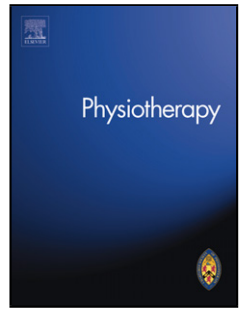
Accepted Manuscript



Title: The reliability and validity of the Figure of 8 Walk test in older people with knee replacement: Does the setting have an impact?

Authors: Karen L. Barker, Martha Batting, Michael Schl*ü*ssel, Meredith Newman

PII: S0031-9406(18)30133-0
DOI: https://doi.org/10.1016/j.physio.2018.07.003
Reference: PHYST 1028

To appear in: *Physiotherapy*

Received date: 7-3-2018
Accepted date: 19-7-2018

The reliability and validity of the Figure of 8 Walk test in older people with knee replacement: Does the setting have an impact?

Karen L Barker, FCSP, PhD;[1-2] Martha Batting, MCSP, MRes;[1] Michael Schlüssel PhD[3], Meredith Newman, MCSP, MSc;[1]

[1]Physiotherapy Research Unit, Nuffield Orthopaedic Centre, Oxford University Hospitals Foundation Trust, Oxford, UK; [2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK, Oxford Clinical Trials Research Unit, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.

Corresponding Author: Professor Karen Barker, Physiotherapy Department, Nuffield Orthopaedic Centre, Oxford University Hospitals Foundation Trust, Windmill Road, Headington, OX3 7HE, UK. Karen.barker@ouh.nhs.uk

Manuscript word count: 3232

Public Trials Registry: BioMed Central, ISRCTN82001483

Abstract:

Objective: To determine the reliability and validity of the Figure of 8 Walk test (F8W) for older people after knee replacement surgery in the home setting.

Design: Observational repeated measures.

Setting: A specialist orthopaedic hospital and participants homes

Participants: Seventy-four older adults aged over 55 years one year following knee replacement surgery participated in two assessments more than one week apart.

Main Outcome Measure: The time to complete, steps, boundary limits and smoothness score from the Figure of 8 Walk Test (F8W), the time to complete the Timed Up and Go (TUG) test and the time to complete the Timed Walk Test (TWT).

Results: Overall, on 95% of occasions, the difference between the measurements for Intra-rater reliability were within 1.8s of the two test times. The difference between the measurements for inter-rater reliability were found to be narrower than for intra-rater reliability with scores within 1.2s overall on 95% of occasions. . The time to perform the F8W was highly correlated to the Timed Up and Go (overall sample r = 0.921) and the variability spread within narrow limits (-0.8 to 0.8 z-scores). This was also true for the Timed Walk Test (overall sample r = 0.834) with a narrow limit of variability on almost all of the observations (-1.16 to 1.16 z-scores).

Conclusions: The F8W test has good reliability and validity when used in either a clinical or home setting for patients around one year following knee replacement surgery.

Contribution of the paper:

- The F8W was a valid and reliable measure of walking ability irrespective of whether it was conducted in the home or clinical setting

- The F8W exposed minor walking difficulties such as stride length and pace that were not apparent with the TUG and TWT making it a more useful measure of functional performance for this population

Keywords: Outcome Assessments, Gait, Balance, Osteoarthritis, Knee replacement and Reproducibility of Results.

INTRODUCTION

In recent years the demand for community rehabilitation has increased which has been caused by the need to reduce hospital length of stay as prolonged periods of admission are known to contribute to poorer mobility and balance. Furthermore there are some political factors driving the need for community rehabilitation including a reduction on dependence and care costs [1] and a need to move from a reactive model of healthcare to a more proactive one [2]. In order to evaluate effectiveness of community rehabilitation,, outcome measures need to be able to accurately portray any changes in mobility.

Despite this, most current measures of balance and gait are conducted in well lit, uncluttered and straight level or single curve paths which do not reflect the reality of mobility in the community. We know that mobilising in a community setting requires us to constantly adapt our balance and walking patterns to avoid obstacles, change direction and carry loads [3]. These measures will not therefore provide an accurate measure of functional mobility required to navigate in the community.

One measure that does recognise the multidirectional and adaptive requirements of gait is the Figure of 8 Walk (F8W) test. The F8W has been developed to measure both straight and curved path walking in clockwise and counter-clockwise directions [4]. However, current studies have only validated the measure in controlled conditions i.e. with smooth floors, well-lit and quiet surroundings. There is therefore the need to validate the measure in the community where it could be used as a gauge for life space mobility.

One group in which the F8W may be a useful tool to measure balance and mobility is older people who have undergone knee replacement surgery due to osteoarthritis (OA). It is recommended that older people reporting a fall or considered at risk of falling should be observed for balance and gait deficits so that their ability to benefit from interventions to improve their balance and strength can be gauged [5]. This population of patients are particularly at risk of falls as 15% of people report a poor outcome following knee replacement surgery due to continuing pain, mobility problems and impaired balance [6-7]. Furthermore it is known that this population has reduced gait speed, endurance and function [8]. The F8W may therefore be a useful outcome measure to use in this population to assess mobility problems, highlight those at risk of falls and monitor interventions.

For this study the F8W will be validated against two other outcome measures commonly used in clinical practice, the Timed Up and Go (TUG) and the Short Physical Performance Battery (SPPB). The TUG includes a single turn around a curve and performance time has been shown to increase significantly with mobility impairments [9]. Furthermore, although the TUG has been investigated in a home setting [10], feedback from community physical therapists suggests performing the TUG is not always feasible due to space constraints in patients' homes. The SPPB includes a timed walk test (TWT) component which provides a measure of dependence and mobility for the elderly [11-12]. However, it only assesses straight path walking and it is not routinely used in the home environment.

The objectives of this study were to examine both the inter-rater and intra-rater (test-retest) reliability of the F8W in both clinical and home settings and to explore its validity by comparing its agreement with the TUG and the TWT component of the SPPB in older people who were 12 months post knee replacement surgery.

METHODS

Participants

Eighty-two participants were recruited from a specialist orthopaedic hospital. Individuals were eligible if they had undergone a primary knee replacement due to OA, were one-year post surgery and were over the age of 55. Individuals were not eligible if they were found to have severe cardiovascular or pulmonary disease, severe dementia or communication difficulties, Rheumatoid Arthritis or a neurological condition that would affect their ability to take part in the walking tests or were due to have further planned physiotherapy or surgical treatment within the next 4 weeks. We enrolled a consecutive series of participants onto the study. Ethical approval was obtained for the study (reference: 15/SW/0138), and all participants were informed about the purpose and procedures of the study and gave their written consent.

Measures

The outcome measures were undertaken following standardised protocols. Background information was collected about an individual's relevant medical history and physical characteristics. Participants also completed the Oxford Knee Score (OKS), which is a brief 12-item questionnaire that measures patient reported outcomes of knee function to provide a picture of the severity and functional impact of their symptoms [13].

Figure of 8 Walk Test

Participants were instructed to stand mid-way between two cones (1.52m (5ft) apart) and to face towards one of the cones. An unmarked 0.6m (2ft) boundary around the testing area was noted by the assessor to check for walking accuracy (See Figure 1). Participants were asked to walk at a comfortable, self-selected speed and direction following a figure of 8 walking path around the cones, stopping when they had returned to their starting positon. The timer began when the participant moved to take their first step and the timer stopped once both feet were back at the starting position. The time and number of steps taken were recorded. The assessor noted whether the participant kept within the unmarked 0.6m boundary and also rated the participant on a 3-item walking smoothness score.

Timed Up and Go

Participants were instructed to begin the test sitting upright with their hands resting beside them in a standard height chair with armrests. They were then asked to stand up from the chair using the armrests if required, walk 3 meters at a comfortable speed, turn around a cone, then return to the chair and sit down. Timing began when the assessor said "Go" and stopped at first contact with the seat of the chair. The time taken to complete the test was recorded in seconds using a stopwatch.

Timed Walk Test

The participant was asked to walk at their usual pace a distance of 2.44m (8ft) and stop. The test was then repeated again and the time it took to complete each test was recorded in seconds using a stopwatch.

Procedures

The F8W was conducted a total of 3 times over 2 separate visits. Half of the participants had their assessments in their own home and half at the orthopaedic hospital. Block randomisation (in blocks of 4) was performed for the testing sequence of each outcome measure using a random numbers table to generate a randomization list based on a 0-9 sequence [14]. The sequence was concealed in sequentially numbered, sealed, opaque envelopes and given to consecutive participants. At the initial assessment all 3 outcome measures were completed and carried out in a varied order.

To establish inter-rater reliability, a second assessment was conducted by two separate assessors; assessor A, who carried out the initial assessment, and a second assessor (assessor B). Both assessor A and assessor B carried out the F8W one time and allowed the patient a 1-minute rest in between the tests. Both assessors were blinded to each other's result. The second assessment was scheduled 7 to 21 days after the first assessment in order to minimise learning effects.

Statistical analysis

Descriptive statistics were used to characterize the study population. Means and standard deviations were calculated for continuous variables; frequencies and percentages for categorical or binary variables. Outcome variables were summarised (mean and 95%CI) by assessment setting.

Intra-rater and inter-rater variation were assessed using Bland and Altman's method of assessing reliability. Agreement between the F8W and the other measures of mobility was also investigated through production of Pearson correlation coefficients and graphically by comparing the time taken to complete the test (TTCT). It is expected in a reliable assessment that the mean difference between measures will be close to zero and the limits of agreement, which contain approximately 95% of the individual differences (2 SDs), should be narrow. To account for the different distances walked in each of the 3 tests z-scores were created to allow the correlation between the TTCT to be calculated.

In order to determine whether the assessment setting contributed to any variability in the TTCT; Bland and Altman plots were produced for the overall sample and subgroups (Home and Clinic).

RESULTS

From the initial 82 participants recruited, data sets for 74 participants were available and included in the analysis (Figure 2). Participant details are reported in Table 1. On average, participants were assessed 15 months (1.9SD) after their knee surgery and had a twelve day (4.9SD) interval between their first and second appointments (Table 1). Thirty-seven (50%) participants were assessed in each setting. The time to complete the F8W was similar when carried out in the home (8.2s) and in clinic (8s) at the initial assessment. The same was true for the TUG (10.3s at home and 10.2 at clinic) and the TWT (3.1s at home and 3s at clinic) (See table 2).

Overall, on 95% of occasions, the difference between the measurements for Intra-rater reliability was within 1.8s. This difference was slightly higher when the assessment setting was in the participant's home (1.9s) and slightly smaller when the setting was in the physiotherapy gym (1.6s) (Figure 3 A,B,C).

The difference between the measurements for inter-rater reliability were found to be narrower than for intra-rater reliability with scores within 1.2s overall on 95% of occasions. This difference was found to be the same when the assessment was in the participants home (1.2s) and slightly higher when in the clinic (1.3s) (Figure 3 D,E,F).

The F8W and TUG times were highly correlated for the overall sample (r = 0.921) and subgroups: Home (r=0.945) and Clinic (r=0.911). The TTCTs for the F8W and TWT were also highly correlated for the overall sample (r = 0.834) and subgroups, Home (r = 0.864) and Clinic (r = 0.793).

The Bland and Altman plots confirm these initial findings and show that the F8W and TUG TTCT had a mean difference of approximately zero, fairly constant variability and almost all scores spread within the narrow limits of agreement (-0.8 to 0.8 z-scores). The same was true for agreement between the F8W and TWT TCTTs with the mean difference close to zero, constant variability and fairly narrow limits of agreement (-1.16 to 1.16 z-scores) (See figure 4).

Finally, all parameters of the F8W were examined to assess any differences in the number of steps, ability to stay within 0.6 m boundary and smoothness ratings. There were comparable performances between groups at Home and in Clinic and for both assessors A and B. While there was complete agreement on the 62 patients who scored *3* on the smoothness score, there was disagreement between the 2 raters on the determination of scores *1* or *2* for 12 participants. Additionally, there was disagreement about whether 4 of the participants had strayed outside of the 0.6 m boundary (Table 3).

DISCUSSION

We chose to investigate the impact of two different test setting, as the majority of research trials report using tests of walking performance exclusively in the clinical setting. Walking within the participants' own homes is a better reflection of their true walking ability and a better predictor of performance and falls risk since it is within the participants' own environment. We have confirmed that the good reliability and validity of the F8W reported in a standardized clinical setting is retained when the test is performed in participants' own homes.

We have also shown that the F8W is a reliable and valid assessment of walking skill in older people who have had a knee replacement within the last 8 to 20 months. The FW8 demonstrated good test-retest reliability and good inter-rater reliability, producing measurements with small inter-rater and intra-rater variability and strong ICCs. The time to complete the F8W was shown to be highly correlated with TTCT from the other mobility assessments and when standardized for different distances walked in each test, a high level of agreement was seen between the F8W, TUG and the TWT.

The study was conducted on a population of patients over the age of 55 who had undergone knee replacement surgery. While the primary purpose of knee replacement is reduction in pain and improved physical function, it is known that around 15% of patients do not report a good outcome from their knee replacement and have continuing pain and mobility problems with impaired balance and decreased walking speed [15-17]. Data from North America indicates that a third of patients report poor patient satisfaction with rehabilitation after knee replacement [18] and systematic reviews have identified that exercise programmes designed to improve function after a knee replacement only result in minor improvements in walking [19,20]. There is agreement that walking is an important outcome after knee replacement as walking speed has been reported to be up to 22% reduced in age-matched men with total knee replacement compared with healthy controls and 17% failed to walk fast enough to cross traffic, walking at less than 1.2m/s [21].

We reported a quicker mean test time for the F8W (8.24 seconds home or 8.01 seconds clinic) when compared to other studies looking at community dwelling older adults with mobility problems: 10.49 seconds [4] and 8.8 seconds [22]. However our mean age of participants was younger than those recruited in both studies with each reporting below average gait speed times for individuals of this age group [23]. Our mean F8W times were more comparable with Welch et al. [24] who reported times of 7.9 seconds for non-fallers and 8.4 seconds for fallers.

The F8W has been shown to pick up subtle changes in walking skill, such as an increase in hesitancy with shorter steps whilst turning about a curve, where other tests of straight path walking have not [4]. This is important as slight changes in gait pattern during daily life, such as stride frequency and length, have been linked with an increased risk of falls [25]. Welch et al. [24] reported that poorer performance of F8W by 1 second was associated with an 8% greater rate of self-reported falls in the previous year.

Other authors have shown the F8W to be a reliable and valid measure of gait performance in community dwelling older adults and stroke participants [4,22,24,26]. However some of these studies failed to recruit an adequate sample size [26] or did not include all components of the F8W scoring system in the final analysis [4,22,24,26].

Although we have concentrated our primary analyses to TTCT, the F8W test has three other components. Our results differed from those of Hess et al. [4] as we report a variation in the smoothness score component between raters with differences in interpretation for the 3 components of hesitancy, stopping and changing pace. Additionally, we also found that the F8W exposed minor walking difficulties that were not apparent on straight line walking, with shorter steps and hesitancy on rounding the cones and a slower pace after direction changes.

We chose to analyse the accuracy of participants, i.e. the ability to stay within the unmarked boundary. For those participants who used a walking stick the accuracy was marked if their feet stayed within the 0.6m boundary while the stick could stray from this boundary without incurring a fail. However, Hess et al. [4] chose not to report accuracy as they found no participants stepped outside the 0.6m boundary, they felt it did not discriminate performance and that the number of steps taken to complete the F8W would act as a proxy measure of accuracy. Although this may have been the case for the sample of community dwelling elders that they tested, in our clinical sample with known lower limb deficits in strength and proprioception there was a minority who could not complete the test within the imposed limits of accuracy (see Table 3). We do not share this view that the number of steps alone will report accuracy as shorter stride lengths will increase the number of steps taken yet the participant may still be within the 0.6m boundary, where the opposite may be true for larger stride lengths. Therefore the number of steps taken may not necessarily account for a walking path outside of the 0.6m exclusion zone and this should be taken into account during the final analysis.

Odonkor et al. [22] also did not evaluate the accuracy component of the F8W. They suggest that navigation of a curved path is more cognitively demanding and suggest the F8W is closely linked to central cognitive-behavioural attributes such as cognitive processing speed which in turn is linked to falls and fall related injuries. By only reporting time to complete the F8W, essential higher level aspects of gait in an elderly population may have been overlooked with undue emphasis placed on cognitive processing speed and less on overall stability and susceptibility to falling. Thus inclusion of the accuracy component of the F8W may be more important for populations with more frailty and may have clinical relevance particularly in identifying subtle problems with performance, for example during dual activities where the additional challenge of balancing speed with accuracy may bring out performance deficits [27].

Study Limitations

To reduce the burden and time commitment for study participant's different groups of participants were assessed either at home or in clinic. To make inferences on the agreement of assessments conducted at different settings, ideally the same group of individuals should have been tested in each setting. However, the close limits of agreement from our results indicates that the F8W is a suitable test to conduct in patients' homes. It also allows us to hypothesise that a good level of agreement could be expected between measures taken in the home or clinical setting in any future F8W studies.

In addition, our sample consisted of 74 volunteers, mobile and healthy enough to complete all 3 outcome assessments at least 1 year after knee replacement surgery. It is reasonable to assume that results could be different for individuals performing the tests in a more acute situation, for example those immediately post joint replacement surgery or with multiple co-morbidities. These acute situations are also more likely to require the participant to use a walking aid which may alter gait patterns and speed of scoring during the F8W. As we only recruited 2 participants who used a stick throughout the testing, we were unable to draw conclusions as to whether the use of a walking aid would affect the F8W scoring.

Finally, it was not the purpose of this study to assess the responsiveness of the FW8 and further work is needed to determine whether the test is sensitive to change and to provide information about the minimally clinically important difference so that it could be effectively used throughout the rehabilitation process.

Conclusions

In a population of older people after knee replacement the F8W was a reliable and valid measure of walking ability irrespective of whether it was conducted in the home or clinical setting. The inclusion of the accuracy and smoothness components of the score is important, potentially giving an early sign of subtle functional deficits that may be linked to cognitive impairments. Additionally, the F8W exposed minor walking difficulties that were not apparent with the TUG and TWT including shorter steps, hesitancy on rounding the cones and a slower pace after direction changes. This makes it a more useful measure of functional performance for this population.

Ethical approval: The study protocol was reviewed and approved by the NRES Committee South West - Exeter UK, reference number 15/SW/0138.

Funding: This work was supported by the National Institute for Health Research (HTA 12/196/08)

Conflicts of Interest: Karen Barker is an Associate Editor for Physiotherapy but had no role in the review of the paper or decision making about its acceptance.

References

[1] Wade D. Community rehabilitation, or rehabilitation in the community?. *Disability and Rehabilitation* 2003;25(15): 875-881.

[2] Edwards N. Community services - how they can transform care [internet]. London: The Kings Fund; 2014 [cited 2018 Mar. 1]. Available from: https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/community-services-nigel-edwards-feb14.pdf

[3] Frank J, Patla A. Balance and mobility challenges in older adults: Implications for preserving community mobility. *American Journal of Preventive Medicine* 2003;25(3):157-163.

[4] Hess R, Brach J, Piva A and VanSwearingen J. Walking Skill can be Assessed in Older Adults: Validity of the Figure-of-8 Walk Test. *Physical Therapy* 2010;90:89-98.

[5] American Geriatrics Society and the British Geriatrics Society. Clinical practice guideline: Prevention of falls in older persons. American Geriatrics Society;2010 [cited 2018 Mar. 1]. Available from: http://www.medcats.com/FALLS/frameset.htm

[6] Jones A, Beaupre A, Johnstone C and Suarez-Almazor E. Total joint arthroplasties: current concepts of participant outcomes after surgery. *Rheumatic Disease Clinics of North America* 2007;33:71-86.

[7] Piva S, Gil A, Almeida G, DiGioia A, Levison T and Fitzgerald K. A Balance Exercise Program Appears to Improve Function for Participants With Total Knee Arthroplasty: A Randomized Clinical Trial. *Journal of the American Physical Therapy Association* 2010;90(6):880–894.

[8] Al-Zahrani K and Bakheit A. A study of the gait characteristics of participants with chronic osteoarthritis of the knee. *Disability and Rehabilitation* 2002;24(5): 275-280.

[9] Bischoff A, Stähelin B, Monsch U, Iversen MD, Weyh A, von Dechend M et al. Identifying a cut-off point for normal mobility: a comparison of the timed ‘up and go’ test in community-dwelling and institutionalised elderly women. *Age Ageing* 2003;32:315-320.

[10] Nikolaus T and Bach M. Preventing Falls in Community-Dwelling Frail Older People Using a Home Intervention Team (HIT): Results From the Randomized Falls-HIT Trial. *Journal of the American Geriatrics Society* 2003;51(3):300-305.

[11] Alexander N and Goldberg A. Clinical gait and stepping performance measures in older adults. *European Review of Aging and Physical Activity* 2006;3:20–8.

[12] Muñoz-Mendoza C, Cabrero-García J, Reig-Ferrer A and Cabañero-Martínez M. Evaluation of walking speed tests as a measurement of functional limitations in elderly people: A structured review. *International Journal of Clinical and Health Psychology* 2010;10(2): 359-378.

[13] Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ et al. The use of the Oxford hip and knee scores. *Journal of Bone Joint Surgery* 2007; 89:1010-1014.

[14] Pocock SJ. Clinical Trials: A practical Approach. Chichester UK: John Wiley & Sons 1984;73-82.

[15] Gage WH, Frank JS, Prentice SD and Stevenson P. Postural responses following a rotational support surface perturbation, following knee joint replacement: frontal plane rotations. *Gait Posture* 2008;27:286–93.

[16] Mandeville D, Osternig LR and Chou LS. The effect of total knee replacement on dynamic support of the body during walking and stair ascent. *Clin Biomech* 2007;22:787–94.

[17] Bruun-Olsen V, Heiberg KE, Wahl AK and Mengshoel AM. The immediate and long term effects of a walking-skill program compared to usual physiotherapy care in patients who have undergone total knee arthroplasty (TKA): a randomized controlled trial. *Disability and Rehabilitation* 2013;35,23:2008-15.

[18] Johnson AJ, Issa K, Naziri Q, Harwin SF, Bonutti PM and Mont MA. Patient dissatisfaction with rehabilitation following primary total knee arthroplasty. *J Knee Surg* 2013;26(6):417-21.

[19] Minns Lowe CJ, Barker KL, Dewey M and Sackley CM. Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2007;335:812.

[20] Artz N, Elvers KT, Minns Lowe C, Sackley C, Jepson P and Beswick AD. Effectiveness of physiotherapy exercise following total knee replacement:

systematic review and meta-analysis. *BMC Musculoskeletal Disorders* 2015;16:15.

[21] Walsh M, Woodhouse LJ, Thomas SG and Finch E. Physical impairments and functional limitations: a comparison of individuals 1 year after total knee arthroplasty with control subjects. *Phys Ther* 1998;78:248–58.

[22] Odonkor C, Thomas J, Holt N, Latham N, Vanswearingen J, Brach JS et al. A Comparison of Straight and Curved Path Walking Tests among Mobility Limited Older Adults. *Journals of gerontology* 2013;68(12):1532-9.

[23] Lusardi MM. Functional Performance in Community Living Older Adults. *Journal of Geriatric Physical Therapy* 2003;26(3):14-22.

[24] Welch SA, Ward RE, Kurlinski LA, Kiely DK, Goldstein R, VanSwearingen J et al. Straight and curved path walking among older adults in primary care: Associations with fall related outcomes. *PM R* 2016;8(8):754-60.

[25] Van Schooten K, Pijnappels M, Rispens S, Elders P, Lips P, Daffertshofer A et al. Daily-Life Gait Quality as Predictor of Falls in Older People: A 1-Year Prospective Cohort Study. *PLoS One* 2016;11(7).

[26] Wong S, Yam M and Ng S. The Figure-of-Eight Walk test: reliability and associations with stroke specific impairements. *Disability and Rehabilitation* 2013;35(22): 1896-1902.

[27] Brach JS, Lowry K, Perera S, Hornyak V, Wert D, Studenski SA et al. Improving motor control in walking: a randomized clinical trial in older adults with subclinical walking difficulty. *Arch Phys Med Rehabil* 2015;96,3:388-94.

Figure 1. Figure of 8 Walk test set up. Arrows indicate direction of walking pattern

Figure 2. Flow of participants throughout study procedures

Figure 3. Bland & Altman plots of Figure of 8 walk test showing intra-rater agreement for assessor A at assessment 1 and 2 (ABC) and inter-rater agreement for assessors A and B at assessment 2 (DEF).

Figure 4. Bland and Altman plots showing limits of agreement (mean 2SD) for the Figure of 8 Walk test and Time to Up and go times (plots ABC) and for the figure of 8 walk test vs Timed Walk Test time (DEF).

Note: Plots A, B and C are z-scores from overall sample, patients' home and physiotherapy clinic respectively; no exclusion of outliers (n = 74, 37 and 37; respectively). Plots D, E and F are z-scores from overall sample, patients' home and physiotherapy clinic respectively; no exclusion of outliers (n = 74, 37 and 37; respectively).

Figure 1.

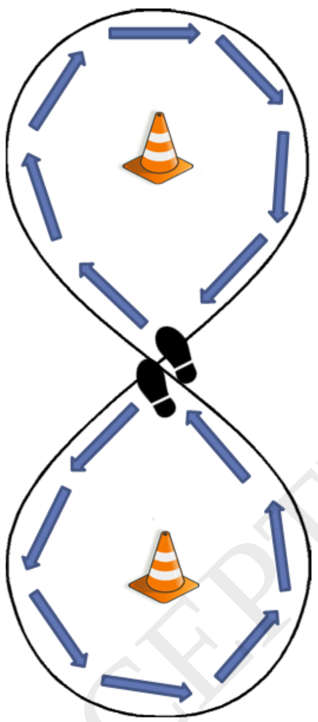


Figure 2.

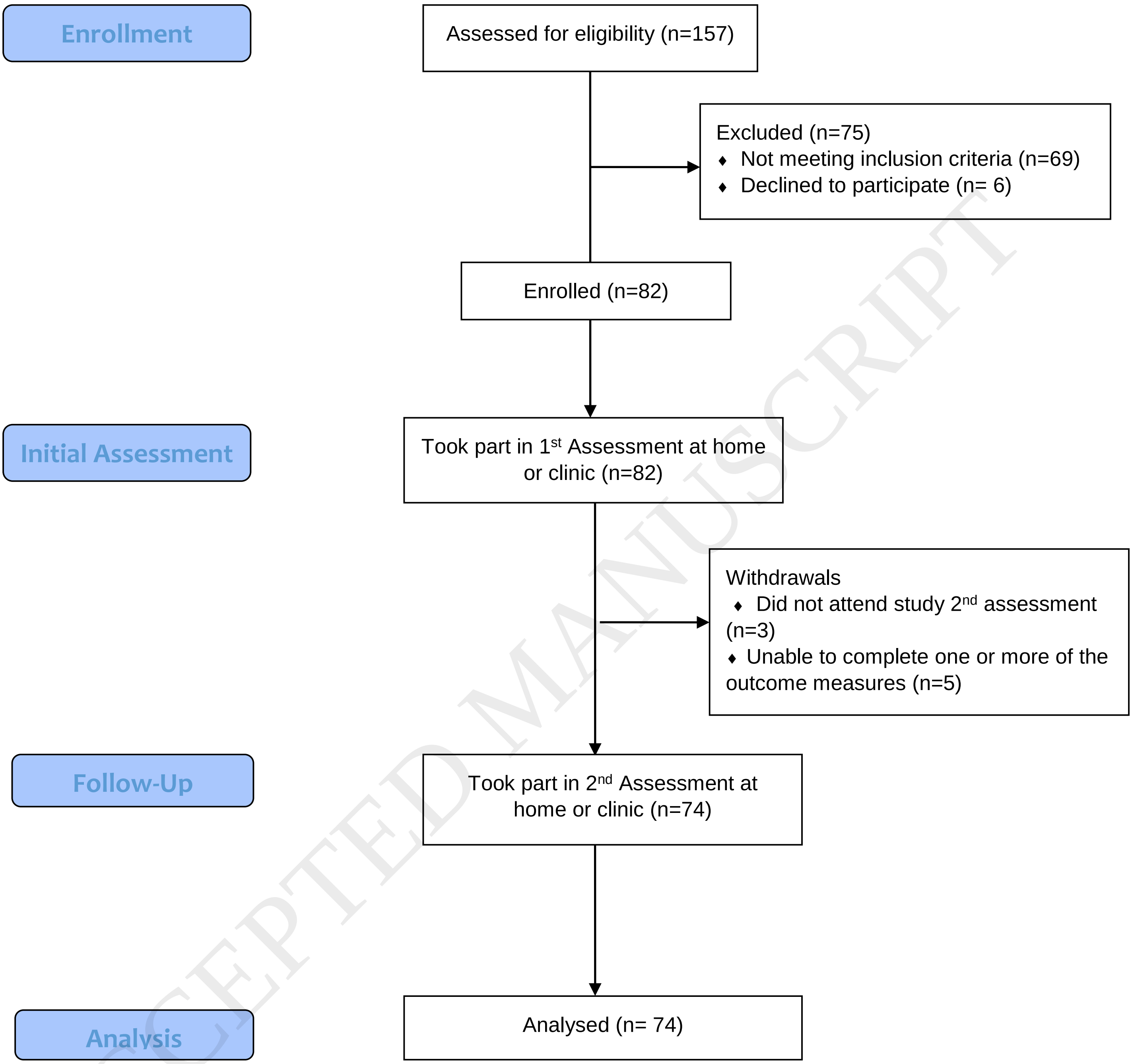

Enrollment
Assessed for eligibility (n=157)
Excluded (n=75)
Not meeting inclusion criteria (n=69)
Declined to participate (n= 6)
Enrolled (n=82)
Initial Assessment
Took part in 1st Assessment at home or clinic (n=82)
Withdrawals
Did not attend study 2nd assessment (n=3)
Unable to complete one or more of the outcome measures (n=5)
Follow-Up
Took part in 2nd Assessment at home or clinic (n=74)
Analysis
Analysed (n= 74)

Figure 3.

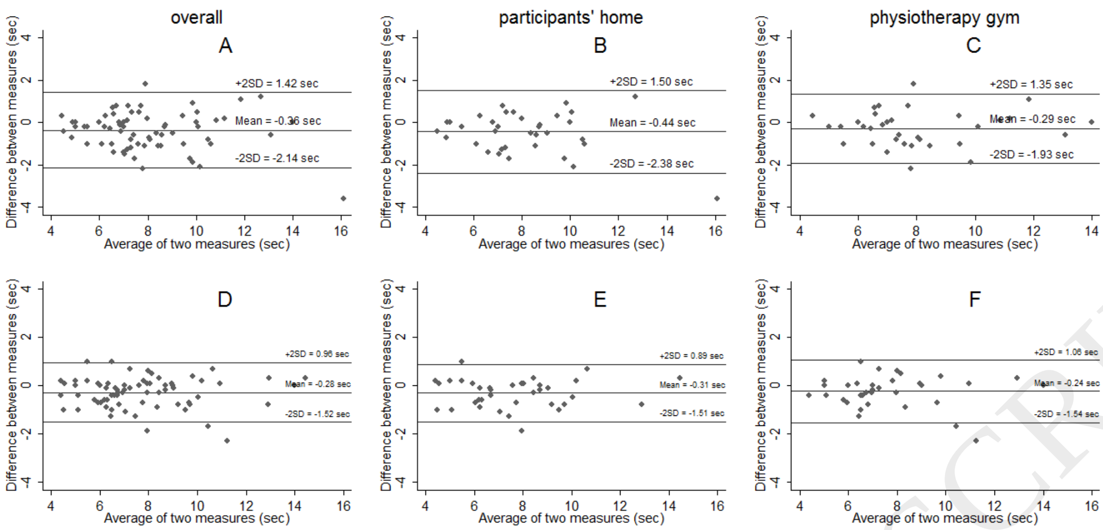

overall
A
+2SD = 1.42 sec
Mean = -0.36 sec
-2SD = -2.14 sec
participants' home
B
+2SD = 1.50 sec
Mean = -0.44 sec
-2SD = -2.38 sec
physiotherapy gym
C
+2SD = 1.35 sec
Mean = -0.29 sec
-2SD = -1.93 sec
D
E
F
Difference between measures (sec)
Average of two measures (sec)

Figure 4.

overall participants' home physiotherapy gym

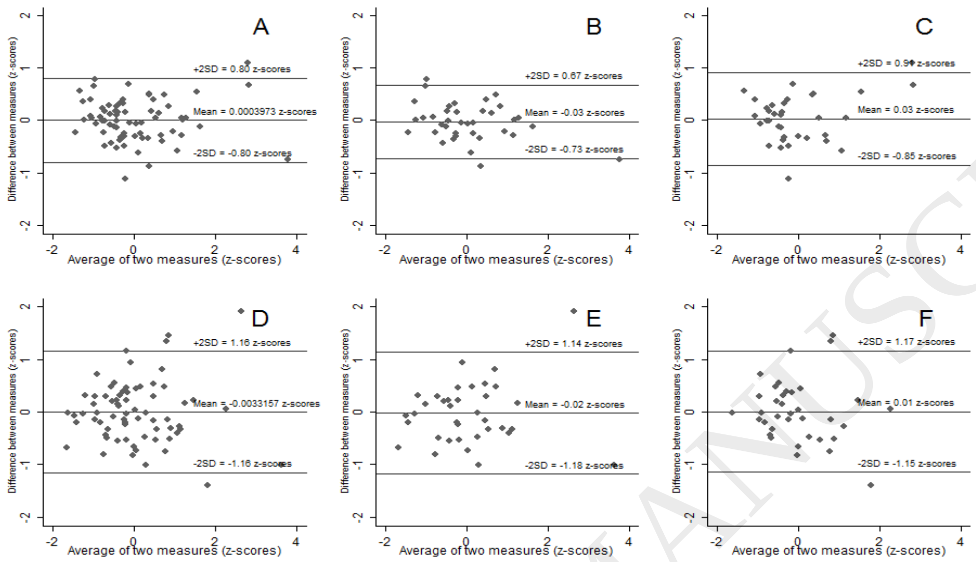


Table 1. Characteristics of participants

	n	Mean (SD)	Median (range)
Age (years)	74	70.3 (7.4)	70 (57 – 88)
BMI (kg/m^2)	74	31.3 (5.4)	30.3 (19.5 – 44.9)
Number of previous surgeries to lower limbs	74	0.9 (0.9)	1 (0 – 4)
Time between surgery and assessment (months)	74	15.1 (1.9)	13 (8 – 20)
Time between assessments (days) [1]	74	12.8 (4.9)	13 (4 – 22)
Oxford Knee Score (OKS)[2]	71	39.8 (7.1)	41 (14 – 48)
	n	%	
BMI (kg/m^2)			
Normal weight (18.00 – 24.99)	7	9	
Overweight (25.00 – 29.99)	28	38	
Obese (30 or more)	39	53	
Surgery type			
Total knee replacement	40	54	
Uni-compartmental knee replacement	34	46	
Other musculoskeletal conditions			
No	43	58	
Yes	31	42	

[1] Time between assessments refers to the day when the first assessment was conducted by assessor A and the day when the second (assessor A) and third (assessor B) assessments were conducted.

[2] Excludes 3 participants because of incomplete OKS forms (missing data).

Table 2. Outcomes summary by Assessment setting.

	Participants' home			Physiotherapy clinic		
	n	Mean (SD)	Median (range)	n	Mean (SD)	Median (range)
First assessment (only assessor A)						
Figure of 8 Time (seconds)	37	8.2 (2.5)	7.9 (4.7 – 17.9)	37	8.0 (2.2)	7.3 (4.3 – 11.1)
Figure of 8 Steps	37	13.7 (2.3)	14 (10 – 21)	37	13.8 (3.0)	13 (9 – 22)
Time to Up and Go (seconds)	37	10.3 (3.3)	9.8 (5.0 – 21.8)	37	10.2 (3.6)	8.6 (6.7 – 21.6)
Timed Walking Test 1 (seconds)	37	3.1 (1.0)	2.9 (1.4 – 6.1)	37	3.0 (0.7)	2.9 (1.7 – 5.0)
Timed Walking Test 2 (seconds)	37	3.1 (1.0)	3.0 (1.5 – 5.9)	37	3.1 (0.8)	3.0 (1.9 – 4.5)
Second assessment						
Assessor A						
Figure of 8 Time (seconds)	37	7.8 (2.3)	7.6 (4.3 – 14.3)	37	7.7 (2.2)	7.1 (4.6 – 14.0)
Figure of 8 Steps	37	13.3 (2.2)	13 (9 – 18)	37	13.6 (3.0)	13 (10 – 22)
Assessor B						
Figure of 8 Time (seconds)	37	7.4 (2.3)	7.0 (4.0 – 14.6)	37	7.5 (2.2)	6.9 (4.2 – 14.0)
Figure of 8 Steps	37	13.0 (2.2)	13 (8 – 18)	37	13.5 (2.9)	13 (10 – 22)

Table 3. Figure of 8 walk test components summary.

	number	% Total
Accuracy - Stayed within the 0.6m surrounding the cones		
First assessment		
No	-	-
Yes	74	100.
Second assessment; assessor A		
No	1	1.35
Yes	73	98.65
Second assessment; assessor B		
No	5	6.76
Yes	69	93.24
Smoothness score		
First assessment		
0	2	2.70
1	6	8.11
2	12	16.22
3	54	72.97
Second assessment; assessor A		
0	1	1.35
1	4	5.41
2	7	9.46
3	62	83.78
Second assessment; assessor B		
0	-	-
1	6	8.11
2	6	8.11
3	62	83.78