Comparison of the Dyspnoea-12 and Multidimensional Dyspnoea Profile in people with COPD

Marie T. Williams[1], David John[2] and Peter Frith[2,3]

Affiliations: [1]Health and Alliance for Research in Exercise, Nutrition and Activity (ARENA), Sansom Institute for Health Research, School of Health Sciences, University of South Australia, Adelaide, Australia. [2]School of Health Sciences, University of South Australia, Adelaide, Australia. [3]Southern Adelaide Health Service and Repatriation General Hospital, Adelaide, Australia.

Correspondence: Marie T. Williams, Health and Alliance for Research in Exercise, Nutrition and Activity (ARENA), Sansom Institute for Health Research, School of Health Sciences, University of South Australia, North Terrace, Adelaide, South Australia 5000, Australia. E-mail: marie.williams@unisa.edu.au

Cite this article as: Williams MT, John D, Frith P. Comparison of the Dyspnoea-12 and Multidimensional Dyspnoea Profile in people with COPD. *Eur Respir J* 2017; 49: 1600773 [https://doi.org/10.1183/13993003.00773-2016].

ABSTRACT We directly compared convergent, discriminant and concurrent validity of the Dyspnoea-12 (D-12) and Multidimensional Dyspnoea Profile (MDP) in people with chronic obstructive pulmonary disease.

Breathlessness measures (D-12, MDP, visual analogue scales and descriptors) were completed for two focal periods (daily life and end of walk test). Instrument structure (D-12 and MDP item grouping) was assessed with factor analysis. Differences between airflow severity stage and focal periods (ANOVA, t-test and Chi-squared test), associations between D-12 and MDP (r, r^2 for static pulmonary function, 6-min walk test and self-reported measures of impairment) and individual consistency for comparable items of the D-12 and MDP (McNemar's test) were assessed.

In 84 participants (mean±SD age 70±9 years, 47 males, forced expiratory volume in 1 s 48±17% predicted), item groupings were confirmed for both focal periods. Developer-recommended single and subdomain scores were highly correlated, and demonstrated similar convergent, discriminant and concurrent validity. Individual consistency differed between the D-12 and MDP according to item/item groups.

At the level of developer-recommended single and subdomain scores, the D-12 and MDP share similar psychometric properties, but these instruments serve different purposes, do not assess the same sensations or emotions and are not interchangeable.

This article has supplementary material available from erj.ersjournals.com

Received: Sept 07 2015 | Accepted after revision: Oct 26 2016

Clinical trials: This study is registered at www.anzctr.org.au with identifier number ACTRN12611000292976.

Support statement: This study was supported by a National Health and Medical Research Council (NHMRC) Project Grant (1010309). The NHMRC did not have input or contribute to the development of the research or the manuscript. Funding information for this article has been deposited with the Open Funder Registry.

Conflict of interest: Disclosures can be found alongside this article at erj.ersjournals.com

Introduction

Dyspnoea is a perceptual experience which differs between daily life (clinical dyspnoea) and the sensation induced by laboratory-based procedures (laboratory dyspnoea) [1, 2]. Various uni- and multidimensional instruments are available to assess dyspnoea, with a greater number of instruments available for assessing the impact of dyspnoea (frequency, functional performance or activity and health-related quality of life) compared with assessments of the sensory-perceptual experience (intensity and sensory quality) or affective distress (unpleasantness and emotional response) [3]. Assessments of the sensory-perceptual and affective responses of dyspnoea include descriptors of breathlessness (volunteered [3, 4] or selected from predetermined lists [3–7]) and visual analogue scales (VASs) [3], although there is no standard VAS instruction or minimum/maximum anchor definition. Two new instruments are the Dyspnoea-12 (D-12) [3, 8] and Multidimensional Dyspnoea Profile (MDP) [3, 9–12]. These instruments share conceptual similarities, but differ in terms of intent, development, items, scale of measurement and scoring (table 1). While psychometric properties (*e.g.* reliability, convergent and discriminant validity) for each of these instruments have been confirmed during development [8, 9, 11, 12], to date no direct comparison between the D-12 and MDP in the same population has been reported.

During 2011, we commenced a clinical trial investigating the addition of a cognitive behavioural therapy programme for the sensation of breathlessness to pulmonary rehabilitation. This provided an opportunity to directly compare the D-12 and MDP in a clinical setting alongside other measures of the sensory-perceptual and affective distress of dyspnoea. The research questions posed for this study were: in a sample of people with stable chronic obstructive pulmonary disease (COPD), do the D-12 and MDP demonstrate similarities in 1) convergent validity with other assessments of the sensory-perceptual and affective distress of dyspnoea, 2) discriminant validity between breathlessness contexts (daily life and exercise), 3) associations with measures of static pulmonary function, functional exercise capacity, anxiety and depression, respiratory-related impairment or quality of life (concurrent validity), and 4) participant responses for comparable items?

Methods

This descriptive study reports cross-sectional, baseline data from a randomised controlled trial (National Health and Medical Research Council (1010309); Australian and New Zealand Clinical Trials Registry, trial number ACTRN12611000292976). Ethical approval was granted by Human Research Ethics Committees of the University of South Australia (P153/07) and Repatriation General Hospital (P56/07) (Adelaide, Australia). All subjects provided written informed consent.

People referred to Comprehensive Pulmonary Rehabilitation (Repatriation General Hospital) were eligible for inclusion if they had a clinical diagnosis of COPD and had at least moderate airways obstruction (post-bronchodilator forced expiratory volume in 1 s (FEV_1) <80% predicted and best recorded ratio of FEV_1/forced vital capacity (FVC) of <70% (*i.e.* Global Initiative for Obstructive Lung Disease (GOLD) [13] grade ⩾II)). Participants were excluded if they had cognitive or memory impairments (Mini-Mental State Examination score <23/30 [14]), clinically unstable COPD (hospitalisation, exacerbation or modification of medication within the past 6 weeks), comorbidities which were likely to render exercise unsafe or were registered for pulmonary surgical interventions.

Post-bronchodilator pulmonary function testing (spirometry and plethysmography [15–17]) was used to confirm the COPD diagnosis, assess the severity of airflow limitation according to GOLD grades (FEV_1 % pred: 50–80% (grade II, moderate), 30–49% (grade III, severe) and <30% (grade IV, very severe)) [13] and hyperinflation. Functional exercise capacity was assessed by the 6-min walk test (6MWT including pre–post rating of perceived exertion) according to the American Thoracic Society recommendations [18, 19]. Self-report of generalised anxiety and depression (Hospital Anxiety and Depression Scale (HADS) [20]), respiratory-related functional performance (modified Medical Research Council (mMRC) scale [21]) and quality of life (Chronic Respiratory Questionnaire (CRQ) [22]) were assessed.

Sensory-perceptual and affective distress assessments of global intensity (VAS-I) and unpleasantness (VAS-U) were done using two 10-cm VASs [3]. An open-ended question was used to solicit volunteered descriptors of sensory quality and affective responses by subjects in their own words [4]. Using the list devised by Mahler *et al.* [7], participants selected (endorsed) up to three descriptors from a list of 15 statements that best reflected their sensation.

The D-12 is a 12-item questionnaire which provides an assessment of dyspnoea severity [8]. Yorke *et al.* [8] recommend the use of the focal period "these days" reflecting breathlessness experienced in daily life. Two subdomains have been identified (Physical: items 1–7, score range 0–21 and Affective: items 8–12, score range 0–15; table 1) with item scores (mild (score 1), moderate (score 2) or severe (score 3) or does not apply (score 0)) summed to provide a total score (score range 0–36) [8], although separate subdomain scores may also be calculated [23]. The MDP contains 11 items which respondents rate on a continuous

TABLE 1 Comparisons of assessments for the sensory-perceptual and affective domains of breathlessness

Items	VAS [3]	Descriptor list [7]	D-12 [8]		MDP [9]	
Intensity	Y					
Unpleasantness	Y				Y (A_1)	MDP-IP
Sensory quality						MDP-IP
I feel that my breathing is rapid		Y			I am breathing a lot (rapidly, deeply or heavily)#	MDP-IP
My breathing is heavy		Y				MDP-IP
My breath does not go out all the way		Y				MDP-IP
My chest feels tight		Y			My chest or lungs feel tight or constricted#	MDP-IP
My chest is constricted		Y				MDP-IP
I feel that I am breathing more		Y				MDP-IP
My breathing requires effort		Y			My breathing requires muscle work or effort#	MDP-IP
My breathing requires work		Y	Y	D-12 Physical		MDP-IP
I feel out of breath		Y		D-12 Physical		MDP-IP
I cannot get enough air		Y	Y	D-12 Physical	I am not getting enough air, I feel hunger for air or I am smothering#	MDP-IP
I feel hungry for air		Y		D-12 Physical		MDP-IP
I feel that I am smothering		Y		D-12 Physical		MDP-IP
I feel that I am suffocating		Y		D-12 Physical		MDP-IP
I have difficulty catching my breath			Y	D-12 Physical		MDP-IP
My breath does not go in all the way		Y	Y	D-12 Physical		MDP-IP
I feel short of breath			Y	D-12 Physical		MDP-IP
My breathing is shallow		Y		D-12 Physical		MDP-IP
My breathing requires mental effort or concentration				D-12 Physical	Y	MDP-IP
My breathing is exhausting			Y	D-12 Physical		MDP-IP
My breathing is uncomfortable			Y	D-12 Physical		MDP-IP
Affective or emotional response						MDP-ER
My breathing makes me feel miserable			y	D-12 Affective		MDP-ER
My breathing is distressing			Y	D-12 Affective		MDP-ER
My breathing makes me feel agitated			Y	D-12 Affective		MDP-ER
My breathing is irritating			Y	D-12 Affective		MDP-ER
My breathing makes me feel depressed			Y	D-12 Affective	Y	MDP-ER
Anxious				D-12 Affective	Y	MDP-ER
Frustrated				D-12 Affective	Y	MDP-ER
Angry				D-12 Affective	Y	MDP-ER
Afraid				D-12 Affective	Y	MDP-ER
Instrument characteristics						
Items	1	15	12		11	
Scoring scale	mm	NA	Categories scored on 0–3 scale		Continuous numeric rating scale (0–10)	
Recommended scoring	Single score	NA	Single score		Individual items±subdomain scores; if single score required=A_1	
Domains assessed [3]						
Sensory quality		Y	Y		Y	
Intensity	Y (VAS-I)		Y		Y	
Affective distress	Y (VAS-U)		Y		Y	
Symptom impact/burden			Y			
Recommended focal period for assessment	Specified event or time		"These days"		Specified event or time	
Development context			Routine clinic visit		Concurrent laboratory and emergency departments	

Volunteered descriptors are not included as subjects may volunteer multiple descriptors reflecting various intensity, unpleasantness, sensory quality or affective responses [4]. VAS-I/U: visual analogue scale (Intensity/Unpleasantness); D-12: Dyspnoea-12; MDP: Multidimensional Dyspnoea Profile; Y: yes (item is included); A_1: MDP item for unpleasantness; NA: not applicable. #: sensory quality "forced-choice" items.

scale (0–10) and a forced-choice question for the five sensory qualities ("most accurately describes") [9]. Items group into two domains: Immediate Perception (IP: unpleasantness and sensory quality, score range 0–60) and Emotional Response (ER: score range 0–50) (table 1) [9]. By design, each item can be analysed

separately or domain scores calculated (IP and ER) [9]. Where a single score is preferred, the item for "unpleasantness" (A_1) is recommended [9]. The focal period is determined by users as appropriate for the intent of the research or clinical situation (*e.g.* "right now" or "at the end of a minute of a particular activity") [9]. A penultimate version of MDP [24] was used during the initial stages of the clinical trial, updated to the final pre-release version on advice of the developers [9].

Protocol

Participants completed all breathlessness assessments on two occasions. Assessments were collated into a booklet commencing with VASs and volunteered descriptors of breathlessness, with the remaining assessments randomised using a predetermined protocol. A study staff member read each question to the participant using a script with standardised prompts (see online supplementary material), transcribed responses (volunteered descriptors) and invited participants to mark (VASs) or circle a response within items that best reflected their sensation. Following post-bronchodilator pulmonary function testing and completion of the mMRC scale, CRQ and HADS, but prior to the initial 6MWT while resting, participants were asked to recall and assess the sensation of breathlessness experienced as part of daily life (focal period: "on average over the past 2 weeks"). On completion, participants were advised that immediately after completion of the first of two 6MWTs, they would work through the breathlessness assessments again but were to focus on breathlessness experienced "during the last minute of the walk test".

Data management and analysis

Characteristics of participants meeting eligibility for the trial, declining participation and those with complete data for breathlessness assessments for both focal periods were compared. Volunteered and endorsed descriptors of breathlessness were allocated to predetermined descriptor categories using processes previously described [4] with the number of participants, rather than the frequency of descriptors, used for analysis. Single scores as recommended by the instrument developers (D-12 Total score [8] *versus* MDP-A_1 (unpleasantness) [9]) and subdomain scores (D-12 Physical *versus* MDP-IP; D-12 Affective *versus* MDP-ER) were calculated. Systematic bias between instrument scores was assessed with Bland–Altman plots.

As we modified the focal period for breathlessness assessments recommended for the D-12 and MDP in order to align with other self-reported assessments recommended for use within pulmonary rehabilitation, confirmatory factor analysis (model fit and factor loadings) and exploratory principal components analysis (varimax rotation and Cronbach's α) were undertaken to determine whether the two-factor structure and internal consistency of items reported by the original developers was retained [8, 11, 12].

We hypothesised that the D-12 and MDP would concur with other assessments of the sensory-perceptual and affective distress of dyspnoea in that they would 1) not discriminate between severity of airways obstruction (FEV_1 % pred cut-offs recommended for GOLD grades) and 2) discriminate between context of breathlessness (difference between last minute of the 6MWT *versus* daily life). These relationships were assessed using ANOVA and paired t-tests (continuous variables) with the Chi-squared test and McNemar's 2×2 test for categorical variables.

In order to assess concurrent validity, linear associations (Pearson's correlation coefficient (r with 95% confidence limits; for sample size n=84, r significant at >0.2) and coefficients of determination (r^2)) were calculated between scores for scalable instruments (VASs, D-12 and MDP) and measures of static pulmonary function, 6MWT distance, HADS, mMRC scale and CRQ. *Post hoc* z-tests on Fisher's transformed correlations were used to test for differences between the D-12 and MDP (single and subdomain scores) [25]. Given the multiplicity of testing and interdependence of data (*i.e.* all scores derived from the same sample of participants), in all analyses, Bonferroni adjusted p-values were required for statistical significance. All statistical analysis was performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

Consistency of individual participant response was assessed for four comparable items/item groups of the MDP and D-12 (unpleasantness/uncomfortable, depressed, air hunger, work/effort [9]). Consistency was calculated as the number of individual participants that selected or rated ($\geqslant$1) the comparable item/item group across instruments (descriptor list [7], D-12 and MDP) expressed as a percentage (with the denominator equal to the total sample minus the number of participants that did not select/rate the item in any instrument). McNemar's test was used to assess inconsistency of participant responses (rated $\geqslant$1 in one but not the other) between the D-12 and MDP items/item groups (the online supplementary material presents details of items and item grouping).

Results

Of the 277 persons screened to participate in the trial, 85 declined and 91 did not meet the inclusion criteria; thus 101 participants met all inclusion criteria, with 84 participants providing complete data for breathlessness assessments in both focal periods (table 2 and online supplementary table S4). Of the 85

TABLE 2 Summary of participant characteristics for the sample with data for both breathlessness focal periods

Subjects	84
Age years	69.7±8.9
Height cm	166±9
Weight kg	75.1±18.6
Male	47 (56)
English spoken at home	81 (96)
Mini-Mental State Examination score	29.2±1.7
mMRC scale score	1.9±1.2
0	5 (6)
1	37 (44)
2	13 (15)
3	17 (20)
4	12 (14)
HADS score	
Anxiety	7.0±4.4
Depression	6.3±4.2
CRQ score	
Dyspnoea	4.7±1.4
Fatigue	3.9±1.3
Emotion	4.6±1.2
Mastery	4.8±1.4
Perceived rate of exertion score	
Pre-6MWT	0.9±1.1
Post-6MWT	3.5±1.8
Maximum distance 6MWT m	385±135
FEV_1 % pred	47.9±16.6
FEV_1/FVC	42.6±14.1
RV/TLC % pred	134±26

Data are presented as n, mean±SD or n (%). mMRC: modified Medical Research Council; HAD: Hospital Anxiety and Depression Scale; CRQ: Chronic Respiratory Questionnaire; 6MWT: 6-min walk test; FEV_1: forced expiratory volume in 1 s; FVC: forced vital capacity; RV/TLC: residual volume/total lung capacity.

persons declining participation, 66 met GOLD grade ⩾II severity. There were no statistically or clinically significant differences between subjects eligible for participation but declining (n=66), those participating in the trial (n=101) or those with data for both focal periods (n=84). For participants included within this analysis (n=84) significant differences between GOLD grade groups (n=40 grade II, n=31 grade III and n=13 grade IV) were evident for measures of pulmonary function, with clinically important differences for respiratory-related impairment [26], depression [27], CRQ-Dyspnoea [28] and 6MWT [29] (details in online supplementary material).

Model fit indices for confirmatory factor analysis of D-12 and MDP scores for each focal period (daily life and last minute of walk test) were comparable to the original development studies with all factor loadings ⩾0.47 and significant (details in online supplementary material) [8, 11, 12]. Exploratory principal components analysis identified two factors with the same allocation of items and similar degrees of variance as originally reported for each instrument [8, 11, 12]. For both instruments and focal periods, Cronbach's α was >0.80. No systematic measurement bias was evident for comparable scores of D-12 and MDP in either focal period (Bland–Altman plots; see online supplementary figure S1).

The D-12 and MDP consistently reflected the findings of VAS-I, VAS-U and descriptor assessments (convergent and discriminant validity). There were no significant differences between GOLD grade groups for any breathlessness assessment for either focal period (table 3, and online supplementary tables S5 and S6). Global measures of breathlessness intensity (VAS-I) or unpleasantness (VAS-U and MDP-A_1) did not differ significantly between focal periods (table 3). The most frequently selected MDP sensory quality that "most accurately describes" differed between focal periods (daily life: tight/constricted n=27 (37%); end of exercise: breathing a lot n=19 (28%)). Descriptors and D-12/MDP specific items scores reflecting sensory qualities (air hunger and tightness) or affective/emotional response were significantly more frequent and higher during daily life compared with the last minute of the walk test (table 3, and online supplementary tables S5 and S6). Volunteered descriptors for work/effort were the exception to this pattern (end of walk test 37% of participants *versus* 19% daily life).

TABLE 3 Breathlessness assessments: daily life responses/scores by Global Initiative for Obstructive Lung Disease (GOLD) grade and comparison between responses/scores for breathlessness in daily life and end of exercise

Descriptor categories[#]	Daily life responses/scores by GOLD grade			Daily life compared with end of exercise	
	II	III	IV	Daily life	Last minute of walk test
Subjects	40	31	13	84	84
Descriptors (categorical)					
Volunteered					
Air hunger/suffocating	18 (45)	11 (35)	4 (31)	34 (40)	30 (36)
Work/effort	10 (25)	5 (16)	1 (8)	16 (19)	31 (37)[¶]
Tight/constricted	5 (13)	1 (3)	1 (8)	7 (8)	5 (6)
Frightening/awful/worried	14 (35)	16 (52)	1 (8)	31 (37)	5 (6)[+]
Uncomfortable/annoying	14 (35)	12 (39)	7 (54)	33 (39)	11 (13)[+]
Helpless/regret/depressed	5 (13)	6 (19)	6 (46)	17 (20)	4 (5)[+]
Endorsed					
Air hunger	28 (70)	20 (65)	8 (62)	56 (67)	32 (38)[+]
Work	14 (35)	18 (58)	6 (46)	38 (45)	36 (43)
Tight	17 (43)	12 (39)	5 (38)	34 (40)	17 (20)[+]
Unnamed	25 (63)	22 (71)	8 (62)	55 (65)	66 (79)
Continuous scale instruments (score range)					
VAS					
Breathlessness intensity (0–100)	45±22	51±21	61±27	50±23	52±27
Breathlessness unpleasantness (0–100)	40±25	56±27	54±29	48±27	41±30
D-12					
Physical subscore (0–21)	8±5	9±6	9±4	8±5	7±6
Affective subscore (0–15)	4±4	4±4	4±3	4±4	1±3[+]
Total score (0–36)	11±8	13±9	13±7	12±8	8±8[+]
MDP					
Unpleasantness (A_1) (0–10)	5±2	5±2	5±2	5±2	4±3
Immediate Perception subscore (0–60)	19±14	20±12	26±14	20±13	17±14
Emotional Response subscore (0–50)	10±12	9±10	16±13	11±11	3±7[+]

Data are presented as n, n (%) or mean±SD. Complete reporting for each item within the Dyspnoea-12 (D-12) and Multidimensional Dyspnoea Profile (MDP), and comparisons between GOLD grades for breathlessness during the last minute of the walk test are presented in online supplementary tables S4 and S5. VAS: visual analogue scale. [#]: descriptor categories reported in Williams *et al.* [4] with data representing the participants volunteering/endorsing the descriptor within each category (participants could volunteer multiple descriptors categorised within different descriptor categories, hence the number and percentages do not add up to total sample size (n=84) or 100%). Bonferroni adjusted p-value for statistical significance: [¶]: p<0.005 (volunteered and endorsed descriptors); [+]: p<0.002 (for scalable instruments). In cases where the majority of cells were small, a Fisher's exact test confirmed findings.

Scores between scalable instruments (VAS-I, VAS-U, D-12 and MDP) were significantly correlated (daily life r=0.37–0.77 (figure 1) and end of exercise r=0.44–0.88 (figure 2)). In both focal periods consistent significant correlations (r −0.63–0.57) existed between scores of D-12/MDP and VAS (I and U), mMRC scale, HADS and CRQ with greater heterogeneity for 6MWT and measures of static pulmonary function (figures 3 and 4; r, 95% confidence intervals and z-tests, online supplementary tables S7 and S8). With the exception of developer-recommended single scores and VAS-I (daily life and end of exercise, online supplementary figure S2) and static measures of residual volume (RV) (RV % pred and RV/TLC % pred, figure 4), there were no significant differences between correlations of D-12 and MDP scores with VAS-U, mMRC scale, HADS, CRQ, static pulmonary function or 6MWT.

In both focal periods, participants volunteered fewer descriptors than they selected/rated in instruments with prescribed items (figure 5). Individual consistency across the descriptor list, D-12 and MDP was generally low (26–45%). For two-descriptor comparisons between the D-12 and MDP, the proportion of participants rating the item ⩾1 in one but not the other instrument differed significantly. All participants who provided a rating for the D-12 uncomfortable item also provided a rating for the MDP unpleasantness item (consistency n=58 (70%) daily life and n=47 (61%) end of exercise); however, 25 (daily life) and 30 (end exercise) participants rated this item in the MDP but not in D-12. For the descriptor work/effort (daily life), 46 participants rated the item in both the D-12 and MDP (consistency 64%), 20 participants provided a rating in the D-12 but not the MDP and six participants provided a rating in the MDP but not the D-12 (figure 5, and online supplementary tables S10 and S11).

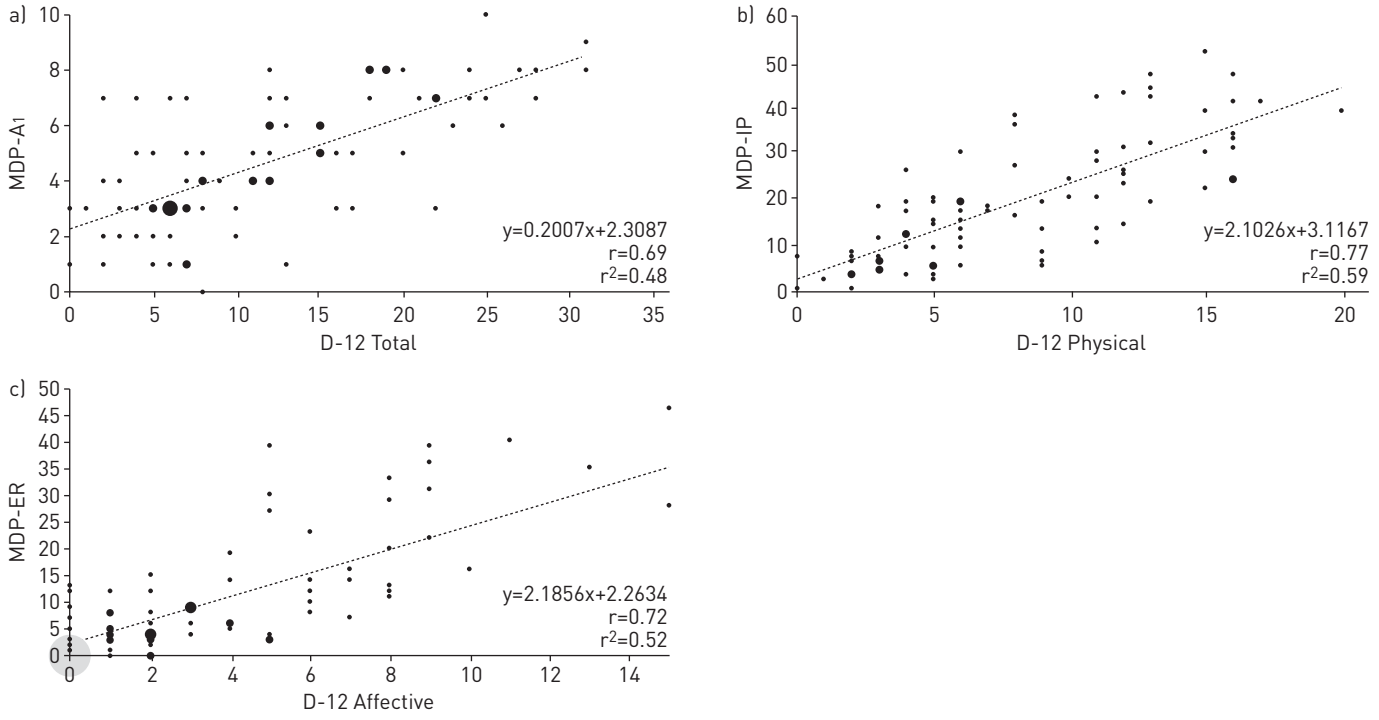

FIGURE 1 Breathlessness in daily life: associations (Pearson's correlation coefficient (r) and coefficient of determination (r^2)) between scores of the Multidimensional Dyspnoea Profile (MDP) and Dyspnoea-12 (D-12). Smallest marker represents n=1 participant (outliers identified using Bland–Altman plots; see online supplementary material). a) MDP-A1 (unpleasantness) (score range 0–10) *versus* D-12 Total (score range 0–36). Largest marker n=4 participants (r=0.80 without outliers (n=6)). b) MDP-IP (Immediate Perception) (score range 0–60) *versus* D-12 Physical (score range 0–21). Largest marker n=2 participants (r=0.79 without outliers (n=2)). c) MDP-ER (Emotional Response) (score range 0–50) *versus* D-12 Affective (score range 0–5). Largest marker (grey circle) n=11 participants (0,0) (r=0.82 without outliers (n=4)).

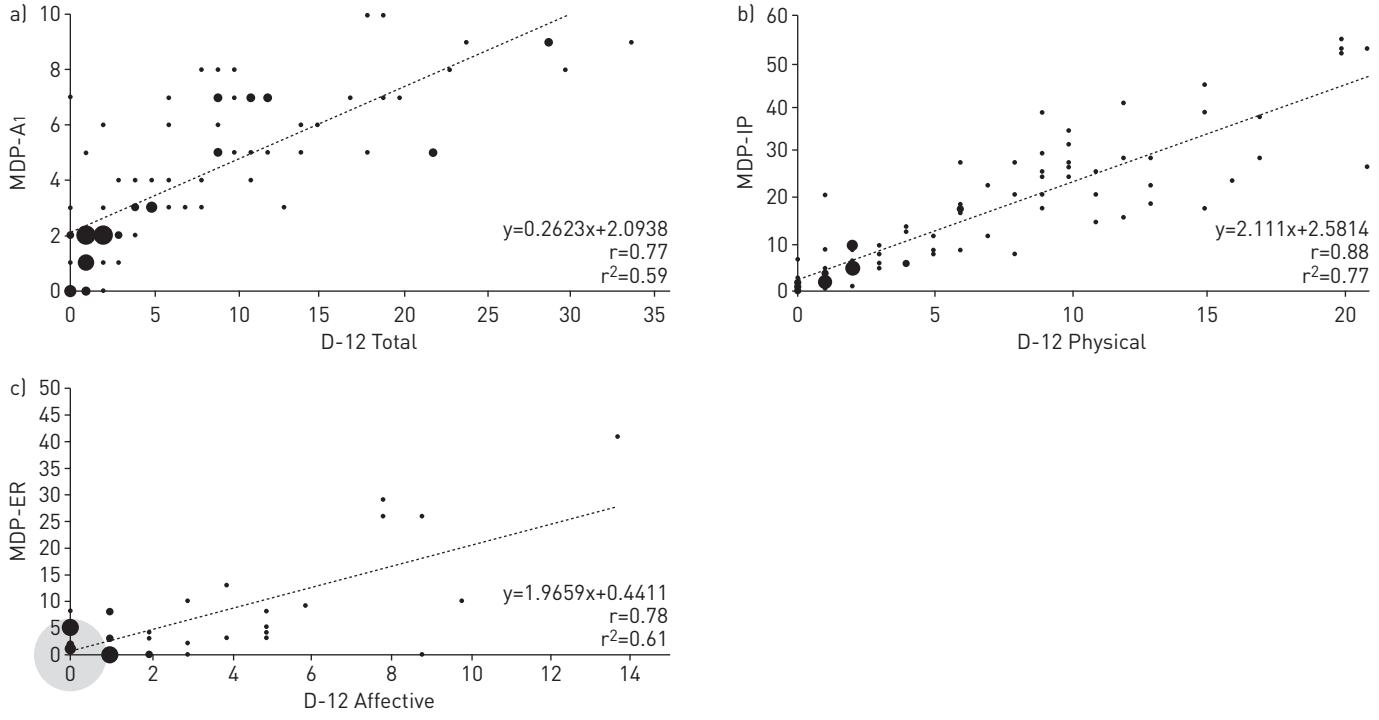

FIGURE 2 Breathlessness during last minute of walk test: associations (Pearson's correlation coefficient (r) and coefficient of determination (r^2)) between scores of the Multidimensional Dyspnoea Profile (MDP) and Dyspnoea-12 (D-12). Smallest marker represents n=1 participant (outliers identified using Bland–Altman plots; see online supplementary material). a) MDP-A1 (unpleasantness) (score range 0–10) *versus* D-12 Total (score range 0–36). Largest marker n=4 participants (r=0.80 without outliers (n=6)). b) MDP-IP (Immediate Perception) (score range 0–60) *versus* D-12 Physical (score range 0–21). Largest marker n=2 participants (r=0.79 without outliers (n=2)). c) MDP-ER (Emotional Response) (score range 0–50) *versus* D-12 Affective (score range 0–15). Largest marker (grey circle) n=11 participants (0,0) (r=0.82 without outliers (n=4)).

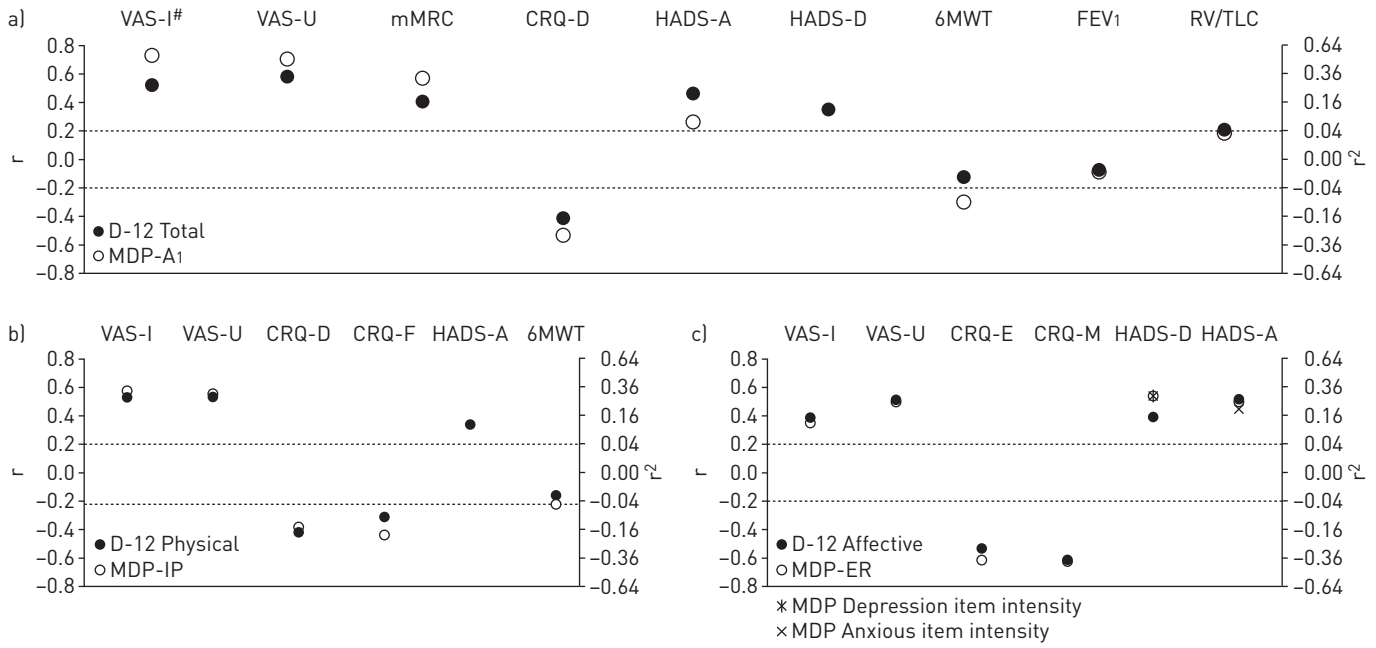

FIGURE 3 Breathlessness during daily life: example associations including Multidimensional Dyspnoea Profile (MDP) Depression and Anxious single-item comparison with the Hospital Anxiety and Depression Scale (HADS) for affective/emotional response subdomain (details in online supplementary material). a) Dyspnoea-12 (D-12) Total and MDP-A1 (unpleasantness). b) MDP-IP (Immediate Perception) and D-12 Physical. c) D-12 Affective and MDP-ER (Emotional Response). VAS-I: visual analogue scale (intensity (daily life)); VAS-U: visual analogue scale (unpleasantness (daily life)); mMRC: modified Medical Research Council scale; CRQ-D/F/E/M: Chronic Respiratory Questionnaire (Dyspnoea/Fatigue/Emotion/Mastery); HADS-A/D: Hospital Anxiety and Depression Scale (Anxiety/Depression); 6MWT: 6-min walk test (maximum distance); FEV1: forced expiratory volume in 1 s % pred; RV/TLC: residual volume/total lung capacity % pred. Dotted lines: r significant at >0.2. #: p≤0.002.

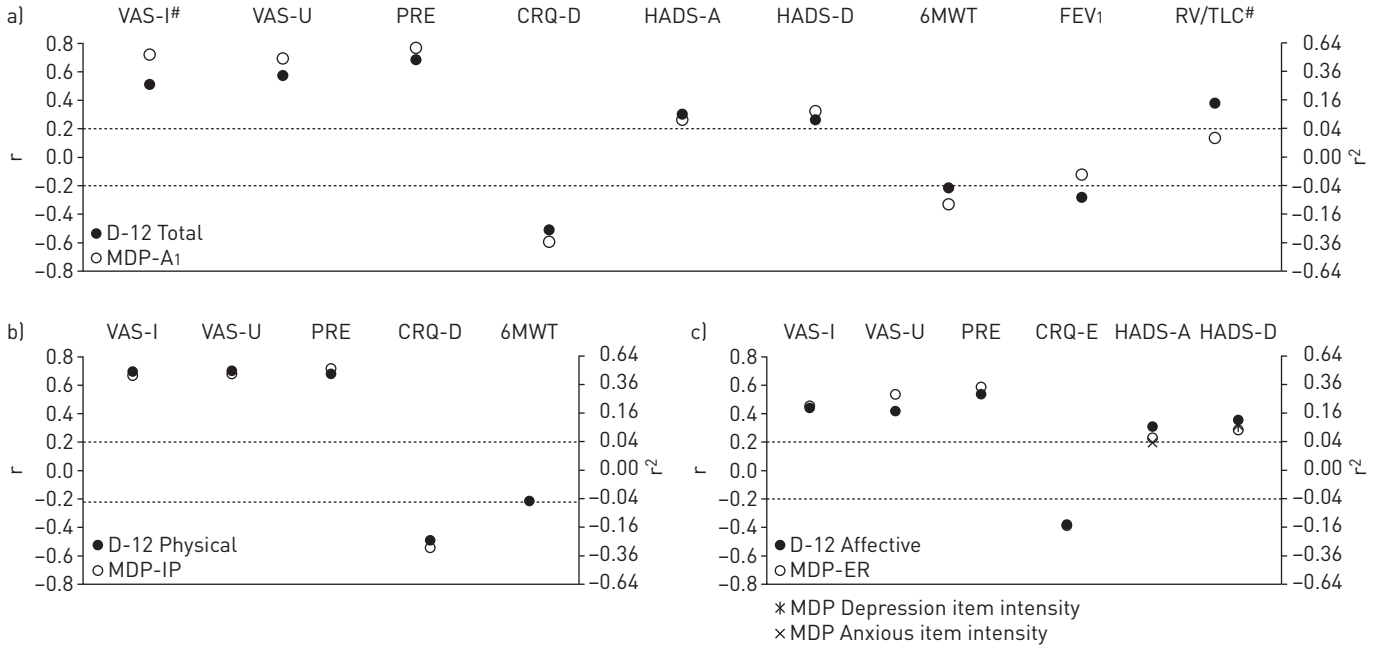

FIGURE 4 Breathlessness during last minute of walk test: example associations including Multidimensional Dyspnoea Profile (MDP) Depression and Anxious single-item comparison with the Hospital Anxiety and Depression Scale (HADS) for affective/emotional response subdomain (details in online supplementary material). a) Dyspnoea-12 (D-12) Total and MDP-A1 (unpleasantness). b) MDP-IP (Immediate Perception) and D-12 Physical. c) D-12 Affective and MDP-ER (Emotional Response). VAS-I: visual analogue scale (intensity (daily life)); VAS-U: visual analogue scale (unpleasantness (daily life)); PRE: perceived rate of exertion; CRQ-D/E: Chronic Respiratory Questionnaire (Dyspnoea/Emotion); HADS-A/D: Hospital Anxiety and Depression Scale (Anxiety/Depression); 6MWT: 6-min walk test (maximum distance); FEV1: forced expiratory volume in 1 s % pred; RV/TLC; residual volume/total lung capacity % pred. Dotted lines: r significant at >0.2. #: p≤0.002.

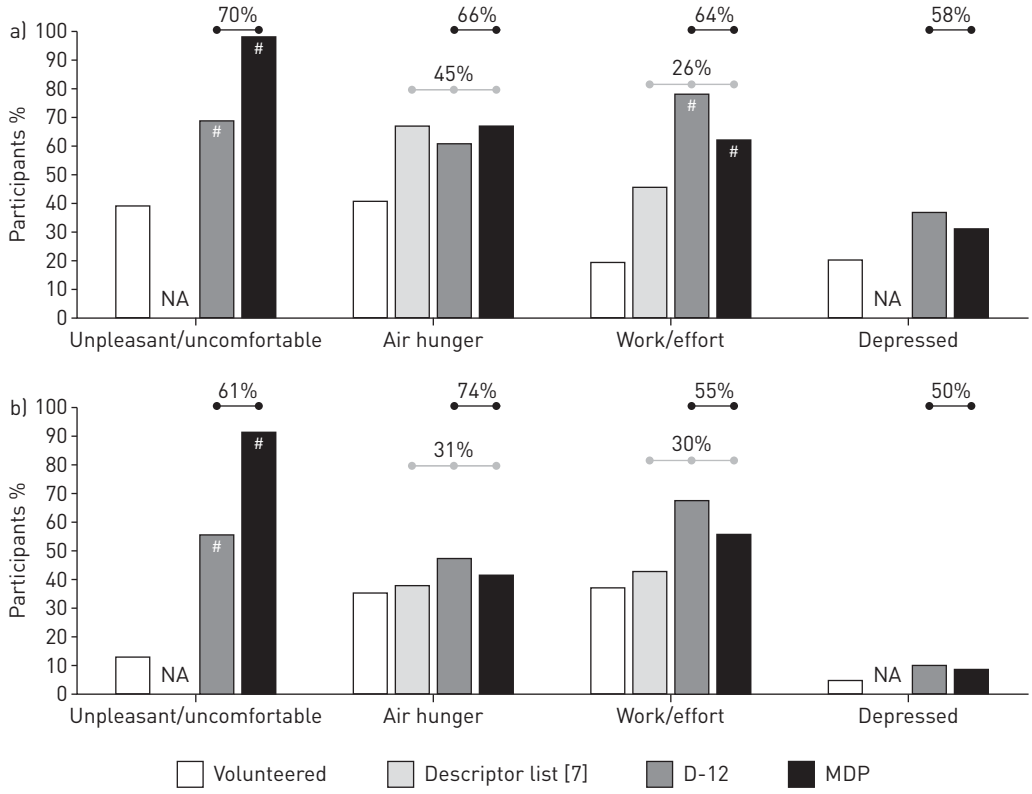

FIGURE 5 Individual consistency of descriptor choice: a) daily life and b) end of exercise. Data are presented as the percentage of participants indicating the item applied in each assessment (n=84). D-12: Dyspnoea-12; MDP: Multidimensional Dyspnoea Profile; NA: item not available within instrument. Percentages stated above dot-and-line markers indicate the same participants selecting/rating (≥1) comparable items across assessments (grey: descriptor list, D-12 and MDP; black: D-12 and MDP). #: Bonferroni adjusted p-value for statistical significance p≤0.006.

Discussion

To the best of our knowledge, this is the first study to directly compare the D-12 and MDP in the same sample of participants. In this cohort of people with moderate to severe COPD, developer-recommended single (D-12 Total and MDP-A1) and item group scores (subdomains) were highly correlated and demonstrated: 1) similar convergent validity with measures of the sensory-perceptual and affective distress of dyspnoea, and 2) similar concurrent validity with measures of static pulmonary function, functional exercise capacity and self-report measures of impairment. Individual consistency between the D-12 and MDP varied according to item, with 26–50% of participants indicating that a comparable item applied in one but not the other instrument. Item groupings (Physical/Affective, Immediate Perception/Emotional Response) reported by the developers of the D-12 and MDP were confirmed independently by our team for focal periods (daily life and end exercise) in this sample of Australians with stable COPD in a rehabilitation setting.

When used at the level of developer-recommended single or subdomain (item group) scores, the D-12 and MDP demonstrate broadly equivalent properties in terms of convergent, discriminant and concurrent validity. In general, D-12 and MDP scores (single or subdomain) tracked consistently with each other across all analyses. Where differences were evident in associations (VAS-I both focal periods, RV and end of walk test), this may be a function of the specific measurement scale (unidimensional scales (VAS-I and MDP-A1) *versus* multidimensional composite measures (D-12 Total)) or the low levels of unpleasantness provoked by the walk test.

There is increasing recognition that FEV_1 is weakly associated with the severity of breathlessness [13]. The functional consequences of airflow limitation in COPD, especially lung hyperinflation, may provide a physiological basis for sensations of dyspnoea both at rest and during exertion [6, 30]. In people with COPD, dynamic hyperinflation has been reported during 6MWTs [31, 32], short duration activities of daily living [33–35] and controlled strenuous exercise tests [5, 6]. While plethysmography has been reported to systematically overestimate static lung volumes in clinical settings [36], in this study the highest significant correlations between static lung function and D-12/MDP scores were between measures

of volume/capacity rather than airflow (daily life: inspiratory capacity % pred and MDP-A1 r=−0.27, 95% CI −0.47–−0.04; end of walk test: RV/TLC % pred and D-12 Total r=0.39, 95% CI 0.17–0.57), although the strength of association was modest.

Sensations of air hunger/tightness and affective/emotional response were perceived as significantly greater in daily life than at the end of the walk test. This discrepancy might be explained by participants operating well within their ventilatory limits during the walk test (suggested by Borg and breathlessness scores within the lower half of scales), self-limiting activity irrespective of context, inaccurate recall or psychological inferences between contexts (variable, unpredictable habitual environments *versus* supervised and safe clinical setting) [1]. An alternative interpretation is that recalled sensations reflect a particular form of recall bias (peak end rule) well documented in pain sciences [37]. As an evolutionary device, recalled sensations are constructed to prioritise salient parts (most meaningful and peak unpleasant/pleasant) of an experience as a basis for future decisions on participation or avoidance of the experience [37, 38]. In theory, if activities of daily living lead to sensations of unpleasantness, this may be sufficient for amplification of the peak experience, *i.e.* recalled sensations perceived as having greater intensity.

To date, there are few studies in people with stable COPD that report data for the D-12 [8] or MDP [39]. Our findings for associations between breathlessness (daily life) and self-reported impairment were similar to those reported in Yorke *et al.* [8] and Morelot-Panzini *et al.* [39], although instrument scores differed most likely as a result of differences in focal period: D-12 Total current study (n=84 average breathlessness in past 2 weeks) 12.0±8.0 *versus* 18±8 (n=53 breathlessness "these days" [8]); MDP median (interquartile range) current study (n=84 average breathlessness in past 2 weeks) MDP-A1 5 (3–7) *versus* 6 (3.5–7.0) (n=97 worst breathlessness in past 2 weeks) [39], MDP-IP 19 (9.0–30) *versus* 25 (13–37) [39] and MDP-ER 7 (3–14) *versus* 10 (3–20) [39].

While the D-12 and MDP share similar psychometric properties, these instruments differ considerably in intent and development, with consequent differences in items, instructions and response options. The intent of the D-12 was to calculate a single score of breathlessness severity for sensation in daily life with items derived from a potential 81 items through Rasch analysis to maximise optimal fit for a prospectively planned, unidimensional model [8]. The MDP was designed to use individual items selected from commonalities in factor analytic studies of descriptors (Immediate Perception) and pain-related instruments (Emotional Response), which can be grouped to calculate separate scores for sensation and emotion during a specified event or time in either laboratory or clinical settings [9]. Consequently, these two instruments assess different sensations and emotions with few directly comparable items [9].

While both the D-12 and MDP allow quantifiable, multidimensional assessments of breathlessness, in this group of people with stable COPD the MDP appeared to more completely capture the most salient sensations and emotions. Descriptors for tight/constricted are not included in the D-12, yet in this cohort this sensation was prevalent in both focal periods for the three other breathlessness assessments and notably, for breathlessness in daily life, was the most frequently selected in the MDP as "most accurately describes" (n=27 (37%)). For two of the four comparable items between the D-12 and MDP, there was a significant proportion of people rating the item in one but not the other instrument. This is likely to be a function of specific instructions/response scales where the D-12 requires respondents to consider whether the sensation/emotion "is troubling you" (implying impact), while the MDP requires only that the sensation/emotion is present ("how your breathing feels"). Alternatively, this may reflect responder interpretation of specific wording (*e.g.* unpleasantness may be interpreted to reflect greater affective distress than the term uncomfortable) or cultural differences in language [40]. At least one Australian participant volunteered the identical word for nine MDP items (four out of six IP items and five out of five ER items) compared with seven D-12 items (six out of seven Physical items and one out of five Affective items). In addition, "frightening" was the most common verbatim descriptor volunteered within our cohort (daily life). While "afraid/frightening" could be considered direct matches (MDP), "distressing" (D-12; a term which none of our participants volunteered) may be less so.

Limitations

The data reported in this study derive from pre-intervention measures for a clinical trial powered for primary end-points of the 6MWT and HADS. Consequently, the sample size may be underpowered for comparisons between breathlessness assessments and factor analysis. We strayed from the focal periods recommended by developers, differences between focal period scores for the D-12/MDP were small and consequently the inclusion of outliers may have underestimated the strength of association for a number of comparisons. Item group scores, rather than individual items for which the MDP was designed, were used for all comparisons and associations for individual MDP items were not systematically explored. Interpretation of the clinical relevance of changes in D-12/MDP scores is restricted by availability of directly applicable data for minimum clinically important differences (MCIDs). In people with lung

cancer, a MCID of 3 units has been suggested for the D-12 [41]. Although the MCID for the MDP has not yet been reported, a change in VAS of 10 mm or 1 unit in numeric rating scales for breathlessness intensity in people with chronic breathlessness has been suggested [42]. People in our study cohort were ambulatory, sufficiently stable and motivated to participate in pulmonary rehabilitation, with scores predominantly in the lower range of scales. Responder bias did not appear to be operating between those accepting/declining participation in the study; however, given these limitations it should not be assumed that the breathlessness assessments would perform similarly in other settings or contexts.

Conclusions

This study confirmed item grouping and convergent, discriminant and concurrent validity of the D-12 and MDP for breathlessness experienced in daily life and at the end of a common clinical exercise test in people with stable, obstructive pulmonary disease. Direct comparison of the D-12 and MDP demonstrated similar psychometric properties, but differences in intent, development, sensation/emotion items, response scales and scoring indicate that the D-12 and MDP serve different purposes, do not assess dyspnoea in the same way and are not interchangeable with each other.

Acknowledgements

The authors would like to thank and acknowledge the tireless work undertaken by Ms Hayley Lewthwaite, Ms Liz Lowe and Dr Louise Wiles, (University of South Australia, Adelaide, Australia) and Ms Anne Southwell (Repatriation General Hospital, Adelaide, Australia) in recruitment, data collection and data management; Dr Janelle Yorke (University of Manchester, Manchester, UK) for permission and advice for using the Dyspnoea-12; Associate Professors Robert Banzett (Harvard Medical School, Boston, MA, USA) and Mark Parshall (University of New Mexico, Albuquerque, NM, USA) for permission and advice on the Multidimensional Dyspnoea Profile and commentary on draft versions of the manuscript; and Mr Paul Cafarella (Director of the Comprehensive Pulmonary Rehabilitation Programme, Adelaide, Australia) and all participants who gave so freely and willing of their time.

Preliminary analyses related to this article were presented in oral form at the Australian Physiotherapy Association National Conference 2013 (Melbourne, Australia, October 17–20, 2013), Thoracic Society of Australia and New Zealand Annual Scientific Meeting 2015 (Gold Coast, Australia, March 27–April 1, 2015) and Dyspnea 2016 (Paris, France, June 16–17, 2016).

Author contributions: M.T. Williams takes responsibility for the content of the manuscript, including data and analysis. M.T. Williams and P. Firth developed the study design, M.T. Williams and D. John contributed substantially to the data management, analysis and interpretation. All three authors contributed to writing, interpretation of findings and critical review of the manuscript.

References

1 Williams MT, Garrard A, Cafarella P, *et al.* Quality of recalled dyspnoea is different from exercise induced dyspnoea: an experimental study. *Aust J Physiother* 2009; 55: 177–183.
2 Nishimura K, Yasui M, Nishimura T, *et al.* Airflow limitation or static hyperinflation: which is more closely related to dyspnea with activities of daily living in patients with COPD? *Respir Res* 2011; 12: 135.
3 Parshall MB, Schwartzstein RM, Adams L, *et al.* An Official Thoracic Society Statement: update on the mechanism, assessment and management of dyspnea. *Am J Respir Crit Care Med* 2012; 185: 435–452.
4 Williams MT, Cafarella P, Olds T, *et al.* Affective descriptors of sensations of breathlessness are more highly associated with severity of impairments than physical descriptors in people with COPD. *Chest* 2010; 138: 315–322.
5 Laveneziana P, Webb KA, Wadell K, *et al.* Does expiratory muscle activity influence dynamic hyperinflation and exertional dyspnea in COPD? *Respir Physiol Neurobiol* 2014; 199: 24–33.
6 Laveneziana P, Webb KA, Ora J, *et al.* Evolution of dyspnea during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011; 184: 1367–1373.
7 Mahler DA, Harver A, Lentine T, *et al.* Descriptors of breathlessness in cardiorespiratory diseases. *Am J Respir Crit Care Med* 1996; 154: 1357–1363.
8 Yorke J, Moosavi SH, Shuldham C, *et al.* Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010; 65: 21–26.
9 Banzett R, O'Donnell CR, Guilfoyle TE, *et al.* Multidimensional Dyspnoea Profile (MDP): an instrument for laboratory and clinical research. *Eur Respir J* 2015; 45: 1681–1691.
10 Lansing R, Gracely RH, Banzett RB. The multiple dimensions of dyspnoea: review and hypothesis. *Respir Physiol Neurobiol* 2009; 167: 53–60.
11 Meek PM, Banzett R, Parshall MB, *et al.* Reliability and validity of the Multidimensional Dyspnoea Profile. *Chest* 2012; 141: 1546–1553.
12 Parshall MB, Meek PM, Sklar D, *et al.* Test–retest reliability of Multidimensional Dyspnoea Profile recall ratings in the emergency department: a prospective, longitudinal study. *BMC Emerg Med* 2012; 12: 6.
13 Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management and Prevention of COPD. www.goldcopd.org Date last accessed: March 2, 2016.
14 Folstein M, Folstein S, McHugh P. "Mini-mental State". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189–198.
15 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
16 MacIntyre N, Crapo RO, Viegi G, *et al.* Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005; 26: 720–735.
17 Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.

18 American Thoracic Society Statement. Guidelines for the six minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.
19 Holland AE, Spruit MA, Troosters T, *et al.* An Official European Respiratory Society/American Thoracic Society technical standard: field walking tests in respiratory disease. *Eur Respir J* 2014; 44: 1428–1446.
20 Zigmond A, Snaith R. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; 67: 361–370.
21 Task Force on Surveillance for Respirator Hazards in the Occupational Setting. Surveillance for respiratory hazards in the occupational setting. *ATS News* 1982; 126: 952–956.
22 Schünemann HJ, Goldstein R, Mador MJ, *et al.* A randomised trial to evaluate the self administered standardised chronic respiratory questionnaire. *Eur Respir J* 2005; 25: 31–40.
23 Yorke J, Spencer LG, Duck A, *et al.* Cross-Atlantic modification and validation of the A Tool to Assess Quality of Life in Idiopathic Pulmonary Fibrosis (ATAQ-IPF-cA). *BMJ Open Respir Res* 2014; 1: e000024.
24 Banzett RB, Pedersen SH, Schwartzstein RM, *et al.* The affective dimension of dyspnea: air hunger is more unpleasant than work/effort. *Am J Respir Crit Care Med* 2008; 177: 1384–1390.
25 Meng XL, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. *Psychol Bull* 1992; 111: 172–175.
26 Crisafulli E, Clini EM. Measures of dyspnea in pulmonary rehabilitation. *Multidiscip Respir Med* 2010; 5: 202–210.
27 Puhan MA, Frey M, Büchi S, *et al.* The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2008; 6: 46.
28 Schünemann HJ, Puhan M, Goldstein R, *et al.* Measurement properties and interpretability of the Chronic Respiratory Disease Questionnaire. *COPD* 2005; 2: 81–89.
29 Puhan MA, Mador M, Held U, *et al.* Interpretation of treatment changes in 6-minute walk distance in patients with COPD. *Eur Respir J* 2008; 32: 637–643.
30 O'Donnell DE, Webb KA, Neder AJ. Lung hyperinflation in COPD: applying physiology to clinical practice. *COPD Res Practice* 2015; 1: 4.
31 Satake M, Shioya T, Uemura S, *et al.* Dynamic hyperinflation and dyspnoea during the six minute walk test in stable chronic obstructive pulmonary disease patients. *Int J Chron Obstruct Pulmon Dis* 2015; 10: 153–158.
32 Chen R, Lin l, Tian JW, *et al.* Predictors of dynamic hyperinflation during the 6-minute walk test in stable chronic obstructive pulmonary disease patients. *J Thorac Dis* 2015; 7: 1142–1150.
33 Silva CS, Nogueira FR, Porto EF, *et al.* Dynamic hyperinflation during activities of daily living in COPD patients. *Chron Respir Dis* 2015; 12: 189–196.
34 Garcia-Rio F, Lores V, Mediano O, *et al.* Daily physical activity in patients with chronic obstructive pulmonary disease is mainly associated with dynamic hyperinflation. *Am J Respir Crit Care Med* 2009; 180: 506–512.
35 Lahaije AJMC, van Hevoort HAC, Dekhuijzen PNR, *et al.* Physiologic limitations during daily life activities in COPD patients. *Respir Med* 2010; 104: 1152–1159.
36 O'Donnell C, Bankier A, Stiebellehner L, *et al.* Comparison of plethysmograghic and helium dilution lung volumes. Which is best for COPD? *Chest*; 2010; 137: 1108–1115.
37 Fredrickson BL. Extracting meaning from past affective experiences: the importance of peaks, ends and specific emotions. *Cogn Emot* 2000; 14: 577–606.
38 Stone AA, Schwartz JE, Broderick JE, *et al.* Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. *Pers Soc Psychol Bull* 2005; 31: 1340–1346.
39 Morelot-Panzini C, Gilet H, Aguilaniu B, *et al.* Real-life assessment of the multidimensional nature of dyspnoea in COPD outpatients. *Eur Respir J* 2016; 47: 1668–1679.
40 Banzett RB, O'Donnell CR. Should we measure dyspnoea in everyone? *Eur Respir J* 2014; 43: 1547–1550.
41 Yorke J, Lloyd-Williams M, Smith Blackhall F, *et al.* Management of the respiratory distress symptom cluster in lung cancer: a randomised controlled feasibility trial. *Support Care Cancer* 2015; 23: 3373–3384.
42 Johnson MJ, Bland JM, Oxberry SG, *et al.* Clinically important differences in the intensity of chronic refractory breathlessness. *J Pain Symptom Manage* 2013; 46: 957–963.