Validation of the 6 min walk test according to the OMERACT filter: a systematic literature review by the EPOSS-OMERACT group

Jérôme Avouac,[1] Otylia Kowal-Bielecka,[2] David Pittrow,[3] Doerte Huscher,[4] Frank Behrens,[5] Christopher P Denton,[6] Ivan Foeldvari,[7] Marc Humbert,[8] Marco Matucci-Cerinic,[9] Peter Nash,[10] Christian F Opitz,[11] Lewis J Rubin,[12] James R Seibold,[13] Oliver Distler,[14] Daniel E Furst[15]; for the EPOSS Group

ABSTRACT

Objective To assess the validity of the 6 min walk test (6MWT) in pulmonary arterial hypertension secondary to systemic sclerosis (PAH-SSc) according to the OMERACT filter.

Methods A systematic literature search was conducted from 1966 through June 2009. The assessment of validation of the 6MWT was based on the OMERACT filter criteria with the domains 'truth', 'discrimination' and 'feasibility'.

Results From the 57 articles identified, 9 (16%) were analysed. The 6MWT had face validity as it has been accepted by the registration agencies as a surrogate of PAH-SSc. It was sensitive to change in response to therapy, with effect sizes ranging from 0.30 to 1.37 with a parallel variation of haemodynamic parameters measured by right heart catheterisation. Feasibility was also validated. Content validity was not confirmed as this test was not specific for PAH-SSc. There was some evidence that 6MWT might meet criterion/construct validity, reproducibility and sensitivity to change over time, but insufficient data were provided to fully validate these components.

Conclusions Current evidence suggests partial validation of the 6MWT in PAH-SSc according to the OMERACT filter. Further dedicated studies are needed to validate completely the 6MWT in PAH-SSc, taking into account the comorbidities interfering with the 6MWT.

INTRODUCTION

Pulmonary arterial hypertension (PAH) has emerged as one of the leading causes of death in systemic sclerosis (SSc).[1] Despite therapeutic advances which improved the prognosis of this condition, outcome measures for PAH are poorly defined and those available are only partially validated or considered as not optimal.[2] Indeed, the evaluation of PAH is made difficult by other cardiopulmonary comorbidities associated with SSc (ie, interstitial lung disease or primary myocardial involvement). Recently, 78 experts in PAH-SSc from the OMERACT-EPOSS (Expert Panel on Outcome Measures in PAH related to SSc) Group defined a core set of outcome measures using the Delphi consensus methods.[3] Among these tools, the 6 min walk test (6MWT) is the most widely used as a primary outcome to assess therapeutic efficacy in PAH clinical trials and is accepted by the Food and Drug Administration. Although this outcome measure is

recommended by this expert group, it is necessary to ascertain if the 6MWT is validated in PAH-SSc since its aetiology and prognosis are not the same as idiopathic pulmonary arterial hypertension. The aim of this study was to analyse the status of validation of the 6MWT in SSc according to the criteria of the OMERACT filter.[4] Different from other approaches performed several years ago, we used an evidence-based systematic literature analysis for this assessment and included recent primary research articles from 1966 through June 2009, taking into consideration the large number of published articles in the field during recent years.

METHODS

Full details of the methodology are shown in the online data supplement.

Systematic literature review

Studies in which the 6MWT was used for the evaluation of patients with PAH/PH-SSc were searched in PubMed and the Cochrane Controlled Trial Register, as were relevant references from retrieved articles during the last 2 years. Inclusion and exclusion criteria emphasised articles in which patients with SSc-PAH were predominant or could be analysed separately and required ≥5 patients.

Quality assessment

Evidence was assessed with respect to the level of the evidence and the quality of randomised controlled trials (RCTs) was assessed according to the Jadad criteria.[5] [6] The quality criteria according to the definition of pulmonary hypertension are summarised in table 1.

Application of the OMERACT filter

The OMERACT process involves consensus on outcome measures and is based on the 'OMERACT filter' which is composed of three key components: truth (face, content and criterion/construct validity), discrimination (reliability, sensitivity to change and discriminant capacity in response to therapy) and feasibility.[4]

RESULTS

Results of the systematic literature search

From the 57 articles identified, 9 were finally included for further analysis (see figure in online supplement).[7–15]

Concise report

Table 1 Quality assessment of studies according to the definition of pulmonary arterial hypertension and the exclusion of other forms of pulmonary hypertension

Definition of PAH	Pulmonary fibrosis and left heart disease excluded	Pulmonary fibrosis and left heart disease not excluded
RHC		
mPAP >25 mm Hg at rest and/or mPAP >30 mm Hg at exercise	A1	A2
Doppler echo		
PASP/TG ≥45 mm Hg	B1	B2
Doppler echo		
35 mm Hg ≤ PASP/TG <45 mm Hg	C1	C2
Other (or not defined)	D1	D2

For detailed definition of quality criteria A–D and category 1/2, please refer to the Methods section and the online supplement. Only if A1 studies were available, specific OMERACT criteria of 6MWT were considered validated (V) or not valid (NV) (see legend to table 2 for detailed definitions).
6MWT, 6 min walk test; mPAP, mean pulmonary artery pressure; PAH, pulmonary arterial hypertension; PASP/TG, pulmonary arterial systolic pressure/tricuspid gradient; RHC, right heart catheterisation.

Quality assessment of retrieved articles

Five studies were RCTs (level 1b),[7 10 13–15] one was a longitudinal study with comparison to a historical control group (level 2b)[12] and three were descriptive studies (level 3).[8 9 11] The mean±SD Jadad score for RCTs was 4.1±1.0. Seven studies included well-defined PAH-SSc subgroups according to quality level A1[7 9 10 12–15] and two studies were classified C1[8 11] (see table in online supplement).

Status of validation according to the OMERACT criteria for the 6MWT

Truth

Face validity

An agreement by a group of experts is required to define the appropriateness of 6MWT to be used for the evaluation of PAH (table 2). As it was accepted by the registration agencies of the USA and Europe as a surrogate of PAH and selected by the EPOSS Group as part of the core set of outcome measures for SSc, this test clearly has face validity.

Content validity

6MWT did not pass this specific filter as no paper evaluated 6MWT in a wide-ranging and comprehensive group of patients with SSc. Moreover, indirect evidence suggests that the 6MWT is not specific for PAH-SSc-related exercise impairment. Thus, in three different studies, 6MWT was also associated in multivariate analysis with measures of myocardial involvement, interstitial lung disease, musculoskeletal disorders and inflammation.[8 9 11] While there is no study available in SSc-PAH with the highest quality level A1, the expert group agreed that, based on the evidence from non-PAH studies, the 6MWT will very likely not meet this criterion and rated content validity as not validated.

Criterion validity/construct validity

Since the 6MWT was never tested against the external or 'gold' standard (right heart catheterisation (RHC)) in any study of patients with SSc-PAH, the 6MWT does not meet this aspect of validation. As an alternative, correlations with other validated parameters were searched to estimate concurrent and convergent validity as surrogates for criterion validity and as indicators of overall construct validity. One study found that 6MWT was significantly associated with mortality in univariate analysis. Survival increased by 24% for every 100 m further a patient could walk (HR 0.76, 95% CI 0.58 to 0.98, p=0.037). However,

Table 2 Validation of 6MWT in PAH-SSc according to the OMERACT filter

OMERACT filter criterion	Validation	Highest quality of PAH definition
Truth		
Face validity	V	NA
Content validity	NV	C1
Criterion validity	ND	No studies
Construct validity	PV	A1
Discrimination		
Sensitivity to change over time	PV	A1 (placebo groups)
Discrimination capacity in response to therapy	V	A1
Reliability (reproducibility)	PV	Studies in SSc-ILD
Feasibility	V	A1

V, valid: a criterion was judged validated if appropriate information was available from studies including exclusively patients with PAH-SSc (quality definition A1, see table 1 for details). Exception is face validity, which is evaluated by the judgement of experts as an appropriate measure rather than by specific studies. NV, not valid: similarly, a criterion was judged not valid if appropriate information was available from studies including exclusively patients with PAH-SSc (quality definition A1, see table 1 for details). PV, partially validated: a criterion was judged partially validated if data from studies lower than quality level A1 indicated that the criterion is validated.
6MWT, 6 min walk test; NA, not applicable; ND, no data; PAH, pulmonary arterial hypertension; PAH-SSc, pulmonary arterial hypertension secondary to systemic sclerosis.

this result was not confirmed in the multivariate analysis. In this study the 6MWT was not an independent risk factor for decreased survival in patients with PAH-SSc (HR 0.81, 95% CI 0.59 to 1.12, p=0.20).[12] At this time the 6MWT has therefore not met the test of criterion/construct validity, but it was rated partially validated for construct validity by the expert group as there is some evidence that it might correlate with survival.

Discrimination

Insufficient data were provided in the analysed studies to assess the reliability and reproducibility of the 6MWT in the specific situation of PAH-SSc (κ coefficient and intraclass and interclass correlation coefficients were not assessed). However, because tests in SSc-ILD indicated reproducibility, it was rated partially validated.[16] It was not possible to assess the sensitivity to change over time because there were no suitable data to analyse. As an alternative, we assessed the sensitivity to change of the 6MWT in the placebo groups of five high-quality RCTs.[7 10 13–15] These studies provided data on exercise capacity for 100 patients included in the placebo groups. Unfortunately, there were insufficient data to calculate the standardised response mean in all studies and thus complete a meta-analysis. However, it is noteworthy that, over time, there was a deterioration in the mean distance walked in this specific group in four studies (deterioration from 2.6 m to 38 m) and no variation in one study (improvement of 3±8 m). Thus, with regard to these data, the 6MWT was rated by the experts as at least partially validated.

Discriminant capacity in response to therapy was validated in the five high-quality RCTs described above. These studies provided data on exercise capacity for 283 patients with SSc. The mean trial duration was 13±4.0 weeks. The 6MWT was sensitive to change in response to therapy with effect sizes ranging from 0.30 to 1.37 (table 3). Moreover, there was also a parallel variation between the effect sizes of the 6MWT and the effect sizes of haemodynamic parameters (mean pulmonary artery pressure (mPAP) and pulmonary vascular resistance (PVR)) during treatment, as measured by RHC in the analysed trials (table 3).

Feasibility

The 6MWT is easy to perform and values are easy to interpret with no special equipment required; minimal training is required

Table 3 Validation of the discriminant capacity over treatment of the 6MWT

Author, year, reference	Quality (Jadad)	SSc patients	Inclusion criteria for PAH	Intervention	Comparator	Study duration	Effect size (95% CI) for 6MWT	Effect size (95% CI) for haemodynamic measures
Denton et al, 2006[7]	5/5	N=52 (79%)	CTD-related PAH WHO III/IV 6MWT 150–500 m mPAP >25 mm Hg, PVR >240	Bosentan 62.5 mg ×2/day during 4 weeks then 125×2 mg/day 250 mg/day	Placebo	12–16 weeks	0.30 (−0.2 to 0.8)	mPAP No data PVR No data
Girgis et al, 2007[14]	4/5	N=38 (46%)	CTD-related PAH WHO II, III and IV 6MWT >450 m excluded mPAP >25 mm Hg at rest PVR ≥240	Sitaxsentan 50 mg/day 100 mg/day 300 mg/day	Placebo	12 weeks	0.57 (−0.2 to 1.3)	mPAP −0.9 (−1.67 to 0.1) PVR −1.1 (−1.9 to 0.3)
Badesch et al, 2007[15]	5/5	N=38 (45%)	CTD-related PAH WHO II-IV 6MWT 100–450 m mPAP ≥25 mm Hg	Sildenafil 20 mg × 3/day 40 mg × 3/day 80 mg × 3/day	Placebo	12 weeks	0.65 (0.0 to 1.3) 0.68 (0.0 to 1.3) 0.31 (−0.2 to 1.0)	mPAP −0.5 (−1.2 to 0.1) −0.4 (0.9 to 0.2) −0.4 (−1.0 to 0.2) PVR −0.5 (−1.1 to 0.1) −0.3 (−0.9 to 0.3) −0.3 (−0.9 to 0.3)
Badesch et al, 2000[13]	3/5	N=110 (100%)	SSc-related PAH mPAP ≥35 mm Hg PVR >240	Epoprostenol intravenously Starting from 2 ng/kg/min intravenously and increased based on symptoms and adverse events plus conventional therapy	Conventional therapy	12 weeks	1.2 (0.8 to 1.6)	mPAP −0.6 (−0.9 to −0.2) PVR −0.9 (−1.3 to −0.5)
Oudiz et al, 2004[10]	3/5	N=45 (50%)	SSc-related PAH mPAP ≥25 mm Hg PVR ≥3 Wood units 6MWT 50–450 m	Treprostinil intravenously Starting from 1.25 to 2.5 ng/kg/min and increases of 1.25–2.5 ng/kg/min every 1–2 weeks. Target dose based on response to therapy and adverse effects and did not exceed 22.5 ng/kg/min	Placebo	12 weeks	1.37 (0.9 to 1.9)	mPAP −1.0 (−1.4 to −0.6) PVR −2.3 (−2.8 to −1.8)

6MWT, 6 min walk test; mPAP, mean pulmonary artery pressure; PAH, pulmonary arterial hypertension; PVR, pulmonary vascular resistance; SSc, systemic sclerosis.

and patient/physician acceptance is good. It is also an inexpensive test that requires a minimum number of healthcare staff and can be performed in an office setting.[11] Thus, availability of the 6MWT at investigational sites and its practicality both in specialised and community practices make it an eminently feasible test.[17] As a limitation, there might be rare disease-related manifestations such as arthritis or severe skin fibrosis where the test cannot be performed.

DISCUSSION

This is the first study to address the validity of 6MWT as an outcome measure in PAH-SSc, with respect to the OMERACT filter, using a systematic literature review. Current evidence suggests that the 6MWT passes the filter for face validity, discriminant capacity with treatment and feasibility. However, content validity is not validated and no robust data allow full confirmation of construct/criterion validity, reliability and sensitivity to change over time (table 3).

The heterogeneity of the trials limited the analysis to only a few studies with patients with PAH-SSc diagnosed according to the highest quality level A1. Many studies included patients with pulmonary hypertension related to PAH, left heart disease and severe interstitial lung disease, without subgroup analysis which allowed one to examine SSc-PAH per se.

Content validity was considered by the expert group as not valid, although one study showed that impairment of exercise performance was associated with pulmonary hypertension.[18] However, this study had a low C2 quality level and was excluded from the initial analysis as exercise capacity was not assessed by the 6MWT. Moreover, recent data have shown that patients with SSc judged free of both cardiac and pulmonary diseases had reduced exercise capacity, suggesting that some SSc-specific or non-specific comorbidities independent of cardiopulmonary involvement, such as articular/soft tissue involvements, may influence exercise capacity in SSc.[8 19 20] This has an impact on future trial design in patients with SSc-PAH. If the 6MWT is chosen as an outcome, patients with significant comorbidities either need to be excluded or need to be controlled for by appropriate tests.

Despite insufficient data to assess the reliability and reproducibility of the 6MWT in PAH-SSc per se, follow-up analyses by Buch et al[16] showed the high reproducibility of the 6MWT in patients with SSc-ILD. The within-subject intertest Pearson correlation coefficient on repeated 6MWT was reported to be 0.95 (p<0.001). Based on all of the above, reproducibility was rated by experts as partially validated.

We used placebo groups of RCTs to assess the sensitivity to change over time. One could argue that placebo is a form of

intervention and may be a bias. This explains why experts rated this item as partially validated (table 2). However, it appears likely that the 6MWT also shows sensitivity to change without placebo intervention. Although the 6MWT had discriminant capacity with treatment, this result has to be taken cautiously because most of the effect sizes calculated were not significant. This may have been related to the variability of the testing procedures or to the low number of patients with SSc included (only five trials allowed subgroup analyses of pivotal RCTs on approximately 20–25% of the total population). On the other hand, two high-quality RCTs, performed in a population of patients with PAH-SSc with sufficient power to detect a significant difference, clearly showed the sensitivity to change of the 6MWT when the patients were treated with epoprostenol and treprostinil.[13]

This structured literature review and the assessment of identified papers according to the OMERACT filter revealed several aspects of 6MWT that need further validation in additional studies (criterion/construct validity, reliability and sensitivity to change over time). Although the OMERACT filter is not an obligatory framework to assess the full validation of 6MWT as an outcome measure in PAH-SSc, it defines a specific research agenda which needs to be addressed.

Definite validation of construct/criterion validity of the 6MWT requires significant correlation between the gold standard RHC and the 6MWT specifically in PAH-SSc cross-sectional studies.

Validation of reliability of the 6MWT in PAH-SSc requires comparisons of repeated 6MWT assessments performed within a short time period by the same investigator (intraobserver variability) and by two independent investigators (interobserver variability) at the same time in patients with well-defined PAH-SSc.

The type of studies required to assess the validity of 6MWT with respect to the sensitivity to change are longitudinal studies including patients with PAH-SSc without treatment and parallel 6MWT and RHC evaluations at different time points.

Our study did not assess the 6MWT in patients with heart or lung disease in general. Further dedicated studies are thus needed to assess the 6MWT in patients with SSc with interstitial lung disease or primary myocardial involvement.

In conclusion, this systematic literature analysis showed that 6MWT in PAH-SSc fulfils the OMERACT criteria only for face validity, discriminant capacity with treatment and feasibility. Content validity was considered as not valid. Further studies are therefore needed to completely validate the 6MWT in PAH-SSc for construct/criterion validity, reliability and sensitivity to change.

Authors affiliations [1]Paris Descartes University, Rheumatology A Department, Cochin Hospital, APHP, Paris, France
[2]Department of Rheumatology and Internal Medicine, Medical University of Bialystok, Bialystok, Poland
[3]Institute for Clinical Pharmacology, Technical University, Dresden, Germany
[4]German Rheumatism Research Centre, Berlin, Germany
[5]Division of Rheumatology/ZAFES, J.W. Goethe University, Frankfurt, Germany
[6]Centre for Rheumatology, Royal Free and University College, Medical School, London, UK
[7]Hamburger Zentrum für Kinder- und Jugendrheumatologie Kompetenz-Zentrum für Sklerodermie im Kindesalter Am Klinikum Eilbek, Hamburg, Germany
[8]Service de Pneumologie et Reanimation Respiratoire, Centre des Maladies Vasculaires Pulmonaires, Hopital Antoine-Beclere, Université Paris-Sud, Clamart, France
[9]Department of Medicine, Division of Rheumatology, Denothe Center, University of Florence, Florence, Italy
[10]Rheumatology Research Unit, Cotton Tree, Sunshine Coast, Queensland, Australia
[11]Medizinische Klinik II, DRK-Kliniken Berlin, Westend, Berlin, Germany
[12]Division of Pulmonary and Critical Care Medicine, University of California, San Diego, School of Medicine, La Jolla, California, USA
[13]University of Michigan Scleroderma Program, Ann Arbor, Michigan, USA
[14]Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland
[15]Division of Rheumatology, Department of Medicine, David Geffen School at UCLA, Los Angeles, California, USA

REFERENCES

1. **Allanore Y**, Avouac J, Wipff J, et al. New therapeutic strategies in the management of systemic sclerosis. *Expert Opin Pharmacother* 2007;**8**:607–15.
2. **Merkel PA**, Clements PJ, Reveille JD, et al. Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. *J Rheumatol* 2003;**30**:1630–47.
3. **Distler O**, Behrens F, Pittrow D, et al. Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. *Arthritis Rheum* 2008;**59**:867–75.
4. **Boers M**, Brooks P, Strand CV, et al. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998;**25**:198–9.
5. **Shekelle PG**, Woolf SH, Eccles M, et al. Clinical guidelines: developing guidelines. *BMJ* 1999;**318**:593–6.
6. **Jadad AR**, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
7. **Denton CP**, Humbert M, Rubin L, et al. Bosentan treatment for pulmonary arterial hypertension related to connective tissue disease: a subgroup analysis of the pivotal clinical trials and their open-label extensions. *Ann Rheum Dis* 2006;**65**:1336–40.
8. **Garin MC**, Highland KB, Silver RM, et al. Limitations to the 6-minute walk test in interstitial lung disease and pulmonary hypertension in scleroderma. *J Rheumatol* 2009;**36**:330–6.
9. **Schoindre Y**, Meune C, Dinh-Xuan AT, et al. Lack of specificity of the 6-minute walk test as an outcome measure for patients with systemic sclerosis. *J Rheumatol* 2009;**36**:1481–5.
10. **Oudiz RJ**, Schilz RJ, Barst RJ, et al. Treprostinil, a prostacyclin analogue, in pulmonary arterial hypertension associated with connective tissue disease. *Chest* 2004;**126**:420–7.
11. **Villalba WO**, Sampaio-Barros PD, Pereira MC, et al. Six-minute walk test for the evaluation of pulmonary disease severity in scleroderma patients. *Chest* 2007;**131**:217–22.
12. **Williams MH**, Das C, Handler CE, et al. Systemic sclerosis associated pulmonary hypertension: improved survival in the modern era. *Heart* 2006;**92**:926–32.
13. **Badesch DB**, Tapson VF, McGoon MD, et al. Continuous intravenous epoprostenol for pulmonary hypertension due to the scleroderma spectrum of disease. A randomized, controlled trial. *Ann Intern Med* 2000;**132**:425–34.
14. **Girgis RE**, Frost AE, Hill NS, et al. Selective endothelin A receptor antagonism with sitaxsentan for pulmonary arterial hypertension associated with connective tissue disease. *Ann Rheum Dis* 2007;**66**:1467–72.
15. **Badesch DB**, Hill NS, Burgess G, et al. Sildenafil for pulmonary arterial hypertension associated with connective tissue disease. *J Rheumatol* 2007;**34**:2417–22.
16. **Buch MH**, Denton CP, Furst DE, et al. Submaximal exercise testing in the assessment of interstitial lung disease secondary to systemic sclerosis: reproducibility and correlations of the 6-min walk test. *Ann Rheum Dis* 2007;**66**:169–73.
17. **Impens AJ**, Wangkaew S, Seibold JR. The 6-minute walk test in scleroderma--how measuring everything measures nothing. *Rheumatology (Oxford)* 2008;**47**(Suppl 5):v68–9.
18. **Morelli S**, Ferrante L, Sgreccia A, et al. Pulmonary hypertension is associated with impaired exercise performance in patients with systemic sclerosis. *Scand J Rheumatol* 2000;**29**:236–42.
19. **Avouac J**, Guerini H, Wipff J, et al. Radiological hand involvement in systemic sclerosis. *Ann Rheum Dis* 2006;**65**:1088–92.
20. **de Oliveira NC**, dos Santos Sabbag LM, Ueno LM, et al. Reduced exercise capacity in systemic sclerosis patients without pulmonary involvement. *Scand J Rheumatol* 2007;**36**:458–61.

Validation of the 6 min walk test according to the OMERACT filter: a systematic literature review by the EPOSS-OMERACT group

Jérôme Avouac, Otylia Kowal-Bielecka, David Pittrow, et al.

	Updated information and services can be found at:
	http://ard.bmj.com/content/early/2010/05/11/ard.2009.120303.full.html

These include:

Data Supplement	*"Web Only Data"* http://ard.bmj.com/content/suppl/2010/05/20/ard.2009.120303.DC1.html
References	This article cites 20 articles, 14 of which can be accessed free at: http://ard.bmj.com/content/early/2010/05/11/ard.2009.120303.full.html#ref-list-1
P<P	Published online May 14, 2010 in advance of the print journal.
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/