INTRODUCTION

Reduced physical fitness is characteristic of patients with chronic heart failure (HF), usually affecting daily physical activities such as walking, but also exercise training and testing (Chien, Lee, Wu, Chen, & Wu, 2008; Piepoli, 2013). The six-minute walk test (6MWT) and the cardiopulmonary exercise test (CPET) are probably the two most common methods for measuring the extent of exercise tolerance in patients with HF (Guazzi, Dickstein, Vicenzi, & Arena, 2009; Ingle, Cleland, & Clark, 2014). The availability of CPET is often limited in clinical practice (Arena, Myers, & Guazzi, 2011; Corra et al., 2010; Ingle et al., 2014), and alternative, easily accessible tests have therefore emerged.

The 6MWT is a simple, standardized, submaximal and well-tolerated method of assessing functional working capacity at a low cost, and needs no expensive equipment (Adsett

Abstract

Purpose: The aims of this study were to determine the test–retest reliability of the duplicated six-minute walk test (6MWT) in patients with chronic heart failure (HF), and to evaluate its variation over time.

Methods: Forty-six patients (9 women) with HF performed duplicated 6MWT every third month for 1 year (5 follow-ups), for a total of 198 paired tests. The patients completed two 6MWT on the same day with a 45-min seated rest between tests.

Results: The mean distance in metres, for the first (6MWT1) versus the second (6MWT2), for each follow-up, was 408 ± 100 versus 411 ± 96, 449 ± 94 versus 465 ± 94, 464 ± 96 versus 473 ± 100, 462 ± 103 versus 468 ± 104 and 472 ± 105 versus 482 ± 107. On average, a marginally, clinically insignificant longer walked distance, 9 m (2.0%), was seen in the second 6MWT. The standard error of a single determination ($S_{\text{method}}$) ranged from 2.4% to 3.9% over the study period, and the intraclass correlation coefficient (ICC) ranged from 0.96 to 0.99 (CI 95% 0.94–0.99). The variation over time of ICC or $S_{\text{method}}$ was not statistically significant.

Conclusion: The 6MWT is highly reliable over time in patients with HF, and one test is, therefore, sufficient in clinical follow-ups.

KEYWORDS

cardiac rehabilitation, exercise test, longitudinal follow-up, reproducibility, walk test
et al., 2011; Faggiano, D’Aloia, Gualeni, Brentana, & Dei, 2004; Ross, Murthy, Wollak, & Jackson, 2010). The distance walked (6MWD) during a 6MWT has also shown good correlation with peak oxygen uptake, (Reilly & Tipton, 2010; Zugck et al., 2000), in patients with HF, which is interesting since the oxygen consumption (VO\textsubscript{2}) has a prognostic value, and formulas for calculating the peakVO\textsubscript{2} from 6MWD are available (Ross et al., 2010). The main outcome in the 6MWT is walked distance, and according to the guidelines of the American Thoracic Society, ATS (ATS statement, 2002), this measure of physical function is suitable for follow-up over time. The 6MWT is often performed for follow-up in clinical practice as well as in clinical trials but the need for repeating the 6MWT twice at follow-up assessments is uncertain (ATS statement, 2002). The repeatability is influenced by methodological factors, learning effects, but may also differ between diseases. (Cahalin, Mathier, Semigran, Dec, & DiSalvo, 1996; Dolmage, Hill, Evans, & Goldstein, 2011; Guazzi et al., 2009; Guyatt et al., 1985; Hanson, McBurney, & Taylor, 2012; Hopkins, 2000; Kervio, Ville, Leclercq, Daubert, & Carre, 2004; O’Keeffe, Lye, Donnellan, & Carmichael, 1998; Purser et al., 1999; Spencer, Alison, & McKeough, 2008; Wu, Sanderson, & Bittner, 2003).

Most patients with heart failure are in need of care for several years, and physical function (shorter 6MWD and lower peakVO\textsubscript{2}) has a prognostic value. We therefore found it relevant to evaluate the 6MWT over a longer period of time to decide whether it is necessary with duplicated tests over time, as the walk itself can be strenuous, especially in the event of deterioration. Furthermore, studies of test–retest reliability of the 6MWT over time are scarce in patients with HF.

The aim of this study was to determine the long-term test–retest reliability of the 6MWT in patients with HF.

2 | METHODS

2.1 | Study population

This longitudinal reliability study reports data from two randomized controlled trials designed to evaluate physical fitness in an exercise programme of which one, thus far, has been published (Lans, Cider, Nylander, & Brudin, 2018). Patients provided measures at baseline and at 3, 6, 9 and 12 mo. Patients were enrolled from the departments of cardiology or clinical physiology at a County Hospital. The study included patients with a left ventricular ejection fraction (EF) by echocardiography ≤ 40%, age ≤ 80 years and New York Heart Association (NYHA) functional classes II and III. Exclusion criteria were physical or mental disorders that limited the ability to perform the 6MWT. The Regional Ethical Review Board, Linköping, Sweden, approved the study (Dnr99266, Dnr02-041), and all included patients gave written informed consent to participate. The investigation conforms to the principles outlined in the Declaration of Helsinki.

2.2 | Protocol

The 6MWT took place in an indoor, flat 80-m hospital corridor, marked at 2.5-m intervals. Cones were placed at both ends of the course. At each follow-up, two 6MWTs (Test 1 and Test 2) were performed on the same day with a 45-min seated rest between tests. Standardized oral instructions on performance of 6MWT, modified from ATS’s recommendations, and how to use the Borg scales were provided (ATS statement, 2002; Borg, 1998; Demers, McKelvie, Negassa, & Yusuf, 2001). The patients were instructed to walk at a self-selected walking pace for 6 min, but still cover as much distance as possible. During the test (while measuring time), the patients were allowed to stop and rest if needed, and then continue the 6MWT as soon as they felt able. No encouragement was given during the test. The total walking distance was measured. Heart rate was monitored (Polar®, Kempele, Finland) before and immediately after the 6MWT, and the patients were asked to rate, on Borg’s scales, their perceived exertion (RPE 6–20), feelings of breathlessness and chest pain (category ratio 0–10 (CR-10)) (Borg, 1998). A stopwatch was used to record the time. All 6MWTs were supervised by a physiotherapist who did not walk with the patients.

All patients underwent an exercise programme for one year, either peripheral muscle training or bicycling and walking, initially in a hospital-setting and continued with home-based training. All walking tests were performed on separate occasions, and not on the days of the training sessions.

2.3 | Statistics

Data were analysed by descriptive and inferential statistics. Reliability between paired measurements was calculated as follows: the Bland–Altman plot was used to describe the difference between the paired 6MWT calculated individually at all follow-ups (Bland & Altman, 1986). The intraclass correlation coefficient (ICC) is the correlation between two observations made on the same subject. An ICC value > 0.75 was considered adequate (O’Keeffe et al., 1998) and >0.9 was considered excellent (Demers et al., 2001). The standard error of a single determination (S\textsubscript{method}) was calculated as $S_{\text{method}} = \sqrt{(\Sigma d_i^2)/(2n)}$, where $d_i$ is the difference between the paired measurements and $n$ the number of differences (Dahlberg, 1940). The $S_{\text{method}}$ was also expressed as the coefficient of variation (COV in %), that is $S_{\text{method}}$/mean of walking distance*100. The data were calculated and analysed in Statistica version 12 (StatSoft®), MedCalc® version 14.10 (MedCalc Software) and Excel 2013 (Microsoft Office).

3 | RESULTS

Baseline characteristics of the patients are presented in Table 1.

The study group comprised 46 patients, including 9 females, with stable HF, EF (29 ± 9%) and an age ≤ 80 years (68.2 ± 8.7 years). Twenty-eight (61%) patients were in NYHA II and 18 (39%) in NYHA
Patients were recruited from the outpatient cardiology clinic or hospital register of patients undergoing echocardiography at the department of clinical physiology, both clinics at Kalmar County Hospital, Sweden.

All 6MWTs were performed safely without complications, and no test was prematurely discontinued. Of 230 theoretically duplicated 6MWTs (in total 460 walk tests), 35 (76%) patients performed all five follow-ups. Forty-six patients completed at least one, and a total of 198 paired tests (396 single walk tests) were performed. The mean distance, walked in m, for the first (6MWT1) versus second (6MWT2), for each follow-up, was 408 ± 100 versus 411 ± 96, 449 ± 94 versus 465 ± 94, 464 ± 96 versus 473 ± 100, 462 ± 10 versus 468 ± 104 and 472 ± 105 versus 482 ± 107.

The increase in walked distance at each follow-up may result from the exercise intervention in the studies (of which one is published; Lans et al., 2018) from where the data were collected.

The ICC was >0.90 at every follow-up (0.98, 0.96, 0.97, 0.99 and 0.98), with an overall ICC of 0.97 (CI 95% 0.94–0.99). Table 2. All performed 6MWTs are shown in Figure 1.

The Bland–Altman analysis, Figure 1, showed a mean test–retest positive bias of 9 m, that is on average a longer walked distance at the second test (95% limits of agreement −29.4 to 47.3 m) calculated on all performed follow-ups.

Reliability, calculated as the error in a single determination ($S_{\text{method}}$) and (ICC), did not significantly differ between the different

### TABLE 1  Patient baseline characteristics

<table>
<thead>
<tr>
<th>Variables</th>
<th>N = 46</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/Female (n)</td>
<td>37/9</td>
</tr>
<tr>
<td>Age (year)</td>
<td>68.2 ± 8.7</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>174 ± 9</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>83 ± 18</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27 ± 5</td>
</tr>
<tr>
<td>NYHA II</td>
<td>28 (61%)</td>
</tr>
<tr>
<td>III</td>
<td>18 (39%)</td>
</tr>
<tr>
<td>Ejection fraction %</td>
<td>29 ± 9</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
</tr>
<tr>
<td>Dilated cardiomyopathy</td>
<td>12 (26%)</td>
</tr>
<tr>
<td>Ischaemic heart disease</td>
<td>28 (61%)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>2 (4%)</td>
</tr>
<tr>
<td>Medication</td>
<td></td>
</tr>
<tr>
<td>ACE inhibitor</td>
<td>37 (80%)</td>
</tr>
<tr>
<td>ARB</td>
<td>5 (11%)</td>
</tr>
<tr>
<td>Diuretic</td>
<td>44 (96%)</td>
</tr>
<tr>
<td>β-blockers</td>
<td>36 (78%)</td>
</tr>
</tbody>
</table>

Abbreviations: ACE, angiotensin-converting enzyme; ARB, angiotensin II receptor blocker; BMI, body mass index; NYHA, New York Heart Association.

### TABLE 2  Repeatability of 6MWT

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>3 months</th>
<th>6 months</th>
<th>9 months</th>
<th>12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>46</td>
<td>42</td>
<td>39</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>First walk (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>408 ± 100</td>
<td>449 ± 94</td>
<td>464 ± 96</td>
<td>462 ± 103</td>
<td>472 ± 105</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>431 (95 to 555)</td>
<td>451 (260 to 616)</td>
<td>455 (260 to 634)</td>
<td>448 (240 to 633)</td>
<td>456 (256 to 668)</td>
</tr>
<tr>
<td>Second walk (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>411 ± 96</td>
<td>465 ± 94</td>
<td>473 ± 100</td>
<td>468 ± 104</td>
<td>482 ± 107</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>431 (138 to 570)</td>
<td>461 (260 to 646)</td>
<td>449 (258 to 643)</td>
<td>447 (244 to 623)</td>
<td>454 (250 to 665)</td>
</tr>
<tr>
<td>Diff 2-1 (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.4 ± 21.7</td>
<td>15.7 ± 19.6</td>
<td>9.7 ± 22.2</td>
<td>5.7 ± 14.7</td>
<td>10.5 ± 15.6</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>5.8 (~58 to 61)</td>
<td>12.5 (~18 to 68)</td>
<td>5.0 (~46 to 89)</td>
<td>5.0 (~48 to 28)</td>
<td>10.0 (~13 to 55)</td>
</tr>
<tr>
<td>Diff %</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.9 ± 8.9</td>
<td>3.8 ± 5.5</td>
<td>2.1 ± 5.4</td>
<td>1.3 ± 3.4</td>
<td>2.3 ± 3.6</td>
</tr>
<tr>
<td>Median (Range)</td>
<td>1.6 (~21 to 45)</td>
<td>2.4 (~5 to 23)</td>
<td>1.2 (~11 to 25)</td>
<td>1.3 (~12 to 7)</td>
<td>2.3 (~3 to 11)</td>
</tr>
<tr>
<td>p Valuea</td>
<td>0.288</td>
<td>&lt;0.001</td>
<td>0.010</td>
<td>0.027</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_{\text{method}}$ (COV%)</td>
<td>15.4 (3.8%)</td>
<td>17.6 (3.9%)</td>
<td>16.9 (3.6%)</td>
<td>11.0 (2.4%)</td>
<td>12.9 (2.7%)</td>
</tr>
<tr>
<td>ICC (CI 95%)</td>
<td>0.98 (0.96 to 0.99)</td>
<td>0.96 (0.94 to 0.98)</td>
<td>0.97 (0.94 to 0.98)</td>
<td>0.99 (0.98 to 0.99)</td>
<td>0.98 (0.97 to 0.99)</td>
</tr>
</tbody>
</table>

Note: 6MWT: 6-min walk test; $S_{\text{method}}$: the standard error of a single determination also expressed as coefficient of variance in % (COV%); and ICC (CI 95%) is intraclass correlation coefficient with 95% confidence interval.

aSignificant level of the difference between first and second walks.
There was no statistically significant difference between 6MWT1 and 6MWT2 in the rating of perceived exertion (RPE 6–20), dyspnoea (CR-10) or heart rate pre- and post-testing at any of the follow-ups.

4 | DISCUSSION

This study aimed to determine the test–retest reliability of the 6MWT. We found an excellent (Demers et al., 2001) reliability at duplicate measurements during the follow-up of the 6MWT in patients with HF, and the 6MWT is therefore suitable for repeated assessments of walked distances.

Encouragement is common and recommended but not imperative, there is no consensus regarding its use and further research is needed to validate its impact on the walked distance of the 6MWT (ATS statement, 2002; Guyatt et al., 1984; Holland et al., 2014; Singh et al., 2014). In addition, Singh et al. (2014) pointed out that no articles have reported effects of encouragement since 2000.

The exercise capability of patients with HF depends on several factors, both pathophysiological and psychological factors, such as motivation, experience of exercising and daily life activity. In clinical practice, there is a need for a repeatable test capable of assessing the patient’s baseline exercise capacity, and detecting any change over time, or as an effect of exercise training or other treatments (Eiser, Willscher, & Dore, 2003).

It is important that the test is not too physically demanding, as this can affect willingness to perform a follow-up test with the previous test in mind (Lainchbury & Richards, 2002). Guyatt et al. (1985) were the first to study, and show the simplicity of, the 6MWT in patients with HF. Later on, Guazzi et al. (2009) found the 6MWT a valid exercise test modality to assess clinical status and functional limitation objectively, and that it can represent a more immediately available and favourable application than CPET. The recommendations of the ATS (ATS statement, 2002) favour a practice test, stating that it is not absolutely necessary but should be considered, at least at the first performed 6MWT in a clinical setting. It has been shown that a walked distance ≤ 300 m indicates a poorer prognosis (Cahalin et al., 1996; Roul, Germain, & Bareiss, 1998). According to that, a 6MWT can be used to easily identify impairments in a patient with HF, and it seems to be enough to perform one 6MWT.

To further explore the need for multiple tests, the patients in the present study walked twice with a 45-min, seated rest between tests. At baseline, there was no statistically significant difference between the repeated 6MWTS. At the follow-up at 3, 6, 9 and 12 months, the walking distance was statistically, but not clinically, significantly longer at the second test, as it constituted only 2%–4% difference in the covered distance. The correlation between the covered distance at Test 1 and Test 2 in this study was very high.

The Bland–Altman analysis showed a positive test–retest bias of only 9 m with a relatively narrow range, which indicates that there was only a minor, clinically insignificant learning effect between test and retest. The range in this study was smaller compared to other studies using a shuttle walk test (Pepera, McAllister, & Sandercock, 2010) or 6MWT in patients with cardiac vascular disease (Bellet et al., 2011). Our data support that in the 6MWT, the
patients with HF will choose the same self-selected walking pace at repeated tests, based on their physical ability, and this indicates that only a single 6MWT is necessary in a follow-up situation.

Several studies have investigated the optimal number of performed 6MWTS in HF patients (Cahalin et al., 1996; Hanson et al., 2012; Kervio et al., 2004; O’Keeffe et al., 1998). However, measurements in those studies were performed at fewer follow-ups than in this study. In a study by Adsett et al. (2011), patients with HF performed two 6MWTS on the same day but with different rest intervals between tests, 15–90 min. They discovered that the difference between Test 1 and Test 2 was dependent on the baseline test (Test 1) performance. In patients who walked less than 300 m, the mean improvement was only 1.8 m, in patients walking 300–450 m the mean improvement was 16.1 m, and for the patients covering the longest distance, the mean improvement was 36.1 m. So, when patients covered a greater distance, the difference between the tests was greater. This was not confirmed in our study.

The instructions for the 6MWT, according to ATS (ATS statement, 2002), are not really consistent with the instructions given to patients with HF since the instructions are designed for patients with chronic obstructive pulmonary disease.

For example, ATS recommends the use of Borg’s CR-10 scale to estimate the exertion rate, while Borg’s RPE scale is commonly used in cardiac rehabilitation. Since instructions to the 6MWT differ widely, it may be difficult to compare studies. Thus, there seems to be a need to standardize the 6MWT protocol for patients with HF.

4.1 | Limitations

A limitation of our study is that only about one fifth of the participants were women. It is a common finding that fewer women participate in exercise studies for patients with HF. Only patients having heart failure with reduced ejection fraction were studied, which makes it difficult to generalize to other populations with HF.

5 | CONCLUSION

This study has shown an excellent test–retest reliability for the 6MWT in patients with HF, also on repeated occasions during the long-term follow-up of one year. Thus, there is apparently no learning effect in the 6MWT over a long-term test–retest duration, which rejects the need for a practice walk.

CONFLICT OF INTEREST

All authors declare no conflicts of interest. All authors have read the final version of the manuscript and approved submission to the Clinical Physiology and Functional Imaging.

REFERENCES


Reilly, T., & Tipton, M. (2010). A sub-maximal occupational aerobic fitness test alternative, when the use of heart rate is not appropriate. Work, 36, 333–337. https://doi.org/10.3233/WOR-2010-1034