R N J | 1 1 1 | B | Dispatch: 23.4.13 | Journal: CE:
Journal Name | Manuscript No. | Author Received: | No. of pages: 10 | PE: Bhuvana M

FEATURE

Test–Retest Reliability and Minimal Detectable Change Scores for Fitness Assessment in Older Adults with Type 2 Diabetes

Rosa M. Alfonso-Rosa[1], Msc, Borja del Pozo-Cruz[1], PhD, Jesus del Pozo-Cruz[1], Msc, Borja Sañudo Corrales[1], PhD & Michael E. Rogers[2], PhD

1 Department of Physical Education and Sports, University of Seville, Seville, Spain

2 Department of Human Performance Studies, Wichita State University, Wichita, KS, USA

Keywords

Reliability; diabetic patients; physical function; elderly.

Correspondence

Borja del Pozo-Cruz, Department of Physical Education and Sports, University of Seville, C/Pirotecnia s/n, Seville 41012, Spain.

E-mail: bpozo@us.es

Accepted April 01, 2013.

doi: 10.1002/rnj.111

Abstract

Purpose: To assess the intraclass correlation coefficients (ICCs) and to determine the minimal detectable change (MDC_{95}) scores of the data for the Hand Grip Strength Test, the Chair Sit and Reach Test (CSRT), the Timed "Up and Go" (TUG) test, the Six-Minute Walk Test (6MWT) and 30 seconds Sit to Stand Test (30s-STS) test in older adults with type 2 NIDDM.

Design: Test–retest reliability.

Methods: Eighteen subject participated in two sessions (1 week apart), which included the different tests.

Findings: High ICCs (≥ 0.92) were found for all tests. The MDC_{95} scores were as follows: 4.0 kg for Hand Grip Strength Tests, 7.5 cm for the right leg-CSRT, 9.0 cm for the left leg-CSRT, 1.0 second for the TUG test, 27 m for the 6MWT, and 3.3 repetitions for the 30s-STS test.

Conclusions: All tests evaluated are reliable outcome measures for type 2 NIDDM patients.

Clinical relevance: This study has generated novel MCD_{95} data, which will assist nursing practitioners in both prescribing the most beneficial exercise and interpreting posttreatment changes after rehabilitation in patients with T2DM.

Background

Type 2 diabetes mellitus (T2DM) is a metabolic disorder characterized by hyperglycemia and insufficiency of secretion or action of endogenous insulin (Maritim, Sanders, & Watkins, 2003). Because T2DM prevalence is increasing within worldwide (Zimmet, Alberti, & Shaw, 2001), public health authorities should encourage the implementation of both clinical and preventive intervention programs to tackle the associated health and economic burden of the disease. Increasing physical fitness in this population could enhance insulin sensitivity and glycemic control (Srikanthan & Karlamangla, 2011) and may attenuate declines in strength, endurance, and function (Balducci et al., 2012), all of which have been reported to be lower in this population when compared with those without the disease (Ozdirenc, Biberoglu, & Ozcan, 2003; Th et al., 2012). On the other hand, low levels of fitness have been considered to independently predict mortality among older adults (Sui et al., 2007) as well as those with T2DM (Nylen et al., 2010). Therefore, interventions designed to improve fitness levels are warranted for this population (Knowler et al., 2002; Tuomilehto et al., 2001).

Considering the importance of physical fitness in people with T2DM, nurses working with older adults with T2DM should be cognizant of the fitness level and abilities of individual patients and encourage appropriate exercise. Moreover, it is important to assess progress or decline in function for an optimal prescription in each individual case. Therefore, specific clinical tools should be tested for reliability with individuals with T2DM. A variety of physical fitness tests are currently available to assess function and to monitor improvements in a clinical setting.

In this study, the Hand Grip Strength Test, Chair Sit and Reach Test (CSRT), Timed "Up and Go" (TUG), Six-Minute Walk Test (6MWT), and 30 seconds Sit to Stand Test (30s-STS Test) were used. With the exception of the CSRT, the Hand Grip Strength Test (Gin et al., 2010), the TUG (Alvarenga, Pereira, & Anjos, 2010; Oliveira, Fachin, Tozatti, Ferreira, & Marinheiro, 2012), the 6MWT (Ozdirenc et al., 2003), and the 30s-STS Test (Lambers et al., 2008) have been previously used among patients with T2DM.

The Hand Grip Strength Test is often used to quantitatively assess maximal voluntary isometric muscle strength of the arms. It has been suggested as a tool to characterize the severity of a specific disease (Aparicio et al., 2011) and disease diagnoses (Aparicio et al., 2011). The CSRT has also been used in different older adult populations (Bautmans, Van Hees, Lemper, & Mets, 2005; Schmid, Van Puymbroeck, & Koceja, 2010) for lower limb flexibility (which is also reduced in T2DM (Herriott, Colberg, Parson, Nunnold, & Vinik, 2004)) assessment purposes. The TUG test is the most widely used test for mobility/agility assessment in clinical populations (e.g., postpolio survivors or Parkinson's disease) (Lehmann, Sunnerhagen, & Willen, 2006; Matinolli et al., 2009). The 6MWT is a physical fitness test that is used to assess cardiorespiratory fitness, which is also reduced in patients with T2DM (Ozdirenc et al., 2003). The 10-repeated Sit to Stand (STS) Test (Takai et al., 2009) and the 5-repeated STS Test (Batista et al., 2012) have been used to quantitate lower limb muscle strength in patients with lower limb muscle strength weakness, a characteristic of patients with T2DM (Th et al., 2012). However, as both tests can suffer from floor effects (i.e., no score awarded unless the subject can complete the required number of stands), the number of full stands completed in 30 seconds (30s-STS Test) has been the preferred method to evaluate lower limb muscle strength in older adults (Rikli, 2001).

Unfortunately, the reliability of these tests for people with T2DM has not been yet determined. The reliability of a test should be expressed as both relative reliability and absolute reliability. Relative reliability may be measured with the intraclass correlation coefficient (ICC), which is used for test–retest reliability (Weir, 2005) or the coefficient of variation, expressed as a percentage (% CV) (Cohen, 1988). Individual performance and measurement error are measured with absolute reliability, which provides information for differentiating a true change in performance from a change due to individual variation and measurement error (Weir, 2005).

Within this context, the aims of this study were to calculate the test–retest relative reliability of commonly used physical fitness tests in older adults with T2DM (Hand Grip Strength Test, the CSRT, the TUG Test, the 6MWT, and the 30s-STS Test) and to calculate the absolute reliability with the standard error of measurement (*SEM*) and minimal detectable change scores at 95% confidence intervals (MDC_{95}).

Research design and methods

Participants and study design

A test–retest reliability study design was conducted. Participants were recruited (between January 1 and March 30, 2012) from a local primary care facility (Seville, Spain). Twenty-five volunteers received detailed information about the aims and study procedures and were included in the study. Inclusion criteria were T2DM diagnoses (Diagnosis & classification of diabetes mellitus, 2012) and to be able to walk independently without pain or a walking aid. Exclusion criteria were less than 65 years of age, a T2DM-related complication (i.e., neuropathy, nephropathy, or vision impairment), uncontrolled diabetes, history of cognitive impairment, severe heart and liver or kidney disease. A total of seven participants did not meet these criteria and were not included in the study. Finally, 18 patients (aged 73.6 [8.1] years) were included.

The study was developed following the ethical guidelines of the Declaration of Helsinki, last modified in 2000 and had local research and ethics committee approval (University of Seville). All subjects gave written consent.

Procedures and outcome measures

Participant characteristics were recorded, including age, gender, annual income, marital status, educational status, blood pressure, resting heart rate, number of oral glycemic medications, fasting blood glucose, hemoglobin A1C, and diabetes duration. Their weight, height, and waist and hip circumferences were measured, so that body mass index (BMI; kg/m^2) and waist-to-hip ratio could be calculated. Body fat percentage (BF%) was also estimated using a handheld impedance analyzer (Omron BF-306; Omron Healthcare Europe BV, Hoofddorp, The Netherlands) according to the manufacturer's instructions (Deurenberg et al., 2001). Mid-arm and mid-calf circumferences were also measured.

Participants performed each of the tests twice, with a 1-week interval between the testing sessions. Every effort was made to keep all factors associated with the testing sessions consistent: day of the week, time of day, climatic characteristics, and area in which the test was performed. Participants performed the Hand Grip Strength Test, the CSRT (Right and Left legs), the TUG Test, the 6MWT, and the 30s-STS Test. Participants were required to rest for 5 minutes between each mode of testing in an effort to allow recovery (Gusi et al., 2011).

Hand Grip Strength Test was used to assess upper body muscular strength (Rodriguez et al., 1998). This test was conducted with a digital dynamometer (TKK 5401 Grip-D; Takei Scientific Instruments, Tokyo, Japan). The participants maintained the standard bipedal position during the entire test with the arm in complete extension. Each participant performed the test twice with each hand allowing a 1-minute rest period between measures. The best value of two trials was chosen as score of the test for each arm (dominant and nondominant arm) and an average score of both hands was computed as bimanual Hand Grip Score. The grip position of the dynamometer was adjusted to each individual's hand size.

The CSRT was used to assess lower body flexibility (Rikli, 2001). A ruler was used to measure the distance between the end of the 3rd digit of the hand and the toes. This value is negative if the fingertip does not reach the toes and positive if fingertip passes the toes. Both sides were measured twice and the maximal score from each leg was recorded.

Motor agility/mobility was assessed by the TUG test (Rikli, 2001). The participant had to stand up from a chair, walk 2.44 m to and around a cone, and return to the chair in the shortest possible time. The best time of two trials (1-minute rest period between each trial) was recorded.

To assess cardiovascular fitness, the 6MWT was used (Rikli, 2001). Participants were instructed to walk as far as they could at a fast, comfortable pace in 6 minutes. The maximum distance (meters) walked was recorded as the score of the test. Participants were discouraged from talking during the test and were notified of each passing minute.

The 30 seconds Sit to Stand Test was used to assess lower body strength (Rikli, 2001). Participants were instructed to perform the task starting and finishing in the seated position. Participants were allowed a practice trial before the beginning of the test. The number of times within 30 seconds that the participant could raise to a full stand from a seated position with back straight and feet flat on the floor "as fast as possible," without pushing off the arms was counted.

Sample size and study power

Before the start of the study, required sample size was calculated following the suggestions of Walter et al. (Walter, Eliasziw, & Donner, 1998) to reach a power of 0.90 on ICC, according to the following standards: alpha = 0.05, under the null hypothesis that the ICC was moderate in accordance with the criteria points used (0.50) (Munro, Visintainer, & Page, 1986), and the alternative hypothesis had excellent ICC (0.9) (Munro et al., 1986). Within these criteria, the required sample size was at least 11 participants for each test; however, additional participants were recruited because of the potential for attrition between the two test trials. The final reported power achieved in each test with these specifications was 0.96 for Hand Grip Strength, CSRT, and TUG Tests; 0.93 for the 6MWT and 0.95 the 30s-STS test.

Data analysis

The SPSS package version 18.0 for Windows (SPSS Inc., Chicago, IL) was used for data analysis. The level of significance was set at $p \leq .05$ for all statistical analyses performed. According to the Kolmogorov–Smirnov test, data were normally distributed for all measures in this study, and therefore parametric statistics were used. Data are presented as means (*SD*), unless otherwise stated. Paired sample *t*-tests were performed to analyze differences between tests and retest sessions for all physical fitness assessments.

The same technician (who had more than 3 years of experience applying the tests used in this study in T2DM patients) administered all of the tests, so the intrarater reliability was calculated. Relative reliability was determinate using the $ICC_{1,1}$ (one-way random effects model analysis of variance) with 95% confidence intervals and across the two test sessions (Shrout & Fleiss, 1979). An ICC above 0.70 was considered to demonstrate good reliability (Munro et al., 1986); although for clinical measures, it has been suggested that the ICC should exceed 0.90 (Portney & Watkins, 2000). In addition, the %CV based on the method error to quantitate the percentage of variation from trial to trial was calculated for a better understanding on reliability. For the CV, a change of 10%–20% in variation was considered to be adequate reliability (Cohen, 1988).

Absolute reliability was determined with the standard errors of measurement (*SEM*) and minimal detectable change scores at 95% confidence interval (MCD_{95}) with the following equations (Weir, 2005): $SEM = SD \sqrt{(1 - ICC)}$. In this equation, *SD* is the mean *SD* of day 1 and day 2, and ICC is the reliability coefficient. $MCD_{95} = 1.96 \sqrt{2SEM}$. In this equation, SEM was calculated as previously described. The 1.96 in the MDC_{95} equation represents the z-score at the 95% confidence level. Bland–Altman plots were also performed for physical fitness tests (Bland & Altman, 1986).

Results

Two participants were unwilling to attend the second testing session. In addition, three participants did not complete the second 6MWT due to the lack of time and one participant could not perform the second 30s-STS Test because of worsened health. No adverse events occurred during testing.

Descriptive statistics for the 18 participants are shown in Table 1. Outcomes of day 1 and day 2 testing values and the results of repeated tests are shown in Table 2. No statistically significant differences were found between testing days for all outcomes of the study except for hand grip measures, where the kg values of day 2 were greater than kg values of day 1 ($p < 0.05$).

The ICCs of each test are presented in Table 3, along with the *SEM* and MCD_{95} values. The ICCs for test–retest reliability were high (>0.90) for all of the outcome measures (Hand Grip Strength Test, the CSRT, the TUG Test, the 6MWT, and the 30s-STS). Also, the %CV for each test across the trials is presented in Table 3. With the exception of CSRTs, the %CVs (between 5% and 17%) shown in this study indicate that participants had low variability when both trials were considered.

Figure 1 shows Bland-Altman plots of the tests on day 1 and day 2. With the exception of the TUG Test and Right CSRT, the bias representing the average difference for measures between day 1 and day 2 was negative. This information indicates that day 2 had higher values than day 1.

The *SEM* was 1.40 kg, 1.56 kg, and 1.49 kg for the Hand Grip Strength Test (dominant arm, nondominant arm, and bimanual hand grip strength, respectively); 2.70 and 3.25 cm for the CSRT (Right and left CSRT, respectively); 0.31 seconds for the TUG Test; 9.88 m for the 6MWT, and 1.21 repetitions for the 30s-STS Test.

The MDC_{95} values were 3.85 kg, 4.32 kg, and 4.13 kg for the Hand Grip Strength Test (dominant arm, nondominant arm, and bimanual hand grip strength, respectively); 7.50 and 9.01 cm for the CSRT (Right and Left CSRT, respectively); 0.85 sec for the TUG Test; 27.37 m for the 6MWT; and 3.35 repetitions for the 30s-STS Test.

Table 1 Characteristics of the type 2 diabetic older adults included in the study (n = 18)

Socio-Demographic Characteristics	
Age (years)*	73.55 (8.13)
Gender, women (%)	44.4
Income (%)	
<USD 1,544	72.2
USD 1,544–2,315	22.2
>USD 2,315	5.6
Educational status (%)	
Unfinished studies	22.2
Primary school	38.9
Secondary school	33.3
University degree	5.6
Marital status (%)	
Married	50
Unmarried	11.1
Separated/Divorced/Widowed	38.9
Clinic and health characteristics	
Oral hypoglycemic agents (number per day)*	1.9 (1.5)
Blood glucose level (mg/dL)*	141.4 (43)
HbA_{1c} (%)	7.2 (1.2)
SBP (mmHg)*	15.06(2.2)
DBP (mmHg)*	6.50 (1)
HR (bpm)*	75.7 (11.3)
Years since clinical diagnosis (number)*	8.7 (7.3)
Body composition	
BMI (Kg/m^2)*	31.5 (7.3)
BF%	38.4 (10.5)
WHR*	0.91 (0.09)

*Values expressed as Mean (*SD*), HbA1c: Glycated hemoglobin. SBP, systolic blood pressure; DBP, diastolic blood pressure; HR, heart rate; BMI, body mass index; BF%, body fat percentage; WHR, waist-to-hip ratio.

Discussion

As a novelty, this study has provided estimates of variability for commonly used physical performance measures

Table 2 Mean differences in physical fitness tests performed by older adults with T2DM between day 1 and day 2 of measurement

Physical Fitness Variables	Day 1 Mean (*SD*)	Day 2 Mean (*SD*)	*t*	*p*
Hand Grip Strength, dominant arm (kg) (*n* = 16)	25.56 (9.83)	27.90 (11.08)	−2.43	.028
Hand Grip Strength, nondominant arm (kg) (*n* = 16)	23.45 (10.42)	26.06 (11.60)	−2.98	.009
Bimanual Grip Strength (kg) (*n* = 16)	24.50 (9.96)	26.98 (11.12)	−2.98	.009
Right Chair Sit and Reach Test (cm) (*n* = 16)	−10.94 (10.64)	−13.68 (11.44)	.932	.366
Left Chair Sit and Reach Test (cm) (*n* = 16)	−11.69 (12.13)	− 12.90 (12.45)	−.470	.645
Timed "Up and Go" Test (s) (n = 16)	8.77 (2.21)	8.75 (2.15)	.675	.510
Six-Minute Walk Test (m) (*n* = 13)	391.14 (97.37)	391.69 (100.15)	−1.757	.101
30-Sit to Stand Test (number of times) (*n* = 15)	12.11 (3.58)	12.93 (4.96)	−1.389	.190

T2DM: Type 2 diabetes mellitus; Day 1: day 1 of measurement; Day 2: Day 2 of measurement; *p*: *p* value from *Student's t* for repeated measures.

Table 3 Reliability analysis of the physical fitness tests performed in T2DM older adults

Physical Fitness Tests	ICC	95% CI of the ICC	*SEM*	%*SEM*	MDC	**%** MDC	%CV
Hand Grip Strength, dominant arm (kg)	.98	(.95 to .99)	1.40	5.2	3.89	14.5	10.62
Hand Grip Strength, nondominant arm (kg)	.98	(.96 to .99)	1.56	6.3	4.32	17.4	10.52
Bimanual Grip Strength (kg)	.98	(.96 to 1.00)	1.49	5.8	4.13	16.1	9.55
Right Chair Sit and Reach Test (cm)	.94	(.84 to .98)	2.70	22.0	7.50	60.9	39.22
Left Chair Sit and Reach Test (cm)	.93	(.82 to .97)	3.25	26.4	9.01	73.3	47.56
Time "Up and Go" Test (s)	.98	(.95 to .99)	0.31	3.5	0.85	9.8	6.46
Six-Minute Walk test (m)	.99	(.96 to 1.00)	9.88	2.5	27.37	7.0	5.12
30-Sit to Stand Test (number of times)	.92	(.79 to .98)	1.21	9.6	3.35	26.7	17.60

T2DM, Type 2 diabetes mellitus; ICC, intraclass correlation coefficient; CI, confidence interval; *SEM*, standard error of measurement; MDC, minimal detectable change; %CV, % coefficient of variation.

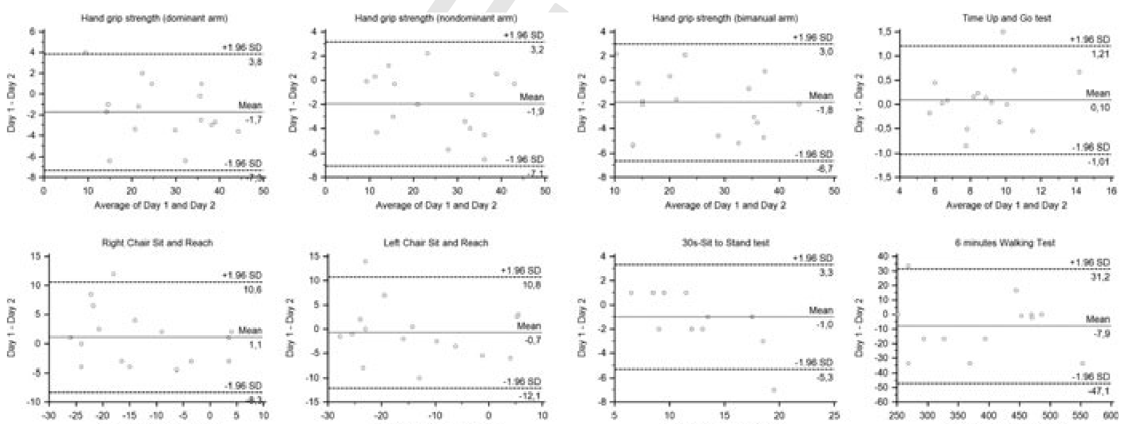

Figure 1 Bland–Alman Plots with limits of agreement of physical fitness tests.

for a group of older adults with T2DM. The main findings of this study were that the test–retest relative reliability of the tests was excellent. All outcome measures were found to have high test–retest relative reliability (ICCs) having values above the threshold of 0.90 for minimal acceptable reliability for a clinical test (Portney & Watkins, 2000).

Reported ICC values for the Hand Grip Strength Test reported in this study were similar to those reported in a clinical population (Segura-Orti & Martinez-Olmos, 2011), healthy adults (Hamilton, Balnave, & Adams, 1994), and older adults (Gusi et al., 2011). Although there is a paucity of data on the use of the CSRT, reported ICCs for this test in healthy older adults are consistent with the results of this study (Gusi et al., 2011), thus confirming that the test has good relative reliability in patients with T2DM.

Along with the 6MWT, the TUG Test is one of the most commonly used tests for functional capacity assessment on chronic diseases (Rasekaba, Lee, Naughton, Williams, & Holland, 2009). In this study, the reported ICC for the test–retest reliability of the 6MWT was 0.99. Although this test has also been used in adults with T2DM (Ozdirenc et al., 2003), no previous study, to the authors' knowledge, has determined the ICC of this test in people with T2DM. In any case, results from this study are consistent with those previously presented in studies of older adults with other special clinical situations (Lin & Bose, 2008; Ries, Echternach, Nof, & Gagnon Blodgett, 2009) as well as healthy older adults (Gusi et al., 2011). Comparable ICCs with those found in dependent older adults (Nordin, Rosendahl, & Lundin-Olsson, 2006) and healthy older adults (Gusi et al., 2011) were found for the TUG Test in this study.

The 30s-STS Test has not been used in people with T2DM; however, the ICCs of this test in older adults populations (Gusi et al., 2011) and people with musculoskeletal problems (Smeets, Hijdra, Kester, Hitters, & Knottnerus, 2006) have been reported to be similar to ICC results from this study.

Factors that may explain the high ICCs in all of the physical fitness tests are consistent timing of the tests (same day and hour of the week) and standardization of the evaluator's instructions. Despite this, using only the ICC can give a false impression about the reliability of a measurement. Bland–Altman analysis can confirm a good reliability. As results from Bland–Alman showed that systematic errors (mean difference between test–retest) for the physical fitness tests assessed in this study were nearly zero and the 95% limits of agreement were narrow, the good reliability of the measurement can be confirmed.

This study also used the method error as an adjunct to test–retest reliability because it reflected the percentage of variation from trial to trial, which is not given by the ICC. For example, a CV of 5% for the 6MWT indicates that walking 300 m in one trial might produce an expected variability of 15 m in the next trial. This knowledge could be applied for the rest of variables of this study.

However, ICC, Bland–Alman, or %CV values are not enough to interpret data from a clinical point of view. Therefore, measurement errors should also be small and the method sufficiently sensitive to detect real changes. In this case, the *SEM* (%*SEM*) was used for these purposes. With the exception of the CSRT, the %*SEM* values in this study were low, between 3.5% and 10%, acceptable values from a clinical point of view, indicating that measurements of these tests can be made reliable for a group of older adults with T2DM. Consequently, intervention studies with improvements less than approximately 10%, in most cases, do not indicate a real change.

This study indicated that the Hand Grip Strength Test, the CSRT, the TUG Test, the 6MWT, and the 30s-STS Test are reliable measures. Health-care providers are encouraged to understand how changes in scores translate to clinical practice. To detect a real change for a single individual, MCD_{95} (independent of the unit of measurement and) was calculated (Beckerman et al., 2001). On the basis of the MCD_{95} observed in this study, if a change exceeding ± 0.85 seconds occurs in the TUG Test; ± 27.37 m in the 6MWT; or ± 3.35 repetitions in the 30s-STS Test, clinicians can be 95% confident that the difference is not due to measurement error or variability among participants. Similar conclusions could be achieved for all of the tests being evaluated (Hand Grip Strength Test and the CSRT).

Limitations

There are several limitations to this study. The sample was one of convenience, and although it represented a wide range in physical function and T2DM clinical characteristics, all of the participants were from a primary care center situated in an urban area and may not be representative of all older adults with T2DM. Because the testing protocol of the performed measures may affect reliability, great care was taken to standardize the tests and to carefully follow the protocol. However, all tests were conducted on the same day and, as a result, some people could not attend or did not have time to complete all the tests on that day. It is therefore advisable that future studies conduct more test days to evaluate test–retest reliability to maximize participation. To minimize examiner-related variability, the same examiner performed all measurements and gave standardized instructions with all measures. Although the logistics of using a single researcher for data collection can influence sample size, this was in accordance with exigencies to achieve a high (.90) power. A 7-day interval between tests was used to avoid any influence of learning, fatigue, or pain on the second application of the test. However, participants in the study achieved greater values in hand grip strength test during day 1 of testing, perhaps representing learning effects for this test. Another limitation was that the testing sessions were not conducted in private, so the influence of other older adults in the room may have affected performance on any particular day. The measurements are acceptably sensitive for groups of patients, but to monitor individual progress reliability, studies by age group (e.g., 65–75 years and 75 years or more) should be conducted. Therefore, a large sample size is required to confirm the results achieved in this study and to determine the reliability of these outcome measures in different age groups of older adults with T2DM.

Conclusions

This study suggests excellent test–retest reliability for the Hand Grip Strength Test, the CSRT, the TUG Test, the 6MWT, and the 30s-STS Test in older adults with well-controlled noninsulin dependent T2DM. Despite very high ICCs for test–retest reliability, there was remarkable individual variability in the performance of some of these tests (i.e., CSRT). Presentation of *SEM* and MDC_{95}, for each of the measurement tools, provides nursing practitioners with meaningful thresholds for identifying changes beyond those expected from measurement error and individual variability (i.e., "true" change) in individuals with T2DM. These findings are applicable for both clinicians and researchers. This study has generated novel MDC_{95} values for the Hand Grip Strength Test, the CSRT Test, the TUG Test, the 6MWT, and the 30s-STS Test that will be helpful in monitoring performance changes over time and assessing the effectiveness of physical therapy and exercise interventions in older adults with T2DM.

Acknowledgments

Authors thank A.C, S.C, A. F, and M.H for their technical contribution to the study. Also we thank the staff from the primary care center "Los Bermejales" (Seville, Spain) for their contribution in the recruitment process of patients included in the study and for providing the facility for testing.

Key Practice Points

- Increasing physical fitness in type 2 diabetes could enhance insulin sensitivity and glycemic control reducing declines in physical function associated with aging and accelerated by diabetes
- Nurses working with older adults need to measure outcomes to assess progress or decline in function, and therefore specific clinical tools should be tested for reliability with individuals with type 2 diabetes.
- On the basis of the MCD_{95} observed in this study, if a change exceeding ± 0.85 seconds occurs in the TUG Test; ±27.37 m in the 6MWT; 4 kg in hand grip; or ±3.35 repetitions in the 30s-STS Test, clinicians can be 95% confident that the difference is not due to measurement error or variability among participants.
- The results from this study suggest that common used physical fitness tests in older adults are also reliable among those with type 2 diabetes.

References

Alvarenga, P.P., Pereira, D.S., & Anjos, D.M. (2010). Functional mobility and executive function in elderly diabetics and non-diabetics. *Revista Brasileira de Fisioterapiar*, *14*(6), 491–496.

Aparicio, V. A., Carbonell-Baeza, A., Ruiz, J. R., Aranda, P., Tercedor, P., Delgado-Fernandez, M. et al. (2011). Fitness testing as a discriminative tool for the diagnosis and monitoring of fibromyalgia. *Scandinavian Journal of Medicine and Science in Sports*, ???, ???–???. 1

Aparicio, V.A., Ortega, F.B., Heredia, J.M., Carbonell-Baeza, A., Sjostrom, M., & Delgado-Fernandez, M. (2011). Handgrip strength test as a complementary tool in the assessment of fibromyalgia severity in women. *Archives of Physical Medicine and Rehabilitation*, *92*(1), 83–88.

Balducci, S., Zanuso, S. et al. (2012). Effect of high- versus low-intensity supervised aerobic and resistance training on modifiable cardiovascular risk factors in type 2 diabetes; the Italian Diabetes and Exercise Study (IDES). *PLoS One*, *7* (11), e49297. 2

Batista, F.S., Gomes, G.A., Neri, A.L., Guariento, M.E., Cintra, F.A., Sousa Mda, L. et al. (2012). Relationship between lower-limb muscle strength and frailty among elderly people. *Sao Paulo Medical Journal*, *130*(2), 102–108.

Bautmans, I., Van Hees, E., Lemper, J.C., & Mets, T. (2005). The feasibility of Whole Body Vibration in institutionalised elderly persons and its influence on muscle performance, balance and mobility: a randomised controlled trial [ISRCTN62535013]. *BMC Geriatrics*, *5*, 17.

Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., & Verbeek, A.L. (2001). Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research*, *10*(7), 571–578.

Bland, J.M., & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307–310.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. 2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Deurenberg, P., Andreoli, A., Borg, P., Kukkonen-Harjula, K., de Lorenzo, A., van Marken Lichtenbelt, W.D. et al. (2001). The validity of predicted body fat percentage from body mass index and from impedance in samples of five European populations. *European Journal of Clinical Nutrition*, *55*(11), 973–979.

agnosis and classification of diabetes mellitus. (2012). ???????. *Diabetes Care*, *35*(Suppl 1), S64–71.

n, H., Rigalleau, V. et al. (2010). Insulin therapy and body weight, body composition and muscular strength in patients with type 2 diabetes mellitus. *Journal of Nutrition and Metabolism*, ???: ???–???.

4

Gusi, N., Prieto, J., Olivares, P.R., Delgado, S., Quesada, F., & Cebrián, C. (2011). Normative Fitness Performance Scores of Community-Dwelling Older Adults in Spain. *Journal of Aging and Physical Activity*, *???*, ???–???.In press.

5

Hamilton, A., Balnave, R., & Adams, R. (1994). Grip strength testing reliability. *Journal of Hand Therapy*, *7*(3), 163–170.

Herriott, M.T., Colberg, S.R., Parson, H.K., Nunnold, T., & Vinik, A.I. (2004). Effects of 8 weeks of flexibility and resistance training in older adults with type 2 diabetes. *Diabetes Care*, *27*(12), 2988–2989.

Knowler, W.C., Barrett-Connor, E., Fowler, S.E., Hamman, R.F., Lachin, J.M., Walker, E.A. et al. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, *346*(6), 393–403.

Lambers, S., Van Laethem, C. et al. (2008). Influence of combined exercise training on indices of obesity, diabetes and cardiovascular risk in type 2 diabetes patients. *Clinical Rehabilitation*, *22*(6), 483–492.

Lehmann, K., Sunnerhagen, K.S., & Willen, C. (2006). Postural control in persons with late effects of polio. *Acta Neurologica Scandinavica*, *113*(1), 55–61.

Lin, S.J., & Bose, N.H. (2008). Six-minute walk test in persons with transtibial amputation. *Archives of Physical Medicine and Rehabilitation*, *89*(12), 2354–2359.

Maritim, A.C., Sanders, R.A., & Watkins, J.B., 3rd (2003). Diabetes, oxidative stress, and antioxidants: a review. *Journal of Biochemical and Molecular Toxicology*, *17*(1), 24–38.

Matinolli, M., Korpelainen, J.T., Korpelainen, R., Sotaniemi, K.A., Matinolli, V.M., & Myllyla, V.V. (2009). Mobility and balance in Parkinson's disease: a population-based study. *European Journal of Neurology*, *16*(1), 105–111.

Munro, B.H., Visintainer, M.A., & Page, E.B. (1986). Statistical Methods for Health Care Research. Philadelphia: JB Lippincott.

Nordin, E., Rosendahl, E., & Lundin-Olsson, L. (2006). Timed "Up & Go" test: reliability in older people dependent in activities of daily living–focus on cognitive state. *Physical Therapy*, *86*(5), 646–655.

Nylen, E.S., Kokkinos, P. et al. (2010). Prognostic effect of exercise capacity on mortality in older adults with diabetes mellitus. *Journal of the American Geriatrics Society*, *58*(10), 1850–1854.

Oliveira, P.P., Fachin, S.M., Tozatti, J., Ferreira, M.C., & Marinheiro, L.P. (2012). Comparative analysis of risk for falls in patients with and without type 2 diabetes mellitus. *Revista Da Associacao Medica Brasileira*, *58*(2), 234–239.

Ozdirenc, M., Biberoglu, S., & Ozcan, A. (2003). Evaluation of physical fitness in patients with Type 2 diabetes mellitus. *Diabetes Research and Clinical Practice*, *60*(3), 171–176.

Portney, L., & Watkins, M. (2000). Foundations of Linical Research: Applications to Practice. (2nd ed.): Upper Saddle River, NJ: Prentice Hall Health.

Rasekaba, T., Lee, A.L., Naughton, M.T., Williams, T.J., & Holland, A.E. (2009). The six-minute walk test: a useful metric for the cardiopulmonary patient. *Internal Medicine Journal*, *39*(8), 495–501.

Ries, J.D., Echternach, J.L., Nof, L., & Gagnon Blodgett, M. (2009). Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Physical Therapy*, *89*(6), 569–579.

Rikli, J. (2001). Senior Fitness Test Manual (Human Kinetics ed.). IL: Champaign.

Rodriguez, F., Gusi, N., Valenzuela, A., Nacher, S., Nogues, J., & Marina, M. (1998). Evaluation of health-related fitness in adults (I): background and protocols of the AFISAL-INEFC Battery [in Spanish]. *Apunts*, *52*, 54–76.

Schmid, A.A., Van Puymbroeck, M., & Koceja, D.M. (2010). Effect of a 12-week yoga intervention on fear of falling and balance in older adults: a pilot study. *Archives of Physical Medicine and Rehabilitation*, *91*(4), 576–583.

Segura-Orti, E., & Martinez-Olmos, F.J. (2011). Test-retest reliability and minimal detectable change scores for sit-to-stand-to-sit tests, the six-minute walk test, the one-leg heel-rise test, and handgrip strength in people undergoing hemodialysis. *Physical Therapy*, *91*(8), 1244–1252.

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Smeets, R.J., Hijdra, H.J., Kester, A.D., Hitters, M.W., & Knottnerus, J.A. (2006). The usability of six physical performance tasks in a rehabilitation population with chronic low back pain. *Clinical Rehabilitation*, *20*(11), 989–997.

Srikanthan, P., & Karlamangla, A.S. (2011). Relative muscle mass is inversely associated with insulin resistance and prediabetes. Findings from the third National Health and Nutrition Examination Survey. *Journal of Clinical Endocrinology and Metabolism*, *96*(9), 2898–2903.

Sui, X., LaMonte, M.J. et al. (2007). Cardiorespiratory fitness and adiposity as mortality predictors in older adults. *JAMA*, *298*(21), 2507–2516.

Takai, Y., Ohta, M., Akagi, R., Kanehisa, H., Kawakami, Y., & Fukunaga, T. (2009). Sit-to-stand test to evaluate knee extensor muscle size and strength in the elderly: a novel approach. *Journal of Physiological Anthropology*, *28*(3), 123–128.

Th, I.J., Schaper, N.C., Melai, T., Meijer, K., Willems, P.J., & Savelberg, H.H. (2012). Lower extremity muscle strength is reduced in people with type 2 diabetes, with and without polyneuropathy, and is associated with impaired mobility and reduced quality of life. *Diabetes Research and Clinical Practice*, *95*(3), 345–351.

Tuomilehto, J., Lindstrom, J., Eriksson, J.G., Valle, T.T., Hamalainen, H., Ilanne-Parikka, P. et al. (2001). Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New England Journal of Medicine*, *344*(18), 1343–1350.

Walter, S., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine*, *17*(1), 101–110.

Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research*, *19*(1), 231–240.

Zimmet, P., Alberti, K.G., & Shaw, J. (2001). Global and societal implications of the diabetes epidemic. *Nature*, *414* (6865), 782–787.

Author Query Form

Journal: RNJ

Article: 111

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: Please provide the volume number, page range for reference Aparicio et al. (2011).**	
2	**AUTHOR: If there are fewer than 7 authors for Reference Balducci et al., (2012) please supply all of their names. If there are 7 or more authors, please supply the first 6 author names then et al. Please check and update all such references found in the list.**	
3	**AUTHOR: Please provide the article title for reference Diagnosis and classification of diabetes mellitus (2012).**	
4	**AUTHOR: Please provide the volume number, page range for reference Gin and Rigalleau (2010).**	
5	**AUTHOR: Please update the volume number, page range for reference Gusi et al. (2011).**	
6	**AUTHOR: Figure 1 has been saved at a low resolution of 338 dpi. Please resupply at 600 dpi. Check required artwork specifications at http://authorservices.wiley.com/bauthor/illustration.asp**	

Proof Correction Marks

Please correct and return your proofs using the proof correction marks below. For a more detailed look at using these marks please reference the most recent edition of The Chicago Manual of Style and visit them on the Web at: http://www.chicagomanualofstyle.org/home.html

Instruction to typesetter	*Textual mark*	*Marginal mark*
Leave unchanged	· · · under matter to remain	(stet)
Insert in text the matter indicated in the margin	∧	∧ followed by new matter
Delete	/ through single character, rule or underline or ⊢—⊣ through all characters to be deleted	₰
Substitute character or substitute part of one or more word(s)	/ through letter or ⊢—⊣ through characters	new character / or new characters /
Change to italics	— under matter to be changed	(ital)
Change to capitals	≡ under matter to be changed	(caps)
Change to small capitals	= under matter to be changed	(sc)
Change to bold type	~ under matter to be changed	(bf)
Change to bold italic	≂ under matter to be changed	(bf+ital)
Change to lower case	≢	(lc)
Insert superscript	∨	∨ under character e.g. ∨
Insert subscript	∧	∧ over character e.g. ∧
Insert full stop	⊙	⊙
Insert comma	∧	∧
Insert single quotation marks	∨ ∨	∨ ∨
Insert double quotation marks	∨ ∨	∨ ∨
Insert hyphen	=	=
Start new paragraph	¶	¶
Transpose	⊔⊓	⊔⊓
Close up	linking ⁀‿ characters	⁀‿
Insert or substitute space between characters or words	#	#
Reduce space between characters or words	⁀‿	⁀‿