administered, standardized versions of the Chronic Respiratory Questionnaire (CRQ). The CRQ provides scores in four different domains (dyspnea, fatigue, emotional, and mastery) each on a scale of one to seven, where one is the worst score indicating very poor HRQL and seven the best.[15, 16] Patients completed the CRQ at baseline and biannually up to the 2-year visit.

Statistical analysis

We used regression analyses with the outcomes as dependent and the exercise tests as independent variables (separate analysis for each outcome and for STS and handgrip strength tests, respectively). We used Cox proportional hazards models to assess the association between the exercise tests and mortality (up to two years). Since exacerbations are usually correlated within patients we used negative binomial regression models to assess the association between the exercise tests and the number of exacerbations (incidence rate ratio). To assess the association between the exercise tests and HRQL from baseline up to two years we used random-effects linear regression models that took into account the within-person (repeated measurements) and between person variability. We adjusted the analyses for age, sex, baseline dyspnea, baseline FEV1 in % predicted and the use of long-acting bronchodilators ± inhaled corticosteroids (at baseline). In sensitivity analyses we imputed a zero for patients who did not perform the STS test at baseline. Finally, we calculated the area under the curve for a number of commonly used predictors such as FEV1 in % predicted, age, dyspnea, body mass index and their combinations.

We did not have any missing variables for the handgrip strength test or the confounders. Not all patients were able to complete the STS test and we did sensitivity analyses as described above. We did not perform formal sample size calculations for this analysis but we have described details for sample size calculations for this cohort before.[11] We conducted all analyses using Stata for Windows (version 11.2, College Station (TX), USA).

Results

Study population

Detailed characteristics of the broad study population of 409 COPD patients can be found elsewhere.[12] In brief, at enrolment mean age was 67.3 years (SD 10.0), 57% were male, 63.8% had GOLD stage II, 21.8% stage III and 14.4% stage IV, respectively, and mean degree of dyspnea on the MRC scale was 1.87 (SD 1.46). 38 patients (9.3%) had died within the first 2 years of observation (total observation time of 782 person years). Only 19 (4.6%) patients dropped out of the study because of worsening of physical condition (n=4), overburden (n=4), psychiatric condition (n=3), lost contact (n=1) or other reasons (n=7). 204 patients (49.9%) had no exacerbation, 101 (24.7%) had 1, 33 (8.1%) had 2 and 71 patients (17.4%) had three or more exacerbations during the first 2 years of observation. 93.3% of the adjudicated exacerbations were treated on an outpatient and 6.7% on an inpatient basis.

STS and handgrip strength test results at baseline

374 (91.4%) patients were able to perform the STS test at baseline. Reasons for not performing a STS were musculoskeletal problems (osteoarthritis, back pain or recent orthopedic surgery). The mean number of repetitions during the STS test was 18.9 (SD 8.8). STS test results were lower for women compared to men, for patients with higher GOLD stage, for patients experiencing more dyspnea and for patients with more comorbidities (Table 1). All patients were able to perform the handgrip strength test at baseline. The mean maximum handgrip strength was 35.8 kg (SD 12.8). There was high reproducibility across the three attempts on the right side (means of 33.4, 32.9 and 32.6 kg on first, second and third attempt) and left hand (means of 30.8, 30.7 and 30.5 kg) with intraclass correlation coefficients of 0.88 for both sides. Again, test results were lower for women, in patients with more dyspnea and more comorbidities, whereas patients with higher GOLD stages did not have lower upper limb strength. The correlation between the number of repetitions during the STS test and maximal handgrip strength was 0.45.

Association of the STS test with mortality, exacerbations and HRQL

The STS test at baseline was substantially lower in patients who died subsequently than in patients who were alive at two years (11.8 vs. 19.5 repetitions, Table 2). We found a strong

association of the STS test with mortality. The adjusted analysis showed a hazard ratio of 0.90 per one more repetition (95% CI 0.83-0.97 p=0.004), equivalent to a hazard ratio of 0.58 (95% CI 0.40 to 0.85 p=0.004) per five more repetitions (Table 2).

We did not find an association of the STS test with exacerbations (incidence rate ratio per 1 more repetition close to 1.0). But we found statistically significant associations of the STS test with all four domains of the CRQ, as assessed over 24 months. The adjusted effect of five more repetitions during the STS test was 0.26 (95% CI 0.19-0.34) for the dyspnea, 0.19 (95% CI 0.13-0.26) for the fatigue, 0.10 (95% CI 0.04-0.16) for the emotional function and 0.08 (95% CI 0.02-0.13) for the mastery domain. Six patients completed only four repetitions during the STS test and two patients 3 and 1 repetitions, respectively. Excluding these patients from the analyses led to almost identical estimates as in the main analyses. Finally, in additional sensitivity analyses we imputed a zero for those patients who were unable to perform the STS test at baseline (n=35). The analysis for the complete set of patients (n=409) showed almost identical results as those in the main analysis.

Association of the handgrip strength with mortality, exacerbations and HRQL

Handgrip strength at baseline was substantially lower in patients who died subsequently than in patients who were alive at two years (30.1 kg vs. 36.4 kg, Table 3). We also found a statistically significant association of handgrip strength with mortality. The adjusted analysis showed a hazard ratio of 0.97 per one more kg (95% CI 0.94-0.99 p=0.04) equivalent to a hazard ratio of 0.84 (95% CI 0.72 to 1.00, p=0.04) per five more kg (Table 3).

As for the STS test we did not find an association of handgrip strength with exacerbations. Handgrip strength was statistically significantly associated with three CRQ domains but not with mastery. As for the STS test, we found the strongest association of handgrip strength with dyspnea (adjusted effect of five more kg 0.14, 95% CI 0.09-0.20, followed by fatigue (0.12, 95% CI 0.08-0.17), emotional function (0.06, 95% CI 0.02-0.10) and mastery (0.01, 95% CI -0.03-0.05).

Comparison of prediction of mortality of STS and handgrip strength test with established predictors

The ADO index, the STS test and combinations of both indices with the STS test were strong predictors of 2-year mortality with areas under the curve of around 0.8 (Figure 1). The STS test alone predicted 2-year mortality almost as well as the ADO index. Handgrip strength and the BODE index without an exercise test were both much weaker predictors of 2-year mortality but a BODE index with the STS test instead of Six-minute walk distance showed excellent prediction of 2-year mortality. Age alone was also a good predictor whereas all other predictors included in the BODE or ADO indices were poor or moderate predictors when used alone.

Discussion

We found that upper limb strength measured by the handgrip test and, in particular, the one-minute STS test as a measure of exercise capacity are strongly and independently associated with mortality and HRQL over 24 months of observation but we did not find significant associations with exacerbations. The STS test alone is as good a predictor of mortality as the ADO index and may replace BODE's Six-minute walk distance if latter is not available.

Tables 1 and 2, where we reported the full regression models, indicate that the STS test was a stronger predictor of mortality than FEV1 % predicted, dyspnea or the use of inhaled drugs. The strong predictive properties are also reflected by the area under the curve of 0.78 (Figure). Earlier studies using the Six-minute walk or incremental exercise tests show similar results.[1, 2, 5, 17] Upper limb strength was less strongly but still statistically significantly associated with mortality but the area under the curve (0.62) suggests it may be of limited use to predict mortality in COPD patients. Although direct comparisons of the predictive properties of exercise tests are scarce the current body of evidence suggests that the simpler STS test and the Six-minute walk or incremental exercise tests can be used to predict outcomes in COPD.

In most practice settings, exercise tests are uncommonly performed in COPD patients because these tests require equipment, space and trained staff. In addition, reimbursement for the Six-minute walk test is not always guaranteed. We believe that the STS test offers an attractive alternative. The STS test, which, according to our protocol, lasts one minute, and appears to measure similar aspects of exercise capacity as the Six-minute walk test.[8, 9] It may also be responsive to change[18]. Its simplicity makes it an ideal test for primary but also for acute care

settings.[19] The handgrip strength test requires a handgrip dynamometer (one time investment of about $300) but it is very simple to perform and most, in our study all, patients are able to perform the test. But we think that this test alone is not sufficient to inform patients and physicians about exercise capacity because it does not test large or multiple muscles as do other tests, which is likely the reason for not being a strong predictor of prognosis. But handgrip strength might serve as a first test to identify patients with low upper limb strength who may then undergo more extensive testing of exercise capacity.

Some questions about the usefulness of the STS and handgrip strength tests remain. We did not find a significant association of the two tests with exacerbations. A possible reason is that patients in our cohort rarely experienced severe exacerbations requiring hospital admission. Given the strong association of the tests with mortality it is possible that in patients with more severe COPD who suffer from severe exacerbations a significant association may exist. In the present analyses the number of hospital admissions was too low to restrict the analysis to severe exacerbations but as soon as more years of follow-up are available this or other studies should assess the association of the STS and handgrip strength tests with severe exacerbations. Additional studies may also directly compare the predictive properties of the STS and Six-minute walk test and explore if the STS would be contribute as much to the BODE index as Six-minute walk distance. Finally, the goal of our study was to assess the predictive properties of the STS and handgrip strength tests but we did not evaluate their properties when used as outcome measures. These simple tests may be used in randomized trials as additional outcome measures. However, before tests like the STS or handgrip strength test are routinely used in trials their responsiveness to change and minimal important difference need to be established as it has been done for Six-minute walk distance.[20]

Strengths of this study include a carefully developed study protocol[11] that we implemented successfully as illustrated by the completeness of data collection and very low number of drop-outs. The large and broad cohort from primary care is another strength and we think that our results are applicable to other populations because our cohort is likely to reflect a typical primary care COPD population. A limitation of this study is that we did not directly compare the predictive properties of the STS and handgrip strength test with those of the Six-minute walk or other tests.

We restricted our study to measurements that are feasible in a primary care setting. Another limitation is that we did not ascertain mild exacerbations, which we defined as worsening of symptoms that did not lead to the prescription of antibiotics or systemic corticosteroids. At the design stage of this cohort study we decided to focus on moderate to severe exacerbations because of their great impact on the patients' HRQL and prognosis and because a central adjudication would be very challenging for mild exacerbations.

In conclusion, we found in this prospective cohort study of primary care COPD patients that the STS test as a measure of exercise capacity is a strong predictor of mortality and HRQL and may be an attractive option to assess exercise capacity in COPD patients in both clinical practice and research. The availability of such a simple test may close an important gap in the evaluation of the prognosis of COPD patients across different settings.

Acknowledgements

We thank Ursula Schafroth (Horten Centre for patient-oriented research, University of Zurich, Switzerland) and Alice Karsten (Academic Medical Centre, Department of General Practice, University of Amsterdam, the Netherlands) and the participating general practitioners and COPD patients in Switzerland and the Netherlands (Stichting Gezondheidscentra Amsterdam Zuidoost and Zorggroep Almere) who make this study possible by their enthusiastic participation.

Funding

Swiss National Science Foundation (grant # 3233B0/115216/1)

Dutch Asthma Foundation (grant # 3.4.07.045)

Zurich Lung League (unrestricted grant)

Authors' contributions

MP, MZ and GtR conceived the study idea, contributed to the statistical analysis and drafted the first version of the manuscript. MP, PM and MZ oversaw all activities related to the conduct of the Swiss part of the cohort study. GtR and LS oversaw all activities related to the conduct of the Dutch part of the cohort study. All authors revised the manuscript and accepted the final version.

Conflicts of interest

None of the authors has a conflict of interest to declare in relation to this manuscript.

References

1. Pinto-Plata VM, Cote C, Cabral H, Taylor J, Celli BR. The 6-min walk distance: change over time and value as a predictor of survival in severe COPD. *Eur Respir J* 2004; Jan;23(1):28-33.

2. Marquis K, Debigare R, Lacasse Y, LeBlanc P, Jobin J, Carrier G, et al. Midthigh muscle cross-sectional area is a better predictor of mortality than body mass index in patients with chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine* 2002; Sep 15;166(6):809-13.

3. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *The New England Journal of Medicine* 2004; Mar 4;350(10):1005-12.

4. Puhan MA, Garcia-Aymerich J, Frey M, ter Riet G, Anto JM, Agusti AG, et al. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. *Lancet* 2009; Aug 29;374(9691):704-11.

5. Spruit MA, Polkey MI, Celli B, Edwards LD, Watkins ML, Pinto-Plata V, et al. Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. *J Am Med Dir Assoc* 2012;13(3):291-7.

6. American Thoracic Society. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002;166(1):111-7.

7. Puhan MA, Zoller M, ter Riet G. COPD: more than respiratory. *Lancet* 2008; Jan 5;371(9606):27,8.

8. Bohannon RW. Measurement of Sit-to-Stand Among Older Adults. *Topics in Geriatric Rehabilitation* 2012;28(1):11-16.

9. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Res Q Exerc Sport* 1999; Jun;70(2):113-9.

10. Ozalevli S, Ozden A, Itil O, Akkoclu A. Comparison of the Sit-to-Stand Test with 6 min walk test in patients with chronic obstructive pulmonary disease. *Respir Med* 2007; Feb;101(2):286-93.

11. Siebeling L, ter Riet G, van der Wal WM, Geskus RB, Zoller M, Muggensturm P, et al. ICE COLD ERIC--International collaborative effort on chronic obstructive lung disease: exacerbation

risk index cohorts--study protocol for an international COPD cohort study. *BMC Pulm Med* 2009;9:15.

12. Siebeling L, Puhan MA, Muggensturm P, Zoller M, Ter Riet G. Characteristics of Dutch and Swiss primary care COPD patients - baseline data of the ICE COLD ERIC study. *Clin Epidemiol* 2011:273-83.

13. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. *Age Ageing* 2011;40(4):423-9.

14. McGarvey LP, John M, Anderson JA, Zvarich M, Wise RA. Ascertainment of cause-specific mortality in COPD: operations of the TORCH Clinical Endpoint Committee. *Thorax* 2007; May;62(5):411-5.

15. Puhan MA, Behnke M, Laschke M, Lichtenschopf A, Brandli O, Guyatt GH, et al. Self-administration and standardisation of the chronic respiratory questionnaire: a randomised trial in three German-speaking countries. *Respir Med* 2004; Apr;98(4):342-50.

16. Wijkstra PJ, TenVergert EM, Van Altena R, Otten V, Postma DS, Kraan J, et al. Reliability and validity of the chronic respiratory questionnaire (CRQ). *Thorax* 1994; May;49(5):465-7.

17. Cote CG, Pinto-Plata V, Kasprzyk K, Dordelly LJ, Celli BR. The 6-min walk distance, peak oxygen uptake, and mortality in COPD. *Chest* 2007; Dec;132(6):1778-85.

18. Patel M, Canavan J, Clark A, Ingram K, Fowler R, Polkey M, Man W. The effect of pulmonary rehabilitation on the sit-to-stand test in COPD. Annual Conference of the European Respiratory Society; 2011, Amsterdam. p. P1232.
https://www.ersnetsecure.org/public/prg_congres.abstract?ww_i_presentation=50878

19. Annweiler C, Schott AM, Abellan van Kan G, Rolland Y, Blain H, Fantino B, et al. The Five-Times-Sit-to-Stand test, a marker of global cognitive functioning among community-dwelling older women. *J Nutr Health Aging* 2011;15(4):271-6.

20. Puhan MA, Chandra D, Mosenifar Z, Ries A, Make B, Hansel NN, et al. The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011; Apr;37(4):784-90.

Figure legend

Figure title: Comparison of predictors to predict 2-year mortality in COPD patients

The figure shows the area under the curve (95% CI) for each predictor and for combinations of predictors. An area under the curve of 0.5 indicates prediction no better than chance and 1.0 perfect prediction.

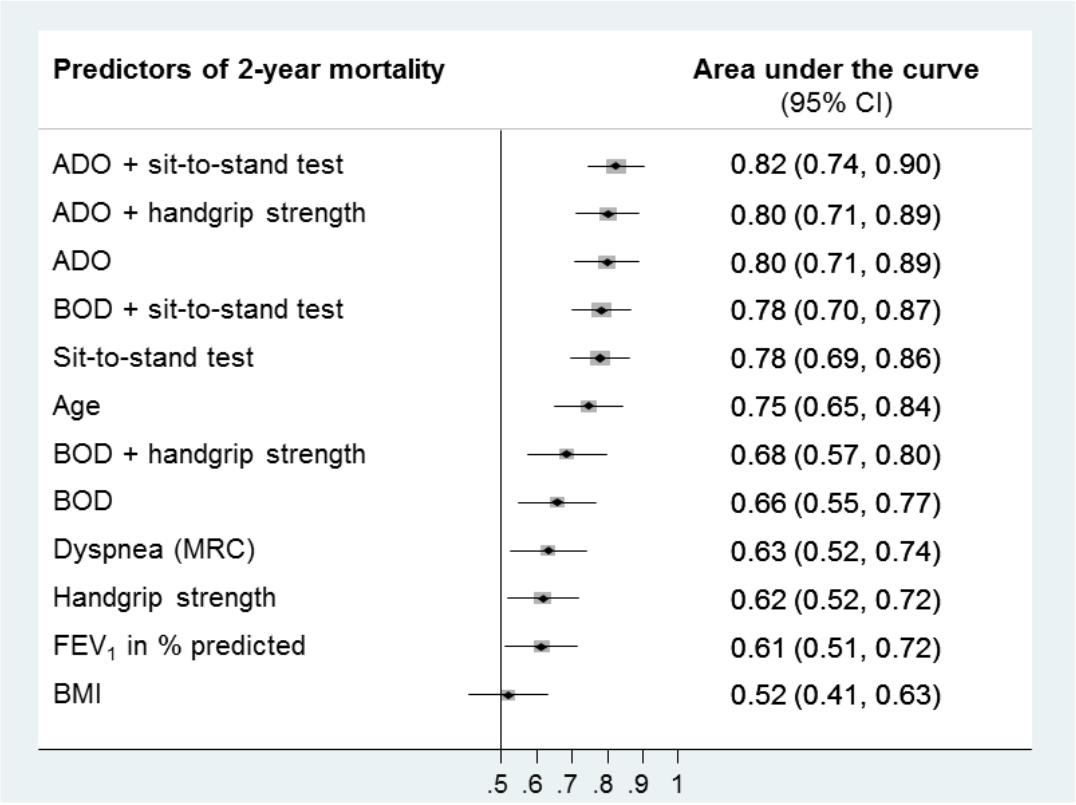

Table 1. Sit-to-Stand test and handgrip strength tests at baseline

		Sit-to-Stand test Mean number of repetitions (SD)	Handgrip strength Mean kg (SD)
All patients	(n1=374; n2=409)[#]	18.9 (8.8)	35.8 (12.1)
Sex			
Male	(n1=212; n2=233)	20.3 (9.5)	42.8 (10.1)
Female	(n1=162; n2=176)	17.2 (7.4)	26.5 (7.3)
GOLD stage			
II	(n1=243; n2=261)	19.2 (8.6)	35.7 (12.3)
III	(n1=80; n2=89)	18.8 (9.7)	34.9 (12.7)
IV	(n1=51; n2=59)	17.9 (8.0)	37.6 (10.3)
MRC dyspnea			
0	(n1=68; n2=71)	24.6 (9.2)	42.7 (12.1)
1	(n1=149; n2=155)	21.3 (8.7)	36.5 (11.8)
2	(n1=40; n2=41)	17.7 (6.2)	36.4 (11.3)
3	(n1=37; n2=42)	13.7 (4.8)	31.5 (10.4)
4	(n1=80; n2=100)	12.8 (5.6)	31.3 (11.1)
Number of comorbidities			
0-2	(n1=168; n2=178)	19.9 (8.6)	37.7 (12.0)
3-4	(n1=112; n2=128)	18.4 (8.3)	34.8 (12.4)
≥5	(n1=94; n2=103)	17.8 (9.5)	33.8 (12.9)

[#] n1 = sample size for results on STS test; n2 = sample size for results on handgrip strength

Table 2. Association of the sit-to-stand test with mortality, exacerbations and health-related quality of life

	Mortality	
	Patients alive at two years	Patients dead at two years
Sit-to-stand test, mean number of repetitions in 1 minute (SD)	19.5 (8.7)	11.8 (6.3)
Association of sit-to-stand test with mortality	Unadjusted association (hazard ratio per 1 more repetition)	0.86 (95% CI 0.81-0.92), p<0.0001, z=-4.73
	Adjusted association (hazard ratio per 1 more repetition) Adjusted association (hazard ratio per 5 more repetitions) Covariates	0.90 (95% CI 0.83-0.97), p=0.004, z=-2.84 0.58 (95% CI 0.40-0.85) p=0.004, z=-2.84 Age 1.08 (SE 0.026, z= 3.17, p=0.002) FEV1 0.98 (SE 0.012, z=-1.76, p=0.078) Dyspnea 1.12 (SE 0.185, z=0.69, p= 0.49) LABA/ICS 0.56 (SE 0.29, z=-1.11, p=0.269)
	Exacerbations	
	Patients with no exacerbation during 2 years of follow-up	Patients with ≥1 exacerbation during 2 years of follow-up
Sit-to-stand test, mean number of repetitions in 1 minute (SD)	19.4 (8.5)	18.1 (8.1)
Association of sit-to-stand test with exacerbations	Unadjusted association (incidence rate ratio per 1 more repetition)	0.99 (95% CI 0.97-1.01), p=0.21, z=-1.26
	Adjusted association (incidence rate ratio per 1 more repetition) Covariates	1.00 (95% CI 0.98-1.02), p=0.94, z=0.08 Age 1.00 (SE 0.007, z=-0.48, p=0.63) FEV1 0.98 (SE 0.005, z=-5.02, p<0.001) Dyspnea 1.06 (SE 0.067, z=0.89, p= 0.38) LABA/ICS 1.52 (SE 0.34, z=1.87, p=0.062)
	Health-related quality of life from baseline to 2-year follow-up	
Association of sit-to-stand test with CRQ dyspnea	Unadjusted difference on scale from 1-7 per 1 more repetition	0.06 (95% CI 0.04-0.07), p<0.001, z=7.55
	Adjusted difference on scale from 1-7 per 1 more repetition Adjusted difference on scale from 1-7 per 5 more repetitions Covariates	0.05 (95% CI 0.04-0.07), p<0.001, z=7.10 0.26 (95% CI 0.19-0.34), p<0.001, z=7.10 Age 0.017 (SE 0.007, z= 2.54, p=0.011)

		FEV1 0.026 (SE 0.004, z=-6.48, p<0.001) LABA/ICS -0.396 (SE 0.16, z=--2.41, p=0.016)
Association of sit-to-stand test with CRQ fatigue	Unadjusted difference on scale from 1-7 per 1 more repetition	0.06 (95% CI 0.04-0.07), p<0.001, z=8.57
	Adjusted difference on scale from 1-7 per 1 more repetition Adjusted difference on scale from 1-7 per 5 more repetitions Covariates	0.04 (95% CI 0.03-0.05), p<0.001, z=6.06 0.19 (95% CI 0.13-0.26), p<0.001, z=6.06 Age 0.007 (SE 0.005, z= 1.34, p=0.18) FEV1 -0.007 (SE 0.005, z=-2.04, p=0.042) Dyspnea 0.361 (SE 0.038, z=9.51, p<0.001) LABA/ICS 0.108 (SE 0.13, z=0.81, p=0.42)
Association of sit-to-stand test with CRQ emotional function	Unadjusted difference on scale from 1-7 per 1 more repetition	0.03 (95% CI 0.02-0.05), p<0.001, z=5.33
	Adjusted difference on scale from 1-7 per 1 more repetition Adjusted difference on scale from 1-7 per 5 more repetitions Covariates	0.02 (95% CI 0.01-0.03), p=0.001, z=3.24 0.10 (95% CI 0.04-0.16), p=0.001, z=3.24 Age 0.013 (SE 0.005, z=2.52, p=0.012) FEV1 -0.009 (SE 0.003, z=-2.81, p=0.005) Dyspnea 0.357 (SE 0.036, z=9.87, p<0.001) LABA/ICS 0.169 (SE 0.13, z=1.32, p=0.19)
Association of sit-to-stand test with CRQ mastery	Unadjusted difference on scale from 1-7 per 1 more repetition	0.03 (95% CI 0.02-0.04), p<0.001, z=5.79
	Adjusted difference on scale from 1-7 per 1 more repetition Adjusted difference on scale from 1-7 per 5 more repetitions Covariates	0.02 (95% CI 0.00-0.03), p=0.005, z=2.84 0.08 (95% CI 0.02-0.13), p=0.005, z=2.84 Age 0.008 (SE 0.005, z= 1.69, p=0.09) FEV1 -0.002 (SE 0.003, z=-0.65, p=0.51) Dyspnea 0.340 (SE 0.033, z=10.41, p<0.001) LABA/ICS -0.053 (SE 0.116, z=-0.46, p=0.64)

z-values are reported to show the relative strength of association of each variable in the multivariable regression models. All regression coefficients for continuous variables except for handgrip strength are per increase of 1 unit (1 year for age, 1 % predicted for FEV1 and 1 point for dyspnea)
LABA: Long-acting beta-agonist; ICS: Inhaled corticosteroids; FEV1: Forced expiratory volume in one second in % predicted; CRQ: Chronic Respiratory Questionnaire

Table 3: Association of handgrip strength with mortality, exacerbations and health-related quality of life

	Mortality	
	Patients alive at two years	Patients dead at two years
Handgrip strength, mean kg (SD)	36.4 (12.2)	30.1 (9.9)
Association of handgrip strength with mortality	Unadjusted association (hazard ratio per 1 more kg)	0.96 (95% CI 0.93-0.98), p=0.002, z=-3.06
	Adjusted association (hazard ratio per 1 more kg)	0.97 (95% CI 0.94-1.00), p=0.04, z=-2.05
	Adjusted association (hazard ratio per 5 more kg)	0.84 (95% CI 0.72-1.00) p=0.04, z=-2.05
	Covariates	Age 1.09 (SE 0.021, z= 4.27, p<0.001) FEV1 0.97 (SE 0.010, z=-3.02, p=0.003) Dyspnea 1.29 (SE 0.167, z=1.97, p= 0.49) LABA/ICS 0.56 (SE 0.26, z=-1.22, p=0.22)
	Exacerbations	
	Patients with no exacerbation during 2 years of follow-up	Patients with ≥1 exacerbation during 2 years of follow-up
Handgrip strength, mean kg (SD)	19.4 (8.5)	18.1 (8.1)
Association of handgrip strength with exacerbations	Unadjusted association (incidence rate ratio per 1 more kg)	0.99 (95% CI 0.98-1.00), p=0.06, z=-1.86
	Adjusted association (incidence rate ratio per 1 more kg)	0.99 (95% CI 0.98-1.00), p=0.14, z=-1.48
	Covariates	Age 0.99 (SE 0.006, z=--0.79, p=0.43) FEV1 0.98 (SE 0.004, z=-5.63, p<0.001) Dyspnea 1.05 (SE 0.055, z=0.99, p= 0.32) LABA/ICS 1.56 (SE 0.34, z=-2.02, p=0.044)
	Health-related quality of life from baseline to 2-year follow-up	
Association of handgrip strength with CRQ dyspnea	Unadjusted difference on scale from 1-7 per 1 more kg	0.03 (95% CI 0.02-0.04), p<0.0001, z=5.05
	Adjusted difference on scale from 1-7 per 1 more kg	0.03 (95% CI 0.02-0.04), p<0.0001, z=5.39
	Adjusted difference on scale from 1-7 per 5 more kg	0.14 (95% CI 0.09-0.20), p<0.0001, z=5.39
	Covariates	Age 0.008 (SE 0.007, z= 1.15, p=0.25) FEV1 0.032 (SE 0.004, z=8.20, p<0.001) LABA/ICS -0.47 (SE 0.17, z=-2.78, p=0.005)

Association of handgrip strength with CRQ fatigue	Unadjusted difference on scale from 1-7 per 1 more kg	0.04 (95% CI 0.03-0.05), p<0.0001, z=7.38
	Adjusted difference on scale from 1-7 per 1 more kg Adjusted difference on scale from 1-7 per 5 more kg Covariates	0.02 (95% CI 0.01-0.03), p<0.0001, z=5.61 0.12 (95% CI 0.08-0.17), p<0.0001, z=5.61 Age 0.003 (SE 0.005, z= 0.52, p=0.60) FEV1 -0.003 (SE 0.003, z=-0.81, p=0.42) Dyspnea 0.401 (SE 0.035, z=11.53, p<0.001) LABA/ICS 0.070 (SE 0.134, z=0.52, p=0.60)
Association of handgrip strength with CRQ emotional function	Unadjusted difference on scale from 1-7 per 1 more kg	0.02 (95% CI 0.01-0.03), p<0.0001, z=4.17
	Adjusted difference on scale from 1-7 per 1 more kg Adjusted difference on scale from 1-7 per 5 more kg Covariates	0.01 (95% CI 0.00-0.02), p=0.006, z=2.75 0.06 (95% CI 0.02-0.10), p=0.006, z=2.75 Age 0.013 (SE 0.005, z= 2.65, p=0.008) FEV1 -0.007 (SE 0.003, z=-2.33, p=0.020) Dyspnea 0.361 (SE 0.033, z=10.90, p<0.001) LABA/ICS 0.141 (SE 0.127, z=1.11, p=0.27)
Association of handgrip strength with CRQ mastery	Unadjusted difference on scale from 1-7 per 1 more kg	0.01 (95% CI 0.00-0.02), p=0.009, z=2.62
	Adjusted difference on scale from 1-7 per 1 more kg Adjusted difference on scale from 1-7 per 5 more kg Covariates	0.00 (95% CI -0.01-0.01), p=0.59, z=0.53 0.01 (95% CI -0.03-0.05), p=0.59, z=0.53 Age 0.008 (SE 0.005, z= 1.73, p=0.08) FEV1 -0.002 (SE 0.003, z=-0.62, p=0.53) Dyspnea 0.364 (SE 0.030, z=11.97, p<0.001) LABA/ICS -0.097 (SE 0.117, z=-0.83, p=0.41)

z-values are reported to show the relative strength of association of each variable in the multivariable regression models. All regression coefficients for continuous variables except for handgrip strength are per increase of 1 unit (1 year for age, 1 % predicted for FEV1 and 1 point for dyspnea)
LABA: Long-acting beta-agonist; ICS: Inhaled corticosteroids; FEV1: Forced expiratory volume in one second in % predicted; CRQ: Chronic Respiratory Questionnaire