Respiratory Medicine 161 (2020) 105856

Contents lists available at ScienceDirect

Respiratory Medicine

journal homepage: http://www.elsevier.com/locate/rmed

Review article

Prognostic value of simple measures of physical function and muscle strength in COPD: A systematic review

Daniela Massierer [a,e,f], Waleed Alsowayan [b], Vanessa Pereira Lima [c], Jean Bourbeau [d,e], Tania Janaudis-Ferreira [a,d,e,f,*]

[a] School of Physical and Occupational Therapy, McGill University, 3654 Prom Sir-William-Osler, Hosmer House, Room #200, Montreal, QC, H3G 1Y5, Canada
[b] Respiratory Division, McGill University Health Center, 1001 Decarie Blvd, Montreal, QC, H4A 3J, Canada
[c] Department of Physical Therapy, Universidade Federal dos Vales do Jequitinhonha e Mucuri-UFVJM, Campus JK, Rodovia MGT 367, Km 583, N° 5000, Alto da Jacuba CEP: 39100-000, Diamantina, Minas Gerais, Brazil
[d] Respiratory Epidemiology and Clinical Research Unit (RECRU), McGill University Health Centre, 5252 Boul. de Maisonneuve Ouest, Montreal, QC, H4A 3S5, Canada
[e] Translational Research in Respiratory Diseases Program, Research Institute of the McGill University Health Centre, 1001 Decarie Blvd, Block E, Montreal, QC, H4A 3J1, Canada
[f] Centre for Health Outcomes Research and Evaluation (CORE), Research Institute of the McGill University Health Centre, 5252 Boul. de Maisonneuve Ouest, Room # 3E.01, Montreal, QC, H4A 3S5, Canada

ARTICLE INFO

Keywords:
Chronic obstructive pulmonary disease
Physical function
Muscle strength
Systematic review

ABSTRACT

Purpose: There has been an increased interest in simple measures of physical function and muscle strength that can be used in all clinical settings to assess individuals with chronic obstructive pulmonary disease (COPD) and predict their prognosis. The main objective was to examine the prognostic value of simple measures of physical function and muscle strength in relation to exacerbation, hospitalization and mortality in individuals with COPD.
Methods: Medline, EMBASE, Cochrane and Web of Science were searched. We included prospective observational studies that examined the prognostic value of simple performed-based tests or self-reported measures of physical function or muscle strength in relation to exacerbation, hospitalization and mortality in individuals with COPD.
Results: Seven articles met the inclusion criteria. The most commonly used tests were the handgrip strength (HGS) (n = 4) and 1-min sit-to-stand (STS) (n = 2). There were considerable variations in terms of characteristics of patients included, setting of recruitment, type of tests used, duration of follow-up and outcome measures of interest. The majority of the studies were classified as having "fair" or "poor" methodological quality.
Conclusions: There is a limited number of studies examining the prognostic value of simple measures of physical function and muscle strength in relation to exacerbations, hospitalizations and mortality in individuals with COPD. To date, the HGS and 1-min STS tests are the most studied tests and seem to be suitable for prognosis purposes in individuals with COPD. However, more studies with better methodological quality are needed to confirm these findings.

1. Background

Chronic Obstructive Pulmonary Disease (COPD) is a major cause of death worldwide and a global healthcare problem [1]. Although the primary pathophysiology of COPD is related to the damage in the lungs, patients with COPD experience secondary non-respiratory related effects of the disease such as lower and upper limb muscle dysfunction and reduced exercise capacity which contribute to functional limitations [2]. These secondary effects of COPD negatively affect symptoms, are strong risk factors for disability [3] in these individuals and can provide relevant information about their prognosis. For example, exercise capacity and quadriceps muscle strength have been shown to be among the strongest predictors of mortality in individuals with COPD [4,5] and their assessment is recommended as part of routine clinical assessment [6,7]. However, assessment of exercise capacity in the laboratory and/or field tests such as Six-Minute Walk Test (6MWT) and Incremental Shuttle Walk Test require trained staff, space and equipment which is rarely feasible in the outpatient clinics. Similarly, proper quadriceps muscle strength assessment requires expensive dynamometers and trained personnel which may not be readily available.

* Corresponding author. 3630 Promenade Sir-William-Osler, Montreal, QC, H3G 1Y5, Canada.
E-mail address: tania.janaudis-ferreira@mcgill.ca (T. Janaudis-Ferreira).

https://doi.org/10.1016/j.rmed.2019.105856
Received 8 October 2019; Received in revised form 17 December 2019; Accepted 18 December 2019
Available online 19 December 2019

In recent years, there has been an increased interest in simple measures of physical function and muscle strength that can be used in all clinical settings to assess individuals with COPD and predict their prognosis [8,9]. These "simpler" tests require less time for the professional and patient to conduct; some require less space; and some have potential suitability in more frail patients [10]. Many of these tests (e.g. the time up and go, sit-to-stand and grip strength tests) are already used and studied in older adults and have shown to be predictors of mortality and healthcare utilization in this population. Previous systematic reviews have described some simple functional exercise tests used in people with COPD and their psychometric properties [10,11]; however, no systematic review has focused on the prognostic value of these tests with respect to important clinical outcomes such as exacerbations, healthcare utilization and mortality.

The objective of this systematic review is to systematically review the literature on the prognostic value of simple measures of physical function and muscle strength in relation to exacerbations, hospitalizations and mortality in individuals with COPD.

2. Methods

2.1. Search strategy and data sources

A librarian performed electronic literature searches in three databases (Medline (Ovid), Embase (Ovid), Cochrane Library (Wiley) and Web of Science) from inception until November 16, 2016, with no limits or language restrictions. An updated search was performed by one investigator (DM) to include articles published to December 2018. The key terms used were related to COPD, physical function measures (both performance-based tests and self-reported measures), muscle strength and hospitalization, exacerbation or mortality. The terms were searched as text words in the Title/Abstract/Keywords as well as in the Subject Headings. The specific search terms and strategy are presented in Appendix A. While performing full-text screening, one investigator (WA) conducted hand searches of the reference list of the articles.

2.2. Inclusion criteria

We included prospective observational studies that, as primary or secondary objective, examined the performance of simple measures of overall physical function and muscle strength in predicting exacerbations, hospitalizations and mortality in patients with COPD or that examined the relationship of these tests with these outcomes.

We considered the term "overall physical function" according to the World Health Organization (WHO)'s International Classification of functioning, Disability and Health (ICF) definition [12] for functioning. We included measures that addressed the following ICF constructs: 1) activities (e.g. functioning evaluated in terms of performance-based measurement focused on limitations in specific activities); 2) measurements assessing participation (involvement in everyday life situations) [12].

Additionally, we defined "simple" performance-based physical function measures or muscle strength if: 1) the test does not require a specialized facility; 2) it takes a relatively short time to be performed (e.g. less than 10 min); 3) it can be performed in any setting (e.g. primary care, acute and/or rehabilitation hospital); and 4) it does not require a professional with specialized expertise. We considered studies that assessed mortality using true observations of all-cause deaths. We excluded self-reported measures that covered quality of life, symptoms, general health perception, emotional and social role functioning and studies that examined the associations of demographic data, symptoms, disease severity, lung function with exacerbation, hospitalization and mortality. We also excluded cross-sectional, pre-post studies, research letters, and congress abstracts.

2.3. Study selection and data extraction

One investigator (WA) reviewed the study titles and abstracts and two investigators (WA and TJF) reviewed the full texts of the articles independently to determine study eligibility. These two investigators met when the full-text screening was complete to resolve any discrepancy.

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [13]. The review protocol was registered in the PROSPERO database (CRD42017059358). Two reviewers (WA and DM) performed the data extraction and tabulation using a standardized form. A second reviewer (VPL) double-checked the extracted data.

2.4. Quality assessment of the included studies

To assess the quality of the studies we used the modified Downs and Black [14] checklist to verify the reporting, external validity, internal validity and statistical power of the studies. The Downs and Black [14] checklist has a total of 27 items but 6 items were considered as 'not applicable' because they are not relevant to non-RCTs. Therefore, the overall possible score was 22 (question 5 gives 2 points for yes). Two assessors (DM and a Research Associate) independently rated each study, and consensus discussion was used to resolve any disagreement. A third assessor (TJF) resolved any outstanding divergences.

3. Results

Fig. 1 shows the PRISMA flow diagram of the search with the number of papers identified, included and excluded (with reasons for exclusion). Seven studies met the inclusion criteria and their characteristics are presented in Table 1. There were considerable variations in terms of characteristics of patients included (e.g. age, stable vs on exacerbation), setting of recruitment, type of tests used, duration of follow-up and outcome measures of interest. The quality assessment of the included studies is presented in Table 2. The total scores ranged from 8 to 18 out of a possible maximum of 22 (average score was 14). No study was classified as "excellent", three studies were classified as "good" [15–17], two classified as "fair" [18,19], and two as "poor" [20,21].

The included studies considered a variety of simple measures of physical function and muscle strength. Four studies included performance-based tests [17,18,21] (handgrip strength (HGS), 1-min sit-to-stand (1-min STS), 5 repetitions sit-to-stand (5STS) and Time Up and Go (TUG) tests) and three included self-reported measures [15,16,19] (Manchester respiratory activities of daily living (MRADL) and Groningen activities for daily living restriction scale (GARS) questionnaires). The handgrip strength (HGS) test was the only muscle strength test that was considered in the included studies. A brief description of these tests and main findings related to each test is shown in Table 3.

3.1. Relationship to all cause or COPD mortality

Six articles examined the relationship of physical function measures or simple measures of muscle strength to mortality [15–19,21]. All articles reported mortality using true observation of all-cause deaths, except one which used probability of survival as the main outcome [16]. Of those, four articles showed some relationship/correlation with mortality.

Puhan et al. [18] demonstrated that the 1-min STS test was strongly associated with mortality at two-year follow up (adjusted hazard ratio (HR) per five more repetitions of 0.58, 95% CI 0.40–0.85; $P = 0.004$) and HGS was moderately associated with mortality (HR 0.84, 95% CI 0.72–1.00; $P = 0.04$). In addition, the authors showed that, 1-min STS test and HGS were substantially lower at baseline in patients who died compared to those who were alive at two-year follow up (11.8 versus 19.5 repetitions; 30.1 kg versus 36.4 kg, respectively). Crook et al. [21]

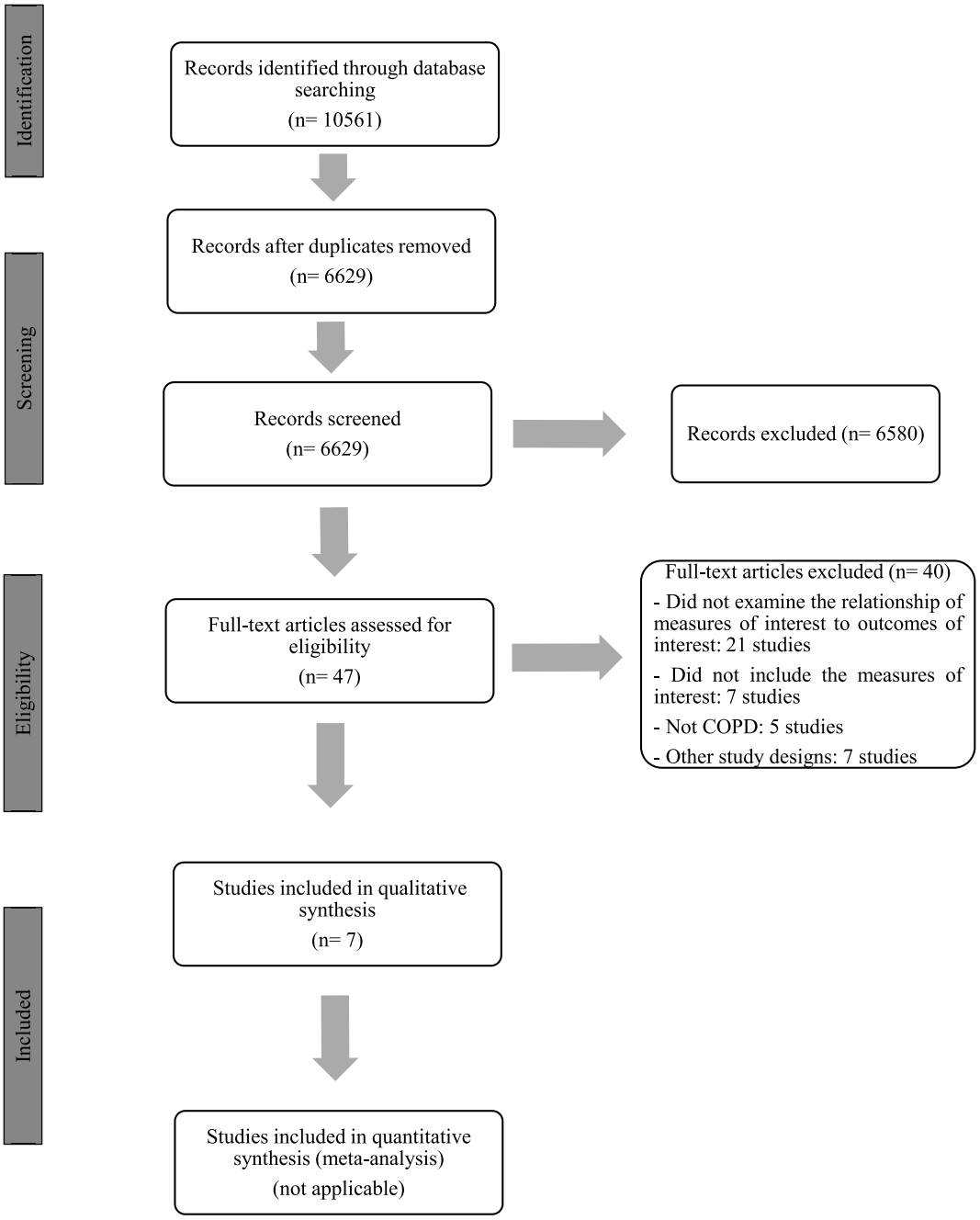

Fig. 1. Flow diagram showing the number of records identified, screened, extracted, and included in the study.

followed up this same cohort for five years. Contrary to the first work, the authors did not observe statistically significantly association between HGS test and mortality. Only the 1-min STS test showed good long-term predictive (5 years) validity for mortality (HR per 3 more repetitions: 0.81, 95% CI 0.65–0.86).

De Buyser et al. [17] compared the predictive value of different physical function measures (HGS, 5STS and TUG tests) in 352 ambulatory older man. Subjects with COPD performed poorer on the all physical function measures than did healthy subjects. ANCOVA analysis further revealed shorter survival time for subjects with COPD ($P = 0.006$). Performance on TUG had no influence on survival time in older men with COPD ($P = 0.202$). The authors posited that the lack of predictive validity of the TUG in subjects with COPD was due to their small sample size.

Yohannes et al. conducted two prospective cohort studies with individuals suffering from COPD; activities of daily living (ADL) were assessed using the MRADL questionnaire [15,19]. The first study found that MRADL was a good predictor of mortality at 2.5 year follow up (Odds ratio (OR) 0.88, 95% CI 0.80–0.97). The second study included patients who had been hospitalized with an acute exacerbation of COPD (AECOPD). At 1-year follow-up, the MRADL was found to be a predictor of mortality in the study population (OR 0.87, 95% CI 0.80–0.94; univariate logistic regression analysis). However, in a multivariate logistic regression analysis this effect was not maintained.

Habraken et al. [16] measured ADL using the GARS questionnaire in 82 COPD patients. The authors stratified the study population into severity subgroups according to the lowest, intermediate and highest tertile of the questionnaire. Higher scores indicated worse daily functioning in all domains. After 2 years follow up, the scores on the GARS declined gradually over time in the majority of patients. Patients in the

Table 1
Characteristics and main findings of the included studies.

Author, Year	Patient and study characteristics	Original objectives of the study	Type of Test	Which outcome of interest was included?			Main findings related to the objective of our systematic review
				Mortality	*Exacerbation*	*Hospitalization*	
Yohannes et al., 2002	COPD outpatients, clinically stable with no change in medication for 1 month and no hospital admission in the previous 6 weeks. N = 137 Age: 60–89 (range) FEV1: 0.89 ± 0.3 (Liters) Follow up: 2.5 years	To evaluate mortality predictors in elderly patients with COPD.	MRADL	Yes	No	No	**MORTALITY**: MRADL score found to be a predictor of mortality (OR = 0.87 (95% CI 0.80–0.97; P < 0.007). Lower MRADL score in those who died versus survivors (7.6 vs 14.0 (p < 0.0001))
Yohannes et al., 2005	Hospitalized AECOPD patients. Groups: Died (n = 36) versus those alive (n = 64). N = 100 Age: 60–89 (range) FEV_1: 0.81 ± 0.3 (Liters) Follow up: 1 year	To examine mortality predictors in patients after AECOPD	MRADL	Yes	No	No	**MORTALITY**: On univariate logistic regression analysis, the MRADL was found to be a predictor of mortality (OR = 0.87 (95% CI 0.80–0.94); P = 0.0001). On multivariate logistic regression analysis MRADL was not a mortality predictor. MRADL score of <12 gave a sensitivity of 86%, specificity of 55%, positive predictive value of 88% and negative predictive value of 52% (in terms of mortality prediction).
Habraken et al., 2011	COPD patients in GOLD stage IV and age ≥60 years. Group Severity based on mean GARS scores at baseline: Low (n = 27), Intermediate (n = 25), High (n = 30). N = 82 Age: 69.4 ± 6.5 FEV_1% predicted: 26 (20–28) Follow up: 5 years	To examine the development of health-related quality of life and functional status over time in COPD patients.	GARS	Yes	No	No	**MORTALITY**: There was a higher probability of survival in the low-severity subgroups. At the 2 year follow up, patients in the high-severity subgroup had a lower probability of survival (hazard ratio 0.53, 95% CI 0.38–0.75)) compared the probability of survival of the low-severity subgroup (0.70 (95% CI 0.55–0.90))
De Buyser et al., 2013	Ambulatory older men recruited from the population register of a semirural community. The subjects self-reported chronic diseases. Comorbidity status was categorized into cardiovascular disease (n = 111), COPD (n = 28) and diabetes mellitus (n = 25). N = 28 Age: 71–86 (all sample) FEV_1% predicted: not reported Follow up: 184 months	To assess and compare the predictive value of physical function measurements for all-cause mortality in older men and to evaluate the Timed Up and Go test (TUG) as a predictor in subjects with underlying comorbidity.	HGS; 5STS test; TUG test	Yes	No	No	**MORTALITY**: Results of the functional tests: 5STS (15.7 ± 4.9 s); TUG test (13.6 ± 5.3 s); Grip strength (22.1 ± 9.6 Kg). Mortality risk doubled for every 8 s increase in time needed to perform TUG (age-adjusted HR per 8 s increase for TUG = 2.08, 95% CI 1.70–2.54, P < 0.001). ANCOVA further revealed shorter survival time for subjects with COPD (P = 0.006). The performance on TUG had NO influence on survival time in subjects with **COPD** (P = 0.202).
Puhan et al., 2013	Multicentre cohort study with COPD patients from primary care in Switzerland and the Netherlands. N = 409 Age: 67.3 ± 10 FEV_1% predicted: not reported GOLD stage II = 63.8% GOLD stage III = 21.8% GOLD stage IV = 14.4% Follow up: 2 years	To assess the predictive performance of the 1-min sit-to-stand test (STS) and handgrip strength tests for mortality and exacerbations.	1-min STS and HGS tests	Yes	Yes	No	**MORTALITY:** 1-min STS test was strongly associated with mortality (adjusted HR per five more repetitions of 0.58, 95% CI 0.40–0.85; p = 0.004). Handgrip strength test was strongly associated with mortality (HR 0.84, 95% CI 0.72–1.00; p = 0.04). 1-min STS test alone was a stronger predictor of 2-year mortality (AUC 0.78), than Handgrip Strength (AUC 0.62). **EXACERBATION:** No association of the 1-min STS test or handgrip strength test with exacerbations was found.
Crook et al., 2017	COPD patients from primary care practices in Switzerland and Netherlands; 409 patients for the handgrip strength test and 371 patients for the 1-min STS test.	To assess the long-term (5-year) predictive performance of the 1-min STS and handgrip strength tests for mortality, health-related quality of life	1-min STS and HGS tests.	Yes	Yes	No	**MORTALITY:** 1-min STS test was strongly associated with long-term mortality (HR per 3 more repetitions: 0.81, 95% CI 0.65–0.86). Handgrip strength test was not

(*continued on next page*)

Table 1 (*continued*)

Author, Year	Patient and study characteristics	Original objectives of the study	Type of Test	Which outcome of interest was included?			Main findings related to the objective of our systematic review
				Mortality	*Exacerbation*	*Hospitalization*	
	N = 409 Age: 67.3 ± 10 FEV_1% predicted: 58 (44–68) GOLD stages II to IV Follow up: 5 years	(HRQoL) and exacerbations in COPD patients.					statistically significantly associated with mortality. **EXACERBATION:** Neither the 1-min STS test nor handgrip strength test was associated with exacerbation.
Martinez et al., 2017	Cross sectional analysis from the COPDGene (Genetic Epidemiology of COPD Study) trial study. For the secondary objective the authors performed longitudinal analysis to verify the long-term associations of HGS with acute respiratory events. N = 272 Age: 64.7 ± 8 FEV_1% predicted: 59 ± 22.5 GOLD 1–2 = 62% GOLD 3–4 = 38%	To test associations of handgrip strength test with pectoralis muscle area (PMA), subcutaneous adipose tissue (SAT), imaging characteristics, and lung function in smokers with COPD, and evaluate the cross-sectional and longitudinal associations of handgrip strength test with acute respiratory events.	HGS test	No	Yes	Yes	**EXACERBATION:** HGS was associated with exacerbation risk. An increment of 5% in the risk of exacerbations for each 1-kg decrement in handgrip strength (risk ratio 1.05; 95% CI 1.01–1.08), and there was a similar risk during follow-up (risk ratio, 1.04; 95% CI 1.01–1.07). Mean follow-up = 2.6 years **HOSPITALIZATION:** An increment of 6% in the risk of hospitalization for each 1- kg decrement in handgrip strength (risk ratio 1.06; 95% CI 1.01–1.12). Risk for severe events (those requiring emergency visits or hospital admission) during follow-up.

OR: odds ratio; HR: hazard ratio; 95% CI: 95% confidence interval; ANCOVA, analysis of covariance; AUC, area under the curve. FEV_1: forced expiratory volume in 1 s; **COPD**: chronic obstructive pulmonary disease; **AECOPD**: acute exacerbation of COPD; **MRADL**: Manchester respiratory activities of daily living questionnaire; **GOLD**: Global Initiative for Chronic Obstructive Lung Disease; **GARS**: Groningen Activities for Daily Living Restriction Scale; **HGS**: handgrip strength; **5STS**: Five-repetitions Sit-to-Stand Test; **TUG**: timed up and go test; **1-min STS**: 1-min sit-to-stand test.

high-severity subgroup had a lower probability of survival (hazard ratio 0.53, 95% CI 0.38–0.75) compared to low severity subgroup (hazard ratio 0.70, 95% CI 0.55–0.90) at the same time point.

3.2. Relationship to COPD exacerbation

Three articles examined the relationship of simple physical function or muscle strength measures to exacerbation. Puhan et al. [18] and Crook et al. [21] assessed in their studies the predictive performance of the 1-min STS and HGS tests for exacerbations at 2-year and 5-year follow up respectively but did not find any relationship to exacerbation [18,21]. Martinez et al. [20] in a cross-sectional analysis of a cohort study, (n = 272) demonstrated that HGS test was associated with exacerbation risk, an increment of 5% in the risk of exacerbations for each 1-kg decrement in HGS test (risk ratio 1.05; 95% CI, 1.01–1.08). Even in their longitudinal analysis and after statistical adjustments, each additional kilogram decrement in HGS test was associated with higher relative risks of having any incident event during follow-up (mean 2.6 years) (risk ratio, 1.04; 95% CI, 1.01–1.07).

3.3. Relationship with COPD-related hospitalizations

One study [20] examined the relationship of our measures of interest to hospitalization (due to acute respiratory events). Martinez et al. [20] found that the risk for severe events (emergency room visits or hospital admissions) during follow-up period (mean 2.6 years) was higher per 1-kg decrement in HGS test (relative risk, 1.06; 95% CI, 1.01–1.12).

4. Discussion

This is the first study that has systematically reviewed information on the predictive validity of simple measures of physical function and muscle strength in relation to outcomes such as mortality, exacerbations and hospitalizations in patients with COPD. Few studies have examined the predictive validity of this type of tests. From the seven studies included in the review, six evaluated the relationship of these measures to mortality; three studies examined this relationship with exacerbations and only one study included the outcome hospitalization. Among the included studies, four used performance-based tests (the most common being the HGS test and 1-min STS) and three used self-reported questionnaires (the MRADL and GARS). The majority of the studies were classified as having “fair” or “poor” methodological quality.

Of the six studies that evaluated the relationship of our tests of interest to mortality, four of them showed that some of these tests are associated with either true observation of death or probability of survival. The tests included in these studies were the 1-min STS, 5STS, TUG, HGS tests and two self-reported questionnaires, the MRADL and GARS. The MRADL and GARS are self-reported questionnaires that evaluate the ability to perform ADL. Limitations in ADL (or disability) have been shown to be a predictor of mortality in different populations [22–24]. The MRADL and GARS questionnaires are easy and quick to be completed and could be used in any setting for prognostic purposes. However, the type of patients to which the questionnaires should be administered should be taken into consideration. Yohannes et al. published two studies examining the MRDAL questionnaire as a mortality predictor in two different populations and found contradictory results. The first study published in 2002 included outpatients with stable COPD and the authors demonstrated that the MRADL was a significant predictor of mortality in that population [19]. The following study (2005) included patients after AECOPD and found that, on the multivariate analysis, MRADL was not a predictor of mortality [15]. One of the main reasons for this divergent result may be that the disease severity as well as the management of these different populations influence their pattern of recovery and functional responses. Moreover, the studies that considered MRADL [15,19] and GARS [16] included older patients (age ≥ 60 years) with GOLD (Global initiative for chronic obstructive lung disease) stage 4, thus the predictive validity of these questionnaires in younger patients who are at earlier stages of the disease is unknown.

The 1-min STS and HGS were shown, in two and three studies, respectively, to be strongly associated with mortality (average follow-up of 2 years) [18]. However, Crook et al. [21] followed a population for 5 years and demonstrated that only the 1-min STS test showed long-term

Table 2
Quality assessment of the included studies.

	Yohannes (2002)	Yohannes (2005)	Habraken (2011)	De Buyser (2013)	Puhan (2013)	Martinez (2017)	Crook (2017)
Reporting							
1. Hypothesis/aim/objective clearly described?	1	1	1	1	1	1	1
2. Main outcomes clearly described in introduction/methods?	1	1	1	1	1	1	0
3. Characteristics of patients included clearly described?	1	1	1	1	1	1	0
4. Interventions of interest clearly described?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
5. Distributions of confounders in each group to be compared clearly described?*	1	1	0	1	1	1	1
6. Main findings clearly described?	1	1	1	1	1	1	1
7. Estimates of the random variability in the data for the main outcomes provided?	0	1	1	1	1	1	1
8. All adverse events potentially from the intervention reported?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
9. Characteristics of patients lost to follow-up described?	0	0	1	1	1	0	0
10. Actual probability values for the main outcomes reported?	1	1	0	1	1	1	1
Subscore (/9)	**6**	**7**	**6**	**8**	**8**	**7**	**5**
External validity							
11. Subjects asked to participate representative of the entire population?	1	1	0	1	1	0	0
12. Subjects who accepted to participate representative of the entire population?	0	1	0	1	0	0	0
13. Staff, places, facilities where patients were treated representative of the treatment most patients receive?	0	0	1	1	0	0	0
Subscore (/3)	**1**	**2**	**1**	**3**	**1**	**0**	**0**
Internal validity - bias							
14. Attempt made to blind study subjects?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
15. Attempt made to blind those measuring the main outcomes?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
16. Any results based on "data dredging" made clear?	1	1	1	1	1	1	0
17. Analyses adjust for different lengths of follow-up or follow-up length the same for cases and controls?	1	0	1	1	0	0	0
18. Statistical tests for main outcomes appropriate?	1	1	1	1	0	1	1
19. Compliance with intervention reliable?	0	0	0	0	0	0	0
20. Main outcome measures used valid and reliable?	1	1	1	1	1	1	1
Subscore (/5)	**4**	**3**	**4**	**4**	**2**	**3**	**2**
Internal validity – confounding (selection bias)							
21. Subjects from both groups recruited from the same population?	1	1	1	0	0	0	0
22. Subjects from both groups recruited over the same time period?	1	1	1	1	1	0	0
23. Study subjects randomized to intervention group?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
24. Randomisation concealed from patients and staff until recruitment complete and irrevocable?	n/a	n/a	n/a	n/a	n/a	n/a	n/a
25. Adequate adjustment for confounding in the analyses for main findings?	1	1	1	1	1	1	1
26. Losses of patient to follow-up taken into account?	0	1	1	1	1	0	0
Subscore (/4)	**3**	**4**	**4**	**3**	**3**	**1**	**1**
Power							
27. Study adequately powered?	0	1	1	0	0	0	0
Subscore (/1)	**0**	**1**	**1**	**0**	**0**	**0**	**0**
TOTAL SCORE (/22)	**14**	**17**	**16**	**18**	**14**	**11**	**8**
TOTAL PERCENTAGE SCORE	**63.6%**	**77.3%**	**72.7%**	**81.8%**	**63.6%**	**50%**	**36.4%**

Scoring: Yes = 1, No = 0, Unable to determinate = 0. * Partially = 1, Yes = 2 points. Not applicable (n/a).
Downs and Black scores: excellent (90–100%), good (69–89%), fair (51–68%), and poor (50% or less).

predictive validity (5 years). The authors attributed the greater long-term predictive validity of the 1-min STS in relation to the HGS to the fact that the 1-min STS is more physically demanding (involves larger muscle groups) and thus more likely to estimate functional exercise performance which may reflect the patients' overall health status.

The performance of the HGS in predicting exacerbations seems to be conflicting. Martinez et al. [20] in a cross-sectional analysis of a cohort study (n= 272) demonstrated that HGS test was associated with exacerbation risk but Puhan et al. [18] and Crook et al. [21] showed that the HGS test did not predict the incidence of exacerbations (period of follow up: 2 and 5 years, respectively). The authors [18,21] attributed the lack of association in their studies to the small number of exacerbations that occurred during the study periods, probably because of the COPD severity (the majority of the patients in their cohort were GOLD stage 2). Moreover, the results of Martinez's study might have been overestimated since they recruited patients from a specialized outpatient clinic and relied on patients to report whether or not they had an exacerbation (while in Puhan's cohort, occurrence of exacerbation had to be medically confirmed and they recruited patients from primary care centres). In addition, Puhan's model of events during follow-up were adjusted for exacerbation history at enrollment.

Despite the conflicting results of the predictive validity of the HGS test in relation to exacerbation, there is strong data on its predictive validity in term of mortality and hospitalization and could be a great test to use for prognostication in many clinical settings. The HGS test requires an inexpensive hand-held dynamometer (approximately US$ 300; one time investment), is simple to perform by individuals of any age and a good option for patients who may not be able to conduct tests where walking or standing is required [21]. In addition, there are normative values [25–27] available to guide clinicians' decision on the best treatment options based on the test results. The 1-min STS test [28] is also a good option since it has been shown to be a strong predictor of mortality in patients with COPD [18], has excellent long-term predictive validity [21], is responsive to pulmonary rehabilitation [29], does not require any specific equipment and has normative data available based on age and gender [28]. The disadvantage of the 1-min STS is that it may

Table 3
Description of the tests and main findings related to each test.

Performance-based tests	Description	Main findings
1-min Sit-To-Stand test (1-min STS)	The test is performed with a standard height (46 cm) chair without arm rests. The patient starts the test sitting on a chair and repeatedly stands up and sits down again during 1 min. The outcome is how many times patients perform this movement in 1 min. This test measures general lower limb strength and is a strong predictor of mortality in older adults.	Associated with 2-year mortality. Adjusted HR = 0.58 (95% CI 0.4–0.85; P = 0.004) per 5 more STS repetitions. Lower performance on the test in patients who had died at 2 year follow up than in patients who were still alive (11.8 rises vs 19.5) (Follow up: 2 years) [19]. 1-min STS test showed good long-term predictive (5 years) validity for mortality (HR per 3 more repetitions: 0.81, 95% CI 0.65–0.86) [21]. Number of repetitions during the STS test was 18.9 ± 8.8. STS test results were lower for females compared with males (17.2 ± 7.4 vs 20.3 ± 9.5) (Follow up: 2 years) [19]. 1-min STS test was NOT associated with exacerbation (Follow up: 2 years) [19].
Five-repetitions Sit-To-Stand test (5STS)	The patient starts the test sitting and repeatedly stands up and sits down five times as quickly as he/she can. Outcome is time (how long it takes for the patients to perform the five repetitions).	5STS was associated with all-cause mortality in ambulatory multimorbidity older men sample (n = 352). (HR 1.46, 95% CI 1.28–1.65, p < 0.001). Authors did not report COPD subgroup analysis (n = 28; not available data) (Follow up: 184 months) [18].
Timed Up and Go test (TUG)	The patient starts the test sitting on chair and is asked to stand up, walk 4 m and come back and sit down again. The outcome is time (how long it takes for the patient to perform the task). This test measures mobility and balance.	ANCOVA analysis further revealed shorter survival time for subjects with COPD (P = 0.006). Performance on TUG had no influence on survival time in older men with COPD (P = 0.202) (Follow up: 184 months) [18].
Handgrip strength test (HGS)	The patient presses a dynamometer with the hand. It is a measure of overall body strength and a strong predictor of mortality in older adults and other clinical populations.	HGS test was strongly associated with mortality (adjusted HR 0.84, 95% CI 0.72–1.00; P = 0.04) (Follow up: 2 years) [19]. HGS test was not associated with mortality at 5-year follow up [21]. HGS was associated with exacerbations risk during follow-up period (Risk ratio 1.04; 95% CI 1.01–1.07) (Mean follow up: 2.6 years) [20]. HGS test was not associated with exacerbations (Incidence ratio 1.00 95% IC 0.98–1.02) (follow up: 5 years) [21]. HGS was associated with hospitalization risk (Risk ratio 1.06; 95% CI 1.01–1.12) (risk for severe events) (Mean follow up: 2.6 years) [20].
Self-reported measures		
Manchester Respiratory Activities of Daily Living questionnaire (MRADL)	MRADL is a 21-item, self-reported questionnaire. It has four domains: Mobility (7 items), Kitchen activities (4 items), Domestic tasks (6 items), and Leisure activities (4 items). The MRADL is scored compositely in the range 0–21, with a score of 21 signifying no physical impairment.	MRADL was a good predictor of mortality at 2.5 year follow up (OR 0.87; 95% CI 0.80–0.97; P < 0.007) (univariate logistic regression analysis). But on multivariate logistic regression analysis, this finding was not confirmed [15]. At 1-year follow-up, the MRADL was found to be a predictor of mortality (OR 0.87 95% CI 0.80–0.94). However, in a multivariate logistic regression analysis this effect was not maintained. (AECOPD) [16].
Groningen Activities of daily living Restriction Scale (GARS)	The GARS measures ADL (personal care) as well as instrumental ADL (domestic activities). The total score has a range of 18–54. Higher scores indicate worse daily functioning in all domains.	At 2-year follow-up, high-severity subgroup based on the GARS had a lower probability of survival (hazard ratio 0.53, 95% CI 0.38–0.75) [17].

HR, hazard ratio; 95% CI, 95% confidence interval; RR, risk ratio; STS, sit to stand; ADL, activities of daily living; 5STS, 5-min sit to stand.

be difficult for patients with musculoskeletal or neurological conditions to perform. It also seems to have a floor effect (older or too debilitated patients may not be able to complete the test) [28] and there is limited evidence for its association with exacerbation and hospitalization in individuals with COPD. It is important to note that even though the HGS and 1-min STS tests seem to be promising tests to be used for prognosis purposes in patients with COPD, the data available are based on 3 and 2 studies, respectively.

The strength of this review is that it compiles, in a systematic manner, information on the predictive validity of simple measures of physical function and muscle strength in relation to exacerbation, hospitalization and mortality in individuals with COPD. Previous reviews had focused on the clinical value and psychometric properties of these tests [8,9,11] but to the best of our knowledge, no review had compiled information on the prognostic value of these tests.

Our systematic review has some limitations. First, the main challenge was that the primary objective of the included studies was not necessarily related to the primary objective of our systematic review which made data extraction difficult. Other limitations were the high diversity of methodologies (patient characteristics, exposures, type of tests used and outcome measures) of the included studies which made it difficult for the authors to conduct a meta-analysis. Finally, the majority of the studies in this review (n = 4) were classified as having a "fair" or "poor" quality assessment according the Downs & Black checklist. The items that received low scores were related to external/internal validity and study's power demonstrating the low methodological quality of the included studies.

5. Conclusions

There is a limited number of studies examining the prognostic value of simple measures of physical function and muscle strength in relation

to exacerbations, hospitalizations and mortality in individuals with COPD. To date, the HGS and 1-min STS tests are the most studied tests and seem to be suitable for prognosis purposes in individuals with COPD. However, more studies with better methodological quality are needed to confirm these findings. It remains unclear whether improvements in these physical function measures can translate into improvement in patient prognosis. Further studies addressing the short-term predictive validity of these tests (e.g. 30-day readmission to hospital) are also needed.

Financial support

None.

Author's contributions

WA and TJF conducted titles and abstracts review. DM and WA screened and accessed the data. All authors made substantial contributions to the interpretation of the data. DM and TJF wrote the paper. WA, VPL and JB critically revised the manuscript for important intellectual content. All authors approved the final version to be submitted for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the research associate Dr Catherine Tansey for helping with the quality assessment of the studies.

List of abbreviations

COPD	Chronic Obstructive Disease
6MWT	Six-minute walk test
ICF	International Classification of functioning, Disability and Health
HGS	Handgrip strength test
STS	sit-to-stand
5STS	Five repetitions sit-to-stand
TUG	Time Up and Go
1-min STS	1-min sit-to-stand
GARS	Groningen activities for daily living restriction scale
MRADL	Manchester respiratory activities of daily living questionnaire
ADL	Activities of daily living
AECOPD	Acute exacerbation of chronic obstructive pulmonary
GOLD	Global Initiative for chronic obstructive lung disease

Appendix A. Electronic search strategy

Terms Used	Variations of Terms to use when searching
Physical Function Measures	Performance Based Tests • Exercise Test/OR ((exercise* or physical activit* or performance or step* or stand* or walk* or strength* or function* or balance or grip or handgrip or sit or lift* or gait or stair or rise) adj3 (test* or measure* or score* or status or assess* or evaluat* or batter* or scale or subscale)).tw,kf. Self-Reported Measures • exp Activities of Daily Living/or ((activit* adj2 (daily living or limit*)) or ADL or ADLs).tw,kf. *AND* • exp "Surveys and Questionnaires"/or (assess* or tool* or evaluat* or test* or scale* or subscale* or score* or inventor* or questionnaire* or measure*).tw,kf. • (self-report* or PFP or PPT or "PFP-10" or SPPB or "SF-36").tw,kf.
Hospitalization/Mortality/ Exacerbation	• exp Mortality/or Prognosis/or exp disease progression/ • exp Hospitalization/ • (hospitali* or ((length or hospital) adj2 stay*) or admission* or admit* or readmission* or readmit* or discharge*).tw,kf. • (mortalit* or fatalit* or death* or progress* or exacerbat* or prognos* or deterioration).tw,kf.
COPD	• exp Pulmonary Disease, Chronic Obstructive/ • Lung Diseases, Obstructive/ • exp Pulmonary Emphysema/ • (obstructive adj2 (pulmonary or lung$ or respirat$ or air$)).tw,kf. • (chronic air$ adj2 (obstruction$ or limitation$ or occlusion$)).tw,kf. • (chronic bronch$ adj2 (obstruction$ or limitation$ or occlusion$)).tw,kf. • (chronic$ adj2 bronch$).tw,kf. • COPD.tw,kf. • COAD.tw,kf. • emphysema$.tw,kf. • (acos and asthm*).tw,kf.

References

[1] J. Vestbo, S.S. Hurd, A.G. Agustí, P.W. Jones, C. Vogelmeier, A. Anzueto, et al., Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary, Am. J. Respir. Crit. Care Med. 187 (4) (2013) 347–365, https://doi.org/10.1164/rccm.201204-0596PP.

[2] M.D. Eisner, P.D. Blanc, E.H. Yelin, S. Sidney, P.P. Katz, L. Ackerson, et al., COPD as a systemic disease: impact on physical functional limitations, Am. J. Med. 121 (9) (2008) 789–796, https://doi.org/10.1016/j.amjmed.2008.04.030.
[3] M.D. Eisner, C. Iribarren, P.D. Blanc, E.H. Yelin, L. Ackerson, N. Byl, et al., Development of disability in chronic obstructive pulmonary disease: beyond lung function, Thorax 66 (2) (2011) 108–114, https://doi.org/10.1136/thx.2010.137661.
[4] E.B. Swallow, D. Reyes, N.S. Hopkinson, W.D. Man, R. Porcher, E.J. Cetti, et al., Quadriceps strength predicts mortality in patients with moderate to severe chronic obstructive pulmonary disease, Thorax 62 (2) (2007) 115–120, https://doi.org/10.1136/thx.2006.062026.
[5] B.R. Celli, Predictors of mortality in COPD, Respir. Med. 104 (6) (2010) 773–779, https://doi.org/10.1016/j.rmed.2009.12.017.
[6] F. Maltais, M. Decramer, R. Casaburi, E. Barreiro, Y. Burelle, R. Debigaré, et al., An official American Thoracic Society/European Respiratory Society statement: update on limb muscle dysfunction in chronic obstructive pulmonary disease, Am. J. Respir. Crit. Care Med. 189 (9) (2014) e15–62, https://doi.org/10.1164/rccm.201402-0373ST.
[7] A.E. Holland, M.A. Spruit, T. Troosters, M.A. Puhan, V. Pepin, D. Saey, et al., An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease, Eur. Respir. J. 44 (6) (2014) 1428–1446, https://doi.org/10.1183/09031936.00150314.
[8] K.L. Bui, A. Nyberg, F. Maltais, D. Saey, Functional tests in chronic obstructive pulmonary disease, Part 1: clinical relevance and links to the international classification of functioning, disability, and health, Ann Am Thorac Soc 14 (5) (2017) 778–784, https://doi.org/10.1513/AnnalsATS.201609-733AS.
[9] K.L. Bui, A. Nyberg, F. Maltais, D. Saey, Functional tests in chronic obstructive pulmonary disease, Part 2: measurement properties, Ann Am Thorac Soc 14 (5) (2017) 785–794, https://doi.org/10.1513/AnnalsATS.201609-734AS.
[10] K.N. Johnston, A.J. Potter, A. Phillips, Measurement properties of short lower extremity functional exercise tests in people with chronic obstructive pulmonary disease: systematic review, Phys. Ther. 97 (9) (2017) 926–943, https://doi.org/10.1093/ptj/pzx063.
[11] G.W. Bisca, A.A. Morita, N.A. Hernandes, V.S. Probst, F. Pitta, Simple lower limb functional tests in patients with chronic obstructive pulmonary disease: a systematic review, Arch. Phys. Med. Rehabil. 96 (12) (2015) 2221–2230, https://doi.org/10.1016/j.apmr.2015.07.017.
[12] W.H. Organization, Towards a Common Language for Functioning, Disability, and Health: ICF, the International Classification of Functioning, Disability and Health, World Health Organization, 2002. https://www.who.int/classifications/icf/icfbeginnersguide.pdf.
[13] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, P. Group, Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement, PLoS Med. 6 (7) (2009), e1000097, https://doi.org/10.1371/journal.pmed.1000097.
[14] S.H. Downs, N. Black, The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions, J. Epidemiol. Community Health 52 (6) (1998) 377–384, https://doi.org/10.1136/jech.52.6.377.
[15] A.M. Yohannes, R.C. Baldwin, M.J. Connolly, Connolly, Predictors of 1-year mortality in patients discharged from hospital following acute exacerbation of chronic obstructive pulmonary disease, Age Ageing 34 (5) (2005) 491–496, https://doi.org/10.1093/ageing/afi163.
[16] J.M. Habraken, W.M. van der Wal, G. Ter Riet, E.J. Weersink, F. Toben, P. J. Bindels, Health-related quality of life and functional status in end-stage COPD: a longitudinal study, Eur. Respir. J. 37 (2) (2011) 280–288, https://doi.org/10.1183/09031936.00149309.
[17] S.L. De Buyser, M. Petrovic, Y.E. Taes, K.R. Toye, J.M. Kaufman, S. Goemaere, Physical function measurements predict mortality in ambulatory older men, Eur. J. Clin. Investig. 43 (4) (2013) 379–386, https://doi.org/10.1111/eci.12056.
[18] M.A. Puhan, L. Siebeling, M. Zoller, P. Muggensturm, G. ter Riet, Simple functional performance tests and mortality in COPD, Eur. Respir. J. 42 (4) (2013) 956–963, https://doi.org/10.1183/09031936.00131612.
[19] A.M. Yohannes, R.C. Baldwin, M. Connolly, Mortality predictors in disabling chronic obstructive pulmonary disease in old age, Age Ageing 31 (2) (2002) 137–140, https://doi.org/10.1093/ageing/31.2.137.
[20] C.H. Martinez, A.A. Diaz, C.A. Meldrum, M.N. McDonald, S. Murray, G.L. Kinney, et al., COPDGene investigators. Handgrip strength in chronic obstructive pulmonary disease. Associations with acute exacerbations and body composition, Ann Am Thorac Soc 14 (11) (2017) 1638–1645, https://doi.org/10.1513/AnnalsATS.201610-821OC.
[21] S. Crook, A. Frei, G. Ter Riet, M.A. Puhan, Prediction of long-term clinical outcomes using simple functional exercise performance tests in patients with COPD: a 5-year prospective cohort study, Respir. Res. 18 (1) (2017) 112, https://doi.org/10.1186/s12931-017-0598-6.
[22] S.M. Dunlay, S.M. Manemann, A.M. Chamberlain, A.L. Cheville, R. Jiang, S. A. Weston, et al., Activities of daily living and outcomes in heart failure, Circ Heart Fail 8 (2) (2015) 261–267, https://doi.org/10.1161/CIRCHEARTFAILURE.114.001542.
[23] L.R. Ramos, E.J. Simoes, M.S. Albert, Dependence in activities of daily living and cognitive impairment strongly predicted mortality in older urban residents in Brazil: a 2-year follow-up, J. Am. Geriatr. Soc. 49 (9) (2001) 1168–1175, https://doi.org/10.1046/j.1532-5415.2001.49233.x.
[24] N.E. Brummel, S.P. Bell, T.D. Girard, P.P. Pandharipande, J.C. Jackson, A. Morandi, et al., Frailty and subsequent disability and mortality among patients with critical illness, Am. J. Respir. Crit. Care Med. 196 (1) (2017) 64–72, https://doi.org/10.1164/rccm.201605-0939OC.
[25] M.A. Spruit, M.J. Sillen, M.T. Groenen, E.F. Wouters, F.M. Franssen, New normative values for handgrip strength: results from the UK Biobank, J. Am. Med. Dir. Assoc. 14 (10) (2013) 775, https://doi.org/10.1016/j.jamda.2013.06.013, e5-11.
[26] S.L. Wong, Grip strength reference values for Canadians aged 6 to 79: Canadian Health Measures Survey, 2007 to 2013, Health Rep. 27 (10) (2016) 3–10. PubMed PMID: 27759870, https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2016010/article/14665-eng.pdf?st=s2lLwXn-.
[27] Y.C. Wang, R.W. Bohannon, X. Li, B. Sindhu, J. Kapellusch, Hand-grip strength: normative reference values and equations for individuals 18 to 85 Years of age residing in the United States, J. Orthop. Sport. Phys. Ther. 48 (9) (2018) 685–693, https://doi.org/10.2519/jospt.2018.7851.
[28] A. Strassmann, C. Steurer-Stey, K.D. Lana, M. Zoller, A.J. Turk, P. Suter, et al., Population-based reference values for the 1-min sit-to-stand test, Int. J. Public Health 58 (6) (2013) 949–953, https://doi.org/10.1007/s00038-013-0504-z.
[29] T. Vaidya, C. de Bisschop, M. Beaumont, H. Ouksel, V. Jean, F. Dessables, et al., Is the 1-minute sit-to-stand test a good tool for the evaluation of the impact of pulmonary rehabilitation? Determination of the minimal important difference in COPD, Int. J. Chronic Obstr. Pulm. Dis. 11 (2016) 2609–2616, https://doi.org/10.2147/COPD.S115439.