Osteoarthritis and Cartilage xxx (xxxx) xxx–xxx

Osteoarthritis and Cartilage

One leg testing in hip and knee osteoarthritis: A comparison with a two-leg oriented functional outcome measure and self-reported functional measures

R. Bendrik # † *, B. Sundström † ‡, K. Bröms #, M. Emtner §, LV Kallings # ¶, M. Peterson # ||

Department of Public Health and Caring Sciences, General Practice, Uppsala University, Uppsala, Sweden
† *Centre for Research and Development, Uppsala University/Region Gävleborg, Gävle, Sweden*
‡ *Department of Public health and Clinical Medicine/Rheumatology, Umeå University, Umeå, Sweden*
§ *Department of Medical Sciences, Uppsala University, Uppsala, Sweden*
¶ *Swedish School of Sport and Health Sciences, GIH, Stockholm, Sweden*
|| *Academic Primary Health Care, Region Uppsala, Sweden*

ARTICLE INFO

Article history:
Received 30 October 2023
Accepted 19 March 2024

Keywords:
Osteoarthritis
Responsiveness
Physical function
Maximal step-up test

SUMMARY

Objective: To compare the responsiveness of two unilateral lower-limb performance-based tests, the one-leg rise test and the maximal step-up test, with the bilateral 30-second chair-stand test and the self-reported measure of physical function (HOOS/KOOS). Specific aims were to evaluate responsiveness, floor/ceiling effect and association between the instruments.

Method: Data was included from 111 participants, mean age 61.3 years (8.3), with clinically verified hip or knee osteoarthritis, who reported less than 150 minutes/week of moderate or vigorous intensity physical activity. Responsiveness, how well the instruments captured improvements, was measured as Cohen's standardised mean difference for effect size, and was assessed from baseline to 12 months following a physical activity intervention. Other assessments were floor and ceiling effects, and correlations between tests.

Results: The maximal step-up test had an effect size of 0.57 (95% CI 0.37, 0.77), the 30-second chair-stand 0.48 (95% CI 0.29, 0.68) and the one-leg rise test 0.12 (95% CI 0.60, 0.31). The one-leg rise test had a floor effect as 72% of the participants scored zero at baseline and 63% at 12 months. The correlation between performance-based tests and questionnaires was considered to be minor (r = 0.188 to 0.226) (p = 0.018 to 0.048).

Conclusion: The unilateral maximal step-up test seems more responsive to change in physical function compared to the bilateral 30-second chair-stand test, although the tests did not differ statistically in effect size. The maximal step-up test provides specific information about each leg for the individual and allows for comparison between the legs.

Introduction

Both performance-based tests and self-reported questionnaires pertaining to physical function are recommended for hip and knee osteoarthritis (OA), as they measure different aspects of function and complement each other.[1] The association between performance-based tests and questionnaires on function is low.[2–4] Performance-based tests of physical function assess what individuals can do, such as walking, going up stairs and getting up from a chair. Self-reported questionnaires concerning physical function usually assess similar domains, though with the disparity that they assess what individuals perceive they can do.[1,5] In comparison, performance-based tests of physical function are generally considered to indicate loss of function earlier than self-reported questionnaires.[6–8] Commonly used self-reported questionnaires in hip or knee OA are the Hip Disability and Osteoarthritis Outcome Score (HOOS) and the Knee Injury and Osteoarthritis Outcome Score (KOOS).[9,10] These questionnaires have

* Correspondence to: Uppsala University, Department of Public Health and Caring Sciences, Box 564, SE-751 22 Uppsala, Sweden.
E-mail addresses: regina.bendrik@pubcare.uu.se (R. Bendrik), bjorn.sundstrom@umu.se (B. Sundström), kristina.broms@pubcare.uu.se (K. Bröms), margareta.emtner@medsci.uu.se (M. Emtner), lena.kallings@gih.se (L. Kallings), magnus.peterson@pubcare.uu.se (M. Peterson).

https://doi.org/10.1016/j.joca.2024.03.111

two subscales in relation to function: activities in daily life and sport/recreation.

The Osteoarthritis Research Society International (OARSI) has recommended a core set of three performance-based tests that evaluate physical function: the 30-second chair-stand test, the 40 m fast-paced walk test and a stair climb test.[1] These tests evaluate forms of function that are important for the osteoarthritis patients to maintain in order to move around and perform daily activities, and the tests are recommended for use in research and monitoring patients with hip or knee OA over time.[1]

However, neither the OARSI core set nor questionnaires evaluate legs separately and systematic reviews have shown that unilateral osteoarthritis is characterised by muscle weakness in the affected leg.[11,12] In the recommended performance-based bilateral tests, the patient may compensate with the other leg. It has been shown that the assumption of bilateral symmetry of lower extremity is low in the bilateral sit-to-stand test.[13] Because improvement in physical function is a goal for many individuals with OA, it is important to use performance-based function tests that are valid, reliable, and responsive to measure change, and that is as truthful as possible. A test that measures each leg separately will provide more specific information for the individual and the caregiver. It is therefore of interest to investigate whether performance-based functional tests that measure each leg separately, i.e. unilateral tests, are more responsive to change, as measured by effect size, compared to tests that evaluate both legs together. The overall aim of this study was to compare two performance-based tests that measure each leg separately, i.e. the maximal step-up test[14] and the one-leg rise test[15] with the bilateral 30-second chair-stand test,[1] as well as self-reported questionnaires on function, the HOOS/KOOS.[9,10] The specific aims were to evaluate responsiveness to changes 12-month post-baseline following a physical activity intervention, including floor and ceiling effect in the three performance-based tests,[1,14,15] and the HOOS/KOOS questionnaires,[9,10] and to evaluate the associations between the performance-based tests and HOOS/KOOS at baseline and after 12 months.

Method

Setting

Data was collected from 111 patients with hip or knee OA participating in a randomised controlled trial (RCT) with the aim of improving physical activity of patients in primary health care in Gävleborg County, Sweden.[16] The participants sought care because of pain or functional problems with either the hip or the knee. The interventions in the RCT were characterised as “individualised advice on physical activity” versus “individualised physical activity on prescription”. Fifty-eight and 53 participants, respectively, were included from the first and second study arms. The primary outcome in the RCT was physical activity. Secondary outcomes were function measured with performance-based tests and questionnaires, and pain and quality of life assessed with questionnaires. The interventions in the two study arms were individualised for the patient to perform the physical activity they preferred. A full description of the RCT assessing the results after 6, 12 and 24 months is provided in previous articles.[16,17] The study was registered at ClinicalTrial.gov (NCT02387034) and was approved by the Regional Ethic Review Board, Uppsala (DNR2010/001). Data used in the current study was collected at baseline and at 12 months.

Participants

Eligibility for the study comprised people aged 40 to 74 with a clinically verified hip or knee OA,[18,19] reporting less than 150 minutes of physical activity of moderate to vigorous intensity per week.[20,21] Exclusion criteria were a diagnosis of hip fracture or a previous hip or knee replacement, meniscal or cruciate ligament injury, neuropathic pain in leg, or concurrent diagnosis of rheumatoid arthritis, severe cardiovascular disease, or cancer. In addition, complete data on the three performance-based tests[1,14,15] was required, as well as the HOOS/KOOS,[9,10] both at baseline and at 12 months.

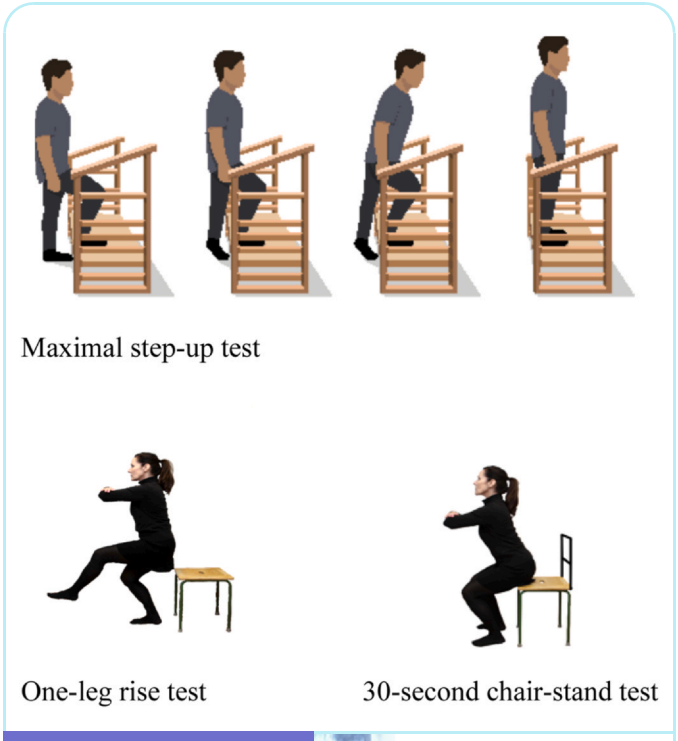

Fig. 1 Osteoarthritis and Cartilage

Illustration of the maximal step-up test, the one-leg rise test and the 30-second chair-stand test.

Measurements

The maximal step-up test (Fig. 1) measures the maximal step-up height in each leg using a step-up box with 3 cm intervals between step heights.[14] The maximal step-up test is significantly correlated to quadriceps strength ($r = 0.68$, $p < 0.001$)[14] measured with Biodex System III PRO.[22] The participant is told to find their balance and take their body weight over the step-up leg placed on the box, and then push the step-up foot into the board while extending the knee. Kick-off with the floor foot or use of the hands while slowly stepping up on to the board is not allowed. The caregiver/examiner assesses whether the test has been carried out correctly. Three attempts are allowed at the current height.[14] The maximal step-up test has shown good reliability for healthy adults between occasions and testers (< 6 cm).[14]

The one-leg rise test (Fig. 1) assesses the total number of repetitions an individual can perform from sitting to standing on one leg, on a chair with a height of 48 cm and performed without a time limit.[15] The one-leg rise test has shown fair reliability between occasions in individuals who underwent meniscectomy,[15] with an intraclass correlation (ICC) of 0.84 (0.69–0.92).

The 30-second chair-stand test (Fig. 1) measures the number of times an individual can go from sitting to standing on a chair with a height of 44 cm, for 30 seconds using both legs.[1] The test has shown

an acceptable level of both between- and within-rater reliability and measurement error (ICC, 0.7–0.8) in people with hip and knee OA, as well as an absence of ceiling and floor effects.[23]

HOOS[9] and KOOS[10] are questionnaires that contain 40 and 42 questions, respectively in five domains and that provide the subscales: pain; other symptoms; function in daily living (ADL); function in sport/recreation; and joint-related quality of life. The analyses used the ADL subscale, consisting of 17 questions scored on a five-point Likert scale. HOOS/KOOS scores can range from 0 (extreme problem) to 100 (no problem).[9,10] Both HOOS and KOOS have shown good reliability between testers and between occasions (ICC ≥ 0.70), and displayed good quality for measurement properties.[24,25]

The learning effects were assumed equal between all participants in the various functional tests and questionnaires.

Descriptive statistics were presented as mean and standard deviation (SD), mean and percent (%), and minimum and maximum scores, at baseline and after 12 months.

Cohen's D, known as standardised mean difference, was used to measure effect size and subsequently compare the responsiveness of change between instruments. Cohen's D is defined as the difference between two means divided by the standard deviation for the data, i.e. between values from baseline and 12 months follow-up and the standard deviation from the paired difference. Effect size was considered as small (d = 0.20–0.49), medium (d = 0.50–0.79) or large (d = ≥0.8).[26,27] A bootstrap method was used to test the difference between the maximal step-up test and the 30-second chair stand test by comparing the respective standardised effect sizes in 10,000 bootstrap samples.

A floor effect was defined as more than 15% of participants achieving the lowest possible score.[28] As the performance-based tests do not have a maximum possible score value in contrast to the questionnaires used, we defined ceiling effects as more than 15% of participants achieving the same maximal score at both baseline and 12 months, and that both values were in the upper 90th percentile.

For reference, the one-leg tests were compared with the two-leg oriented 30-second chair-stand test,[1] and with HOOS/KOOS questionnaires on ADL function. Correlation between performance-based tests and HOOS/KOOS ADL was analysed at baseline and at 12 months using Pearson's correlation coefficient. Definition of the correlations were; negligible 0.00–0.09, small 0.10–0.29, moderate 0.30–0.49 and strong > 0.5.[28]

Results

Data was used in the analyses from 111 patients (Table I). The majority of the participants were women (73%), with the knee joint most commonly affected (72%) in the study population.

The maximal step-up test, which measures each leg separately showed improvement in effect size after 12 months that was considered as medium (0.57; 95% CI 0.37, 0.77) and the effect size of the two-leg oriented 30-second chair-stand test was considered as small (0.48; 95% CI 0.29, 0.68).[26,27] There was no statistical difference between the two tests (p = 0.501). The one-leg rise test, which also evaluates each leg separately, showed an effect size of 0.12 (95% CI −0.60, 0.31) (Table II).

A floor effect was observed in the one-leg rise test, as 72% of the participants at baseline and 63% at 12 months were not able to perform the test. The other performance-based tests of physical function had no significant floor effects and there were no ceiling effects in any of the performance-based tests of physical function (data not shown).

The 30-second chair-stand test showed moderate correlation with the maximal step-up test (r = 0.405, p < 0.001 and r = 0.451, p < 0.001) and with the one-leg rise test (r = 0.301, p < 0.001 and r = 0.450, p < 0.001) measured at baseline and at 12 months, respectively. The correlations between a performance-based test and the HOOS/KOOS ADL questionnaire were small (r < 0.25) measured at baseline and at 12 months (Table III).

Characteristics	n (sd) or n (%)
Age in years	61.3 (8.3)
Women	81 (73)
BMI, kg/m^2	30.2 (4.2)
Joint involved	
Hip	31 (28)
Knee	80 (72)
HOOS[a]/KOOS[b]	
Pain	54.2 (15.7)
Other symptoms	57.3 (18.4)
Activity of daily living	62.2 (17.4)
Sports and recreation	31.2 (22.8)
Quality of life	39.2 (16.0)

[a] HOOS = Hip disability and Osteoarthritis Outcome Score, ranges from 0 (worst) to 100 (best).
[b] KOOS = Knee injury and Osteoarthritis Outcome Score, ranges from 0 (worst) to 100 (best).

Table I Osteoarthritis and Cartilage

Characteristics of 111 individuals with OA participating in physical activity intervention.

A sensitivity analysis produced no significant differences between outcomes in participants with, respectively, hip and knee OA (data not shown).

Discussion

The effect size for the maximal step-up test, which measures each leg separately, was slightly higher than for the two-leg oriented 30-second chair stand test when evaluating physical function after 12 months.[26,27] The effect sizes were 0.57 (95% CI 0.37, 0.77) and 0.48 (95% CI 0.29, 0.68) respectively which were considered medium (between 0.50–0.79) and small (between 0.20–0.49) respectively. However, there was no statistical difference in effect size between the two tests (p = 0.501), and both tests may be useful in clinical practice. Neither of these tests had any significant floor or ceiling effects.

Most participants found the one-leg rise test difficult to perform, with 72% of subjects unable to perform even a single repetition of the test. However, in a previous study in a younger population (mean age of 44.8 years), this test was prognostic in predicting the development of radiographic knee OA 5 years later in subjects with chronic knee pain.[29] In other words, this test might be better suited to a younger population than the one in the present study, where the mean age was 61.3 years. The significant floor effect for this test in our study contributed to a poor effect size of 0.12 (95% CI −0.60, 0.31). The test can thus be deemed inappropriate in an older population.

We found that compared with the HOOS/KOOS questionnaires, the scores for the one-leg tests and the two-leg 30-second chair-stand test showed only a small correlation (r < 0.25). This is consistent with previous studies that have also shown a small correlation between performance-based tests and questionnaires.[2–4] Individuals with hip or knee OA were evaluated before and 3–6 months after joint replacement and increased by up to 43% in self-reported function,[2–4] but did not increase at all in performance-based tests, such as timed up and go[2–4] and five times chair-stand test.[4] An explanation was that they felt they had better function due

	Baseline		12 months		Change		
	Mean (SD)	Min-max	Mean (SD)	Min-max	Mean change	Change (%)	Effect size (95% CI)
Maximal step-up test (cm)	23.4 (6.4)	9–39	26.5 (5.3)	15–39	3.1	13.2	0.57 (0.37, 0.77)
One-leg rise test (numbers)	2.8 (6.4)	0–31	3.5 (7.7)	0–41	0.7	25.0	0.12 (−0.60, 0.31)
30-s chair-stand test (numbers)	11.5 (3.5)	0–20	12.7 (3.2)	5–22	1.2	10.4	0.48 (0.29, 0.68)
HOOS/KOOS ADL (score 0–100)[a]	62 (17)	22–98	74 (19)	16–100	12	19.4	0.65 (0.44, 0.85)

[a] 0 = extreme problem, 100 = no problem.

Table II Osteoarthritis and Cartilage

Scores of performance-based tests and self-reported questionnaire (HOOS/KOOS ADL) at baseline and at 12 months among 111 individuals with hip/knee OA in a physical activity intervention.

	30-s chair-stand test	One-leg rise test	HOOS/KOOS ADL
Correlation at baseline			
Maximal step-up test	r = 0.405, p < 0.001	r = 0.284, p = 0.003	r = 0.226, p = 0.018
30-s chair-stand test		r = 0.301, p < 0.001	r = 0.209, p = 0.028
One-leg rise test			r = 0.188, p = 0.048
Correlation at 12 months			
Maximal step-up test	r = 0.451, p < 0.001	r = 0.415, p < 0.001	r = 0.192, p = 0.043
30-s chair-stand test		r = 0.450, p < 0.001	r = 0.210, p = 0.027
One-leg rise test			r = 0.246, p = 0.009

Table III Osteoarthritis and Cartilage

Correlations between the three different performance-based tests, and between a performance-based test and the self-reported questionnaires on functioning (HOOS/KOOS ADL), in 111 individuals with hip/knee OA participating in a physical activity intervention. Data measured at baseline and at 12 months.

to less pain, but it was too soon after joint replacement for improved function as measured with performance-based tests.[3] Self-reported questionnaires and performance-based tests evaluate different constructs of function and both are needed.[2–4]

There are several benefits to using a test that measures each leg separately in hip and knee OA. First, the result of a one-leg test shows the consequences of OA more clearly than the two-leg tests, as the individual is not able to compensate with the other leg such as in the 30-second chair stand test.[13] Second, the individual consequently obtains information that is more specific if a leg has impaired function. In clinical work, the performance-based tests allow the healthcare professional and the patient to evaluate initial functional capacity and then plan and implement an intervention period and subsequently to evaluate the results of this intervention.[30,31] If single-leg tests are available, both legs can be compared, which is a common strategy in rehabilitation of younger individuals with meniscal or cruciate ligament injuries at risk of osteoarthritis of the knee.[5,32] Furthermore, it is important to evaluate the functional capacity of individuals at risk of falling.[33] Performance-based tests, such as the 30-second chair-stand test and the timed up and go test have been proposed to evaluate functional capacity.[33] However a one-leg test would be useful because functional impairment and asymmetry in leg function is an important risk factor for falls in older adults.[34]

A recent study evaluated a new one-leg test, the step up and down test.[35] Similarly to the maximal step-up test, this test is also considered to be a standardised climbing test, which was valid, reliable and responsive in detecting improvements after an intervention in people with OA.[35] However, even if the step up and down test is described as a single-leg test that can be completed in 15 seconds,[35] we believe it is possible to compensate with the other leg if one leg has lower function. When stepping up and down from the 18 cm high step in 15 seconds, kick off is performed with the foot that is on the lowest step. When performing the maximal step-up test each leg is evaluated separately. The test only evaluates the leg with the foot placed on the step-up box, with the caregiver/examiner checking that there is no kick off with the floor-foot.[14] On the other hand, even though the maximal step-up test was considered a reliable leg function test when evaluated on healthy subjects, the inter-tester and inter-occasion reliability has a disadvantage in that the results may differ by as much as 6 cm between different testers and occasions.[14]

When supporting individuals with osteoarthritis to improve in functions such as climbing stairs and rising from a chair, performance-based tests of function are useful.[1] When using the maximal step-up test in clinical practice, it is possible for the individual to set some goals. It has been argued that there is a possible theoretical maximum step-up height that can be reached when starting with the knee angled at approximately 90° and the femur parallel to the floor.[14] Reaching the theoretical step-up height can be a goal for some individuals. For others, a reasonable goal can be to reach the same step-up heights in both legs, if one leg is weaker.

The strength of this study is that we evaluated tests on one leg in people with hip or knee osteoarthritis and found that a one-leg test can be at least as good as a two-leg test in evaluating change (effect size) after an intervention.

A limitation of this study is that there is no single accepted gold standard performance-based test available with which to compare the one-leg tests. We consequently used the widely used 30-second chair-stand test [1] and HOOS/KOOS questionnaire as reference. [9,10] This is in line with the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist, which suggests comparing with a global instrument if no gold standard is available for comparison.[36]

The maximal step-up test measures the step-up height with 3 cm intervals between step-heights, and the "one-leg rise test" and the "30-second chair stand test" measures times of standing from a standardised chair height. Using a standardised chair makes it easier to perform the latter two tests for a short individual compared with a tall one. A possible limitation is that we did not adjust for the height of the individuals, but rather had the aim of comparing the effect sizes in the same way as in the clinical situations. The conclusion that the maximal step-up test was the most responsive test may be due to its design using multiple step-up heights, easily allowing for individualisation.

Another limitation and issue that needs to be considered is that performance-based tests have a lower effect size than self-reported functional measures,[2–4] and patients therefore do not reach the level of minimum important change when gauged with performance-based measures. Evaluating the smallest within-subject change in score that participants consider to be important has been suggested at an effect size of 0.5,[37] a level that the maximal step-up test reached. However, if both functional outcome measure and self-reported functional measures are used, it is possible to capture the patient's opinion.

Conclusion

The maximal step-up test, a performance-based test that measures each leg separately, seems more responsive to change in physical function, compared to the 30-second chair-stand test which evaluates both legs together, although the tests did not differ statistically in effect size. The maximal step-up test provides more specific information about each leg for the individual and allows for comparison between the two legs if one leg has impaired function.

Role of the funding sources

The project received funding from the Uppsala-Örebro Regional Research Council on three occasions (RFR-81931, RFR-154901, RFR-213801) and from the Centre for Research and Development Uppsala University/Region Gävleborg on two occasions (CFUG-158521, CFUG-572551). Financiers have not taken part in the project design, collection, analysis or interpretation of data, or in writing manuscripts or the decision to submit the manuscript for publication.

Author contributions

R.B., M.P. and B.S.: Study conception and design together with M.E., K.B. and L.V.K. R.B.: Performed the analysis and interpretation of data and wrote the first draft together with M.P. and B.S. R.B., M.P., B.S., L.K., M.E. and K.B.: Participated in the discussions and revisions of the manuscript and approved the final version.

Declaration of competing interest

The authors declare that they have no competing interest.

References

1. Dobson F, Hinman RS, Roos EM, *et al.* OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthr Cartil 2013;21:1042–52.
2. Gandhi R, Tsvetkov D, Davey JR, *et al.* Relationship between self-reported and performance-based tests in a hip and knee joint replacement population. Clin Rheumatol 2009;28:253–7.
3. Dayton MR, Judd DL, Hogan CA, Stevens-Lapsley JE. Performance-based versus self-reported outcomes using the hip disability and osteoarthritis outcome score after total hip arthroplasty. Am J Phys Med Rehabilit 2016;95:132–8.
4. Huber EO, Meichtry A, de Bie RA, Bastiaenen CH. Construct validity of change scores of the Chair Stand Test versus Timed Up and Go test, KOOS questionnaire and the isometric muscle strength test in patients with severe knee osteoarthritis undergoing total knee replacement. Man Ther 2016;21:262–7.
5. Kroman SL, Roos EM, Bennell KL, *et al.* Measurement properties of performance-based outcome measures to assess physical function in young and middle-aged people known to be at high risk of hip and/or knee osteoarthritis: a systematic review. Osteoarthr Cartil 2014;22:26–39.
6. Reuben DB, Valle LA, Hays RD, Siu AL. Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. J Am Geriatr Soc 1995;43:17–23.
7. Brach JS, VanSwearingen JM. Physical impairment and disability: relationship to performance of activities of daily living in community-dwelling older men. Phys Ther 2002;82:752–61.
8. Juhakoski R, Tenhonen S, Anttonen T, *et al.* Factors affecting self-reported pain and physical function in patients with hip osteoarthritis. Arch Phys Med Rehabilit 2008;89:1066–73.
9. Klassbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the Western Ontario and McMaster Universities Osteoarthritis Index. Scand J Rheumatol 2003;32:46–51.
10. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:64.
11. Loureiro A, Mills PM, Barrett RS. Muscle weakness in hip osteoarthritis: a systematic review. Arthr Care Res 2013;65:340–52.
12. dos Santos Siqueira M, Souto LR, Martinez AF, *et al.* Muscle activation, strength, and volume in people with patellofemoral osteoarthritis: a systematic review and meta-analysis. Osteoarthr Cartil 2022;30:935–44.
13. Lundin T, Grabiner M, Jahnigen D. On the assumption of bilateral lower extremity joint moment symmetry during the sit-to-stand task. J Biomech 1995;28:109–12.
14. Nyberg LA, Hellenius ML, Kowalski J, *et al.* Repeatability and validity of a standardised maximal step-up test for leg function–a diagnostic accuracy study. BMC Musculoskelet Disord 2011;12:191.
15. Bremander A, Dahl L, Roos E. Validity and reliability of functional performance tests in meniscectomized patients with or without knee osteoarthritis. Scand J Med Sci Sports 2007;17:120–7.
16. Bendrik R, Kallings LV, Bröms K, *et al.* Physical activity on prescription in patients with hip or knee osteoarthritis: a randomized controlled trial. Clin Rehabilit 2021;35:1465–77.
17. Bendrik R, Kallings LV, Bröms K, Emtner M. Follow-up of individualised physical activity on prescription and individualised advice in patients with hip or knee osteoarthritis: a randomised controlled trial. Clin Rehabil 2024. 2692155241234666.
18. Altman RD. Criteria for classification of clinical osteoarthritis. J Rheumatol Suppl 1991;27:10–2.

19. Zhang W, Doherty M, Peat G, *et al*. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis 2010;69:483–9.
20. The Swedish National Board of Health and Welfare. National guidelines for disease prevention methods 2011, Indicators, Appendix; 2011.
21. Olsson SJ, Ekblom Ö, Andersson E, *et al*. Categorical answer modes provide superior validity to open answers when asking for level of physical activity: a cross-sectional study. Scand J Public Health 2016;44:70–6.
22. Drouin JM, Valovich-mcLeod TC, Shultz SJ, *et al*. Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. Eur J Appl Physiol 2004;91:22–9.
23. Dobson F, Hinman RS, Hall M, *et al*. Reliability and measurement error of the Osteoarthritis Research Society International (OARSI) recommended performance-based tests of physical function in people with hip and knee osteoarthritis. Osteoarthr Cartil 2017;25:1792–6.
24. Thorborg K, Roos EM, Bartels EM, *et al*. Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: a systematic review. Br J Sports Med 2010;44:1186–96.
25. Collins NJ, Prinsen CA, Christensen R, *et al*. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthr Cartil 2016;24:1317–29.
26. Cohen J. The t test for means. Statistical Power Analysis for the Behavioural Sciences. Hillsdale, NJ: Earlbaum; 1988. p. 567.
27. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989:S178–89.
28. Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
29. Thorstensson CA, Petersson IF, Jacobsson LT, *et al*. Reduced functional performance in the lower extremity predicted radiographic knee osteoarthritis five years later. Ann Rheum Dis 2004;63:402–7.
30. Holden MA, Button K, Collins NJ, *et al*. Guidance for implementing best practice therapeutic exercise for patients with knee and hip osteoarthritis: what does the current evidence base tell us? Arthr Care Res 2021;73:1746–53.
31. Holden MA, Metcalf B, Lawford BJ, *et al*. Recommendations for the delivery of therapeutic exercise for people with knee and/or hip osteoarthritis. An international consensus study from the OARSI Rehabilitation Discussion Group. Osteoarthr Cartil 2023;31:386–96.
32. Berg B, Urhausen AP, Øiestad BE, *et al*. What tests should be used to assess functional performance in youth and young adults following anterior cruciate ligament or meniscal injury? A systematic review of measurement properties for the OPTIKNEE consensus. Br J Sports Med 2022;56:1454–64.
33. Skelton DA, Kennedy J, Rutherford OM. Explosive power and asymmetry in leg muscle function in frequent fallers and non-fallers aged over 65. Age Ageing 2002;31:119–25.
34. Jiménez-Lupión D, Chirosa-Ríos L, Martínez-García D, *et al*. Effects of power training on functional capacity related to fall risk in older adults: a systematic review and meta-analysis. Arch Phys Med Rehabilit 2023;104:1514–25.
35. Almeida GPL, Monteiro IO, de Oliveira Dantas RG, *et al*. Reliability, validity and responsiveness of the Step Up and Down (StUD) test for individuals with symptomatic knee osteoarthritis. Musculoskelet Sci Pract 2021;56, 102454.
36. Mokkink LB, Terwee CB, Patrick DL, *et al*. COSMIN Checklist Manual. Amsterdam: University Medical Center,; 2012.
37. Terwee CB, Roorda LD, Dekker J, *et al*. Mind the MIC: large variation among populations and methods. J Clin Epidemiol 2010;63:524–34.