Osteoarthritis and Cartilage (2010) **18,** 365–371

doi:10.1016/j.joca.2009.10.001

Osteoarthritis and Cartilage

An examination chair to measure internal rotation of the hip in routine settings: a validation study

S. Reichenbach†‡§*[a], P. Jüni†§[a], E. Nüesch†§, F. Frey‖, R. Ganz¶ and M. Leunig#

† *Division of Clinical Epidemiology and Biostatistics, Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland*

‡ *Department of Rheumatology, Clinical Immunology and Allergology, Inselspital, University of Bern, Switzerland*

§ *CTU Bern, Inselspital, University of Bern, Switzerland*

‖ *Military Medical Branch, Armed Forces Logistics Organisation (AFLO), Switzerland*

¶ *Department of Orthopaedic Surgery, University of Bern, Switzerland*

Hip Service, Schulthess Clinic, Zürich, Switzerland and University of Bern, Switzerland

Summary

Objective: To determine the performance of a newly developed examination chair as compared with the clinical standard of assessing internal rotation (IR) of the flexed hip with a goniometer.

Methods: The examination chair allowed measurement of IR in a sitting position simultaneously in both hips, with hips and knees flexed 90°, lower legs hanging unsupported and a standardized load of 5 kg applied to both ankles using a bilateral pulley system. Clinical assessment of IR was performed in supine position with hips and knees flexed 90° using a goniometer. Within the framework of a population-based inception cohort study, we calculated inter-observer agreement in two samples of 84 and 64 consecutive, unselected young asymptomatic males using intra-class correlation coefficients (ICC) and determined the correlation between IR assessed with examination chair and clinical assessment.

Results: Inter-observer agreement was excellent for the examination chair (ICC right hip, 0.92, 95% confidence interval [CI] 0.89–0.95; ICC left hip, 0.90, 95% CI 0.86–0.94), and considerably higher than that seen with clinical assessment (ICC right hip, 0.65, 95% CI 0.49–0.77; ICC left hip, 0.69, 95% CI 0.54–0.80, *P* for difference in ICC between examination chair and clinical assessment ≤0.001). The average range of motion (ROM) obtained with examination chair and clinical assessment were similar (difference 1.1°, 95% CI − 0.7–2.8°, $P = 0.23$), and the correlation was strong (Pearson's coefficient, 0.75, 95% CI 0.62–0.84).

Conclusions: The use of the examination chair resulted in a precise assessment of hip IR in our population-based inception cohort study of young asymptomatic males. It was strongly correlated with standard clinical assessment of IR but was considerably more reliable.

Key words: Internal rotation, Hip, Examination, Chair, Femoroacetabular impingement, Validation.

Introduction

Hip osteoarthritis (OA) is a common form of joint disease and a leading cause of pain and disability in older people.[1] Based on a population-based survey in Johnston County, NC, it has been estimated that among US adults ages 45 and older, the prevalence of radiographic hip OA is 27%.[2] The etiology of hip OA is multifactorial.[3] Current classification systems differentiate idiopathic OA in individuals without established risk factors from secondary OA, which may result from trauma, or metabolic, inflammatory, or developmental diseases.[4] Recently, it was proposed that the majority of cases of hip OA traditionally classified as idiopathic may in fact be secondary, due to subtle developmental abnormalities such as femoroacetabular impingement.[5,6] The 'cam' type of femoroacetabular impingement is found predominantly in males. It is caused by a deformity of the femoral head and neck, with a nonspherical extension of the femoral head and/or a decreased head–neck offset.[7] The increasing radius of the femoral head entering the acetabulum results in shearing forces at the acetabular cartilage, especially during flexion and internal rotation (IR). High velocity movements, frequently occurring during athletic exercises, may result in an outside-in abrasion of the anterosuperior acetabular cartilage and to an avulsion of the cartilage at the labrum and of the subchondral bone at the anterior–superior rim.[8] The most frequent location of the deformity is the anterosuperior region. A limited range of motion (ROM) in IR is a standard criterion of the American College of Rheumatology for the assessment of hip OA.[9] However, 100 years ago, Preisser[10] suggested that limited IR was not only a sign of manifest hip OA, but a precursor, actually detectable before symptomatic hip OA occurred. Subsequently, it was recognised that a limited IR of the hip in symptomatic individuals

[a]Both authors contributed equally to this work.

*Address correspondence and reprint requests to: Stephan Reichenbach, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, 3012 Bern, Switzerland. Tel: 41-31-631-35-29; Fax: 41-31-631-35-20; E-mail: rbach@ispm.unibe.ch

Received 11 June 2009; revision accepted 1 October 2009.

may suggest the presence of a cam type deformity with a nonspherical extension of the femoral head and a decreased anterior head–neck offset,[6,11] which in turn may result in the development of hip OA.[12,13] IR may therefore not only be useful in the assessment of individuals with manifest disease but valuable for detecting asymptomatic individuals without established radiographic signs of OA, who are at risk to develop secondary OA caused by cam type impingement mechanisms. However, the usual clinical assessment of IR with a goniometer has limited reproducibility[14] because of measurement error and potential systematic differences in the use and interpretation of the goniometer between observers. Random misclassification due to measurement error will result in bias towards the null, i.e., an underestimation of associations between initial clinical features and subsequent radiographic or clinical outcomes. Systematic differences in the use or interpretation of the goniometer may result in both, an under or overestimation of associations. A more precise, reproducible assessment of IR would therefore be desirable in routine clinical settings to ensure reliable clinical classification of patients, and in epidemiological studies to minimise bias.

Several techniques were developed in recent years, including the use of an inclinometer[15] or a plurimeter[14] to measure IR in the supine position. Other techniques used a sitting position, with legs of examined individuals unsupported over the edge of a bed or a chair, assessing IR with a plurimeter,[16] goniometer,[17] or an inclinometer.[18] The use of plurimeters[14,16] and inclinometers[15,18] by experienced research staff with formal clinical training in orthopaedics or rheumatology may increase reproducibility. However, their use has not gained wide acceptance in routine clinical settings and has not been established in epidemiological studies, which typically are subject to time constraints and may have research personnel with only limited training in one of the relevant clinical fields, such as orthopaedics or rheumatology.

This paper describes an examination chair as a new and simple device to be used by research staff with limited clinical training in time constrained settings to measure ROM of passive IR of the hip joint in a sitting position. Our analyses determined the chair's performance as compared with the clinical standard of assessing ROM of IR in a supine position with a goniometer.

Methods

PARTICIPANTS

Participants in this study were consecutively recruited from a large population-based inception cohort study of young males undergoing conscription for the Swiss army.[19] All male individuals in Switzerland, regardless of their health status, are required by the Swiss army to attend a 3-day recruitment process in specialized centers. Consecutive individuals seen in one center were asked to participate in this study. Participants were excluded if they had had an operation of the hip joint, an inflammatory or metabolic rheumatic disease, or a history of hemophilia. Participants completed questionnaires pertaining to pain, stiffness, and physical function, using the respective subscales of the Western Ontario and McMaster University Osteoarthritis Index (WOMAC, version 3.1).[20] The WOMAC pain subscale ranged from 5 (no pain) to 25 (most severe pain), the WOMAC stiffness subscale from 2 (no symptoms) to 10 (severe symptoms), and the WOMAC function subscale from 17 (no function impairment) to 85 (most severe function impairment). Approval for this study was obtained from the Research Ethics Committee of the Canton of Bern. Written informed consent was obtained from all participants prior to the data collection.

ASSESSMENT OF IR USING EXAMINATION CHAIR

IR was measured on an examination chair that allowed ROM measurement in a sitting position, with the hips and knees flexed 90° and the lower legs unsupported over the edge of the bed (Fig. 1). To prevent compensatory movements, the pelvis was stabilized using a belt to secure it to the chair. A second belt was placed around the knees to ensure that the thighs remained parallel to each other. Stretching increases the passive ROM of IR.[17] Therefore, a standardized load of 5 kg was applied to both ankles simultaneously using a bilateral pulley system, and IR of both hips was measured after an adaptation period of 30 s. The choice of 5 kg was based on the results from pilot investigations in asymptomatic individuals with healthy hips in whom variable loads ranging from 2 to 10 kg were used. 5 kg appeared to be the optimal compromise between the passive ROM attained and the level of discomfort experienced by the examined individuals. The loads were applied at the ankles, 3 cm above the transmalleolar axis, through slow manual release of a pulley system with inextensible cables and straps (Fig. 1). The choice of an adaptation period of 30 s duration was in accordance with the American College of Sports Medicine's Guidelines for Exercise Testing and Prescription.[21] Ten seconds of static stretching are generally sufficient for an adaptation of Golgi tendon organs, with no relevant difference in the ROM between 10, 20 and 30 s.[22] 30 s were deemed sufficiently long for

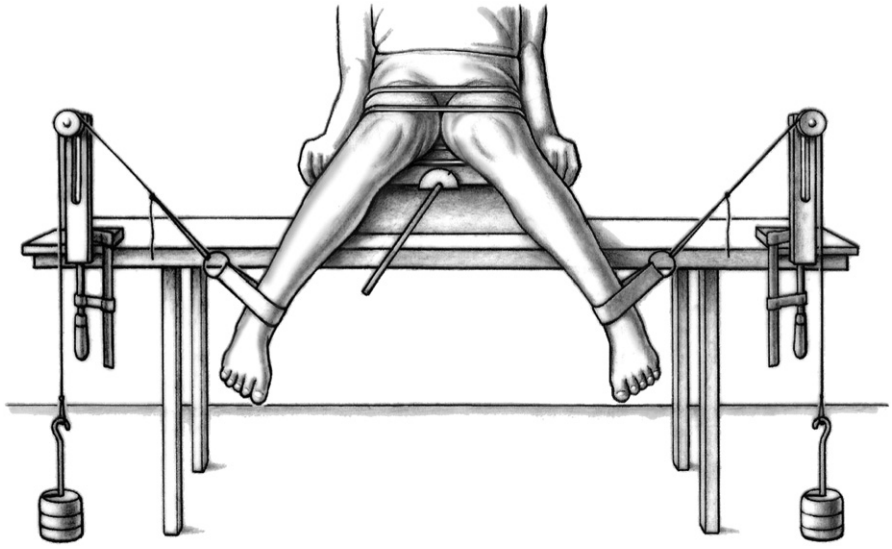

Fig. 1. IR measured by a new device. The participant is placed on a chair, with the hip flexed at 90°. See text for detailed description.

the muscle to stretch, the participant to relax and the maximum ROM to be achieved.[17] The ROM was measured using a protractor fixed to the chair. The arm of the protractor was set parallel to the longitudinal axis of the lower leg, and the angle between the vertical null position and the protractor used to calculate IR. Participants were required to wear shorts or non-restricting clothes.

CLINICAL ASSESSMENT OF IR USING GONIOMETER

The clinical assessment of IR with a goniometer was considered the clinical standard. The participant was in supine position on an examination table, with the hip passively flexed at 90° (0° position). The examiner then rotated the hip internally with the lower leg held parallel to the coronal plane of the patient, while avoiding abduction (Fig. 2). The maximum angle was measured using a two-arm goniometer after maximal passive IR was achieved. The goniometer (Kirchner & Wilhelm GmbH +Co, Asperg, Germany) is a plastic short-armed (35 cm) instrument with a 180° scale marked in 1° increments. The center of the goniometer was positioned over the midpoint of the patella, with one arm set along the longitudinal axis of the lower leg defined by patellar tendon and tibia axis and the other arm set parallel to the sagittal plane.[23,24] A research technician recorded the angle between longitudinal axis and sagittal plane of the patient. The measurement was performed twice, with the leg returned to the neutral position before the second measurement was recorded and the mean of the two evaluations was used.

PROCEDURES

For logistic reasons, two inter-observer agreement studies were performed separately for examination chair and clinical assessment with the goniometer in independent samples of consecutive participants. Two observers, a rheumatologist (SR) and a research technician, measured IR independently on the examination chair in the first sample, two other observers, an experienced senior orthopaedic surgeon (RG) and a physician trained in rheumatology (PJ), measured IR with the two-arm goniometer independently in the second sample. Measurements took place on the same day using the same devices, with all observers blinded to each other's measurements. The participants included in the second inter-observer agreement study of the clinical assessment with the goniometer were additionally assessed on the examination chair by an independent observer (SR), who was unaware of the assessments of the other observers.

STATISTICAL ANALYSIS

Using a two-way analysis of variance we calculated intra-class correlation coefficients (ICC) as a measure of inter-observer agreement. ICCs are estimates of the average concordance between different assessments[25] and values of >0.75 can be interpreted a s excellent, 0.4–0.75 as moderate, and <0.4 as poor agreement.[26] An ICC of 0.50 was expected for clinical assessments with the goniometer[14,16] and was considered the null value against which the ICC of the examination chair was tested. We estimated that a sample size of $\geq$65 participants would provide $>$80% power to detect a minimal clinically relevant ICC of 0.70 for the inter-observer agreement of the examination chair to exceed an ICC of 0.50 at a one-sided α of 5%.[27] Then, we estimated two-sided *p*-values for the difference in ICCs between examination chair and clinical assessment from approximate *z*-tests. We quantified the magnitude of measurement error using the within-subject standard deviation (SD) ζ_w and calculated repeatability coefficients defined as 2.77 ζ_w. The difference between two measurements for the same hip is expected to be less than this coefficient for 95% of pairs of measurements.[28] Smaller repeatability coefficients indicate that measurements of IR on the same subject are less variable. Bland–Altman plots were used to display differences between two measurements against their means, allowing visual inspection of the relationship between the magnitude of IR and measurement error.[29] If the measurement error is unrelated to the size of IR, a random scatter can be expected. Then, we determined the correlation between examination chair and clinical assessment using Pearson's coefficient. Clinical assessments with the goniometer were carried out in duplicate by two independent observers, therefore, we randomly selected one assessment for each hip. A correlation of >0.70 was considered to indicate a strong correlation. All analyses were performed in STATA version 10 (Stata Corporation, College Station, Texas).

Results

1141 individuals were eligible for the study, and 1080 were included in the inception cohort study. Table I summarizes the characteristics of the cohort. All participants were young males with a mean age of 20 years and an average body mass index (BMI) of 23 kg/m^2. In general, they were asymptomatic, as indicated by a low WOMAC score. Measurements of IR on the examination chair typically took 3 min. The mean ROM of IR was 36°.

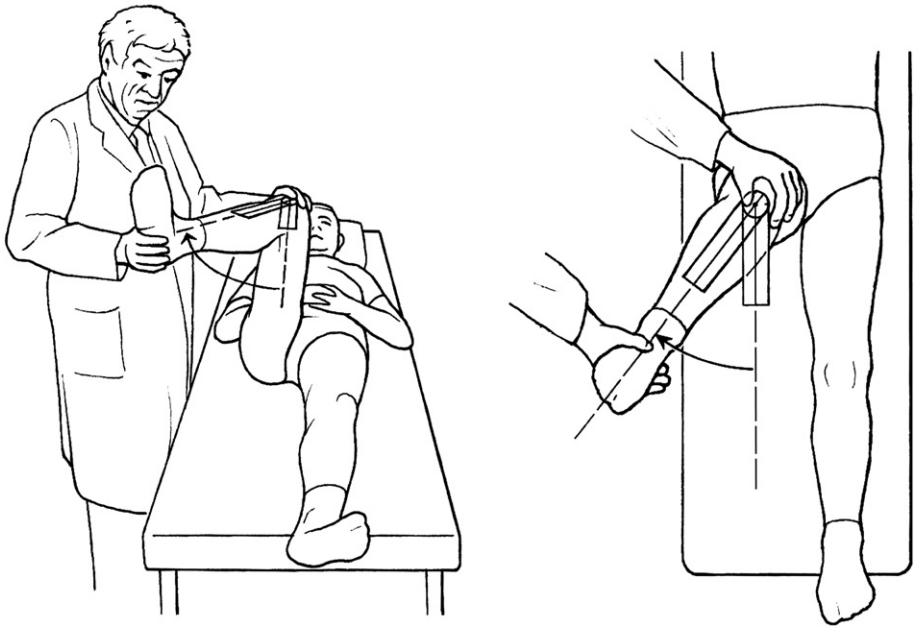

Fig. 2. Clinical assessment of IR, with the participant in supine position and the hip flexed at 90°. Rotation is assessed while avoiding abduction of the hip.

Table I
Characteristics of 1080 male volunteers

Characteristics	Mean ± SD
Age (years)	19.9 ± 0.8
Height (cm)	178.0 ± 6.4
Weight (kg)	73.3 ± 12.6
BMI (kg/m^2)	23.1 ± 3.8
WOMAC overall*	25.4 ± 3.9
WOMAC pain subscale	5.2 ± 1.1
WOMAC stiffness subscale	2.4 ± 1.1
WOMAC function subscale	17.7 ± 2.6
IR − right hip (°)	35.9 ± 6.7
IR − left hip (°)	36.6 ± 6.3

SD: standard deviation, BMI: body mass index, WOMAC: Western Ontario and McMaster University Osteoarthritis Index, IR: internal rotation (degree).

84 consecutive participants were included in the inter-observer agreement study of the examination chair. Table IIA presents means and SDs of IR, ICCs as measures of concordance between observers, within-subject SDs as an estimate of the measurement error and repeatability. The ROM of IR showed excellent agreement with ICCs of 0.92 (95% confidence intervals [CI] 0.88−0.95) for the right hip and 0.90 (95% CI 0.85−0.93) for the left. The estimated measurement errors were low (within-subject SD 1.57° for the right and 1.75° for the left hip), and no clear relationship with the size of ROM measurement was found by visual inspection [Fig. 3(A)]. Repeatability coefficients of 4.8 and 4.9° indicated that the difference between two measurements for the same hip could be expected to be less than 5°.

67 consecutive participants were included in the inter-observer agreement study using the goniometer in supine position of participants (Table IIB). Inter-observer agreement of was moderate with ICC's of 0.65 (95% CI 0.49−0.77) for the right hip and 0.69 (95% CI 0.54−0.80) for the left. The estimated measurement errors were approximately three times higher than for the examination chair. Accordingly, repeatability coefficients indicated that the absolute difference between two measurements for the same hip could be as high as 15°. Figure 3B suggests a relationship between measurement error and IR. The *P*-values for differences in concordance between examination chair and goniometer were <0.001 for both hips in favour of the examination chair.

Figure 4 presents a plot of clinical assessments with goniometer on the *x*-axis against examination chair on the *y*-axis in 67 consecutive participants. The corresponding correlation coefficients were 0.75 for both hips (95% CI 0.62−0.84). The average ROM obtained with examination chair and clinical assessment were similar (difference 1.1°, 95% CI − 0.7−2.8°, $P = 0.23$). With repeatability coefficients of 10.5 and 10.4° for right and left hip, however, the absolute difference in IR between examination chair and clinical assessment with the goniometer could be expected to be as high as 11°.

Discussion

The examination chair presented in this study showed excellent inter-observer agreements, which were clearly higher than those found for the clinical assessment. The clear-cut differences in agreement did not translate into a lack of correlation between the two approaches: a Pearson's coefficient of 0.75 suggested a strong correlation between the two approaches. Taken together, this means that the examination chair measures the same clinical construct as the clinical assessment, but with a considerable increase in precision. The increase in precision means that bias towards the null − resulting in an underestimation of associations between initial clinical features and subsequent radiographic or clinical outcomes − is minimized. In addition, the variation in the use and interpretation across observers appeared minimal with the examination chair in view of the small repeatability coefficients. The examination was fast and easy and could be performed by research personnel without full clinical training in a population-based research setting.

Our study was adequately powered and our population-based sample of young healthy men included the full range of internal hip rotation, from under 20° up to nearly 60° of ROM. In view of its wide use and acceptance, we considered the clinical assessment of IR using a goniometer as the comparator in our study. Some may argue that we should have compared the examination chair to assessments with inclinometers[15] or plurimeters.[16] These tools have not been widely adopted, however, neither in routine settings, nor in clinical research, and lack appropriately powered population-based validation studies in comparable populations. The study was performed within the framework of the baseline examination of a prospective population-based cohort study, which aims at exploring the prognostic importance of cam type femoroacetabular impingement for the development of hip OA. Cam type impingement is predominantly found in males and the study was designed as an inception cohort. Therefore the baseline examination was restricted to young asymptomatic males who did not have formal evidence of hip OA and our results may not be generalizable to other populations, such as elderly

Table II
Inter-observer agreement measures of internal rotation

	IR (Mean ± SD)		ICC (95% CI)	Within-subject SD	Repeatability coefficient
A. Examination chair					
	Observer A $n = 84$	Observer B $n = 84$			
Right hip	36.9° ± 5.8	36.9° ± 5.4	0.92 (0.88−0.95)	1.8°	4.8°
Left hip	38.1° ± 5.5	38.4° ± 5.6	0.90 (0.85−0.93)	1.8°	4.9°
B. Clinical examination					
	Observer C $n = 67$	Observer D $n = 67$			
Right hip	34.3° ± 7.5	36.4° ± 9.7	0.65 (0.49−0.77)	5.1°	14.2°
Left hip	35.8° ± 6.8	36.9° ± 9.9	0.69 (0.54−0.80)	4.7°	13.1°

IR: internal rotation, SD: standard deviation, ICC: intra-class correlation coefficient, 95% CI: 95% confidence interval.

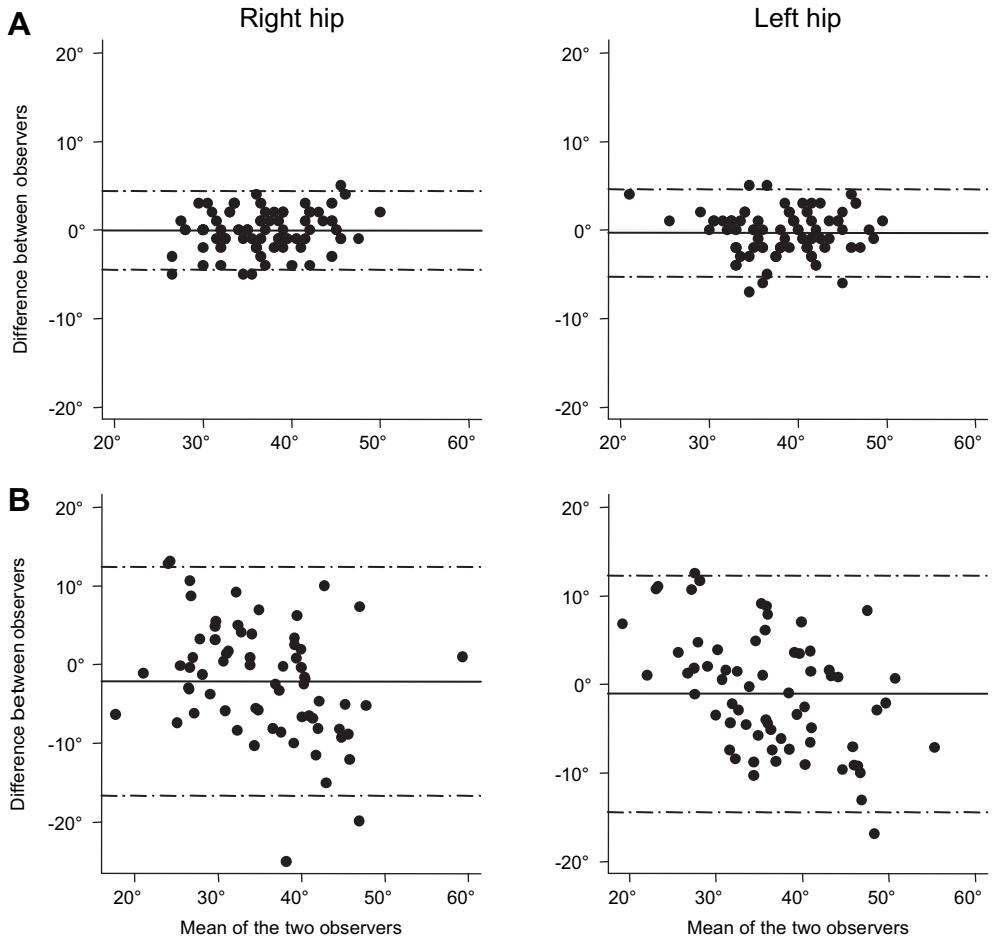

Fig. 3. Bland–Altman Plots of IR. Means of the two measurements of each subject's IR on the *x*-axis are plotted against the difference in each subject's two measurements on the *y*-axis for A) the examination chair ($n = 84$ participants), and B) clinical examination using the goniometer ($n = 67$ participants).

individuals or females. We assessed IR of both hips simultaneously, but this may not work in individuals with symptomatic hip OA, particularly if the disease is unilateral: the stretching of the painful leg could affect the ROM of the contralateral healthy hip. Comorbid musculoskeletal conditions found in elderly individuals, including problems with the lower back and pelvis, may lead to discrepancies between assessments performed in sitting and supine

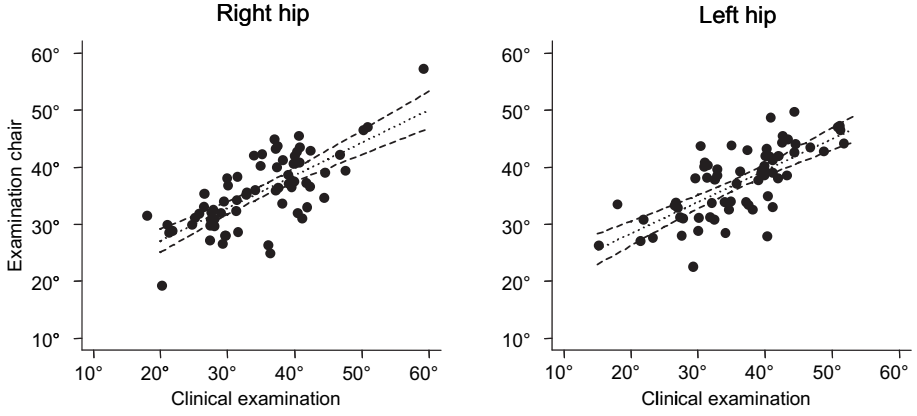

Fig. 4. Difference in assessed IR between clinical examination and examination chair for the right and the left hips. IR of the hip as measured on the examination chair was plotted against the IR as measured *via* clinical examination. The dotted line indicates the correlation between the two measurements. Dashed lines represent the 95% CIs.

positions. The sitting position is likely to offer better stabilization of lower back and pelvis and to provide more comfort to the patient than the supine position. Therefore, we would give precedence to results obtained with the examination chair in case of discrepancies. For logistic reasons, we could not perform a meaningful intra-observer study: the availability of participants during the process of conscription was limited, and we were unable to do repeated ROM testing with a sufficiently long interval between measurements that minimized bias during the second assessment. The examination chair as a device has also several limitations. It is large and may not be conveniently stored in a cabinet for routine clinical use. However, the design can be improved, making the device smaller and possibly foldable. Adjustment of the protractor arm parallel to the longitudinal axis of the lower leg must be done manually, and the reading of the angle between the vertical null position and the protractor is done visually. Even though the ICCs for the inter-observer agreement were excellent and much higher compared with those found for the clinical assessment, automating these processes may reduce residual variation. Therefore, we are currently developing an approach towards measuring IR electronically.

Stretching may increase the passive ROM of IR.[17] To our knowledge, ours is the first study to use a bilateral pulley system to apply a standardized load to each ankle with the aim of achieving maximum IR of the hips in the sitting position after an appropriate adaptation period, which allowed stretching of relevant muscles. As Bierma-Zeinstra *et al*,[15] we did not find a relevant difference in the average ROM assessed in sitting and supine position, but a marked difference in inter-observer variation. In our view, this decrease in variation was mainly related to the belts used to fix pelvis and thighs and the standardized, symmetrical application of a constant load to both legs simultaneously, which minimized compensatory movements and avoided asymmetries in procedures. A limited range of motion is a standard criterion of the American College of Rheumatology[9] in the assessment of hip OA, with IR of $<15^{\circ}$ considered pathological. Future studies should determine whether the examination chair evaluated in our study is also useful in individuals with manifest femoroacetabular impingement or symptomatic hip OA and whether the assessment of IR can be used to identify asymptomatic individuals at risk of developing symptomatic impingement or OA.

We conclude that the use of the evaluated examination chair resulted in a precise assessment of the IR of the hip in our population-based inception cohort study of young asymptomatic males. It was strongly correlated with standard clinical assessment of IR with a goniometer, but was considerably more reliable than the clinical assessment.

Funding/role of sponsor

Supported by the Swiss National Science Foundation's National Research Program 53 on musculoskeletal health (grant no. 405340-104778). SR was a recipient of a research fellowship funded by the Swiss National Science Foundation (grant No PBBEB-115067). PJ was a PROSPER (programme for social medicine, preventive and epidemiological research) fellow funded by the Swiss National Science Foundation (grant No 3233-066377). CTU Bern is supported by the Swiss National Science Foundation. The sponsor had no role in study design, data collection, data analysis, data interpretation, writing of the manuscript, or decision to submit the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank Kathrin Beer for her help in measuring IR, Madeleine Dähler for coordination, Malcom Sturdy for database development, and Nicola Maffiuletti for helpful comments.

References

1. Elders MJ. The increasing impact of arthritis on public health. J Rheumatol Suppl 2000;60:6–8.
2. Lawrence RC, Felson DT, Helmick CG, *et al*. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008;58:26–35.
3. Harris WH. Etiology of osteoarthritis of the hip. Clin Orthop Relat Res 1986;20–33.
4. Flores RH, Hochberg MC. Definition and classification of osteoarthritis. In: Brandt K DM, Lohmander LS, Eds. Osteoarthritis. 2nd edn. Oxford: Oxford University Press; 2003.
5. Ganz R, Parvizi J, Beck M, Leunig M, Notzli H, Siebenrock KA. Femoroacetabular impingement: a cause for osteoarthritis of the hip. Clin Orthop Relat Res 2003;112–20.
6. Ganz R, Leunig M, Leunig-Ganz K, Harris WH. The etiology of osteoarthritis of the hip: an integrated mechanical concept. Clin Orthop Relat Res 2008;466:264–72.
7. Ito K, Minka 2nd MA, Leunig M, Werlen S, Ganz R. Femoroacetabular impingement and the cam-effect. A MRI-based quantitative anatomical study of the femoral head-neck offset. J Bone Joint Surg Br 2001;83:171–6.
8. Beck M, Kalhor M, Leunig M, Ganz R. Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. J Bone Joint Surg Br 2005;87:1012–8.
9. Altman RD. The classification of osteoarthritis. J Rheumatol Suppl 1995; 43:42–3.
10. Preiser G. Statische Gelenkerkrankungen. Stuttgart: Ferdinand Enke Verlag; 1911.
11. Wyss TF, Clark JM, Weishaupt D, Notzli HP. Correlation between internal rotation and bony anatomy in the hip. Clin Orthop Relat Res 2007; 460:152–8.
12. Clohisy JC, Knaus ER, Hunt DM, Lesher JM, Harris-Hayes M, Prather H. Clinical presentation of patients with symptomatic anterior hip impingement. Clin Orthop Relat Res 2009;467:638–44.
13. Sink EL, Gralla J, Ryba A, Dayton M. Clinical presentation of femoroacetabular impingement in adolescents. J Pediatr Orthop 2008;28:806–11.
14. Theiler R, Stucki G, Schutz R, *et al*. Parametric and non-parametric measures in the assessment of knee and hip osteoarthritis: interobserver reliability and correlation with radiology. Osteoarthritis Cartilage 1996;4:35–42.
15. Bierma-Zeinstra SM, Bohnen AM, Ramlal R, Ridderikhoff J, Verhaar JA, Prins A. Comparison between two devices for measuring hip joint motions. Clin Rehabil 1998;12:497–505.
16. Croft PR, Nahit ES, Macfarlane GJ, Silman AJ. Interobserver reliability in measuring flexion, internal rotation, and external rotation of the hip using a plurimeter. Ann Rheum Dis 1996;55:320–3.
17. Aalto TJ, Airaksinen O, Harkonen TM, Arokoski JP. Effect of passive stretch on reproducibility of hip range of motion measurements. Arch Phys Med Rehabil 2005;86:549–57.
18. Pua YH, Wrigley TW, Cowan SM, Bennell KL. Intrarater test-retest reliability of hip range of motion and hip muscle strength measurements in persons with hip osteoarthritis. Arch Phys Med Rehabil 2008;89: 1146–54.
19. National Research Program NRP 53. Musculoskeletal health – chronic pain. Etiology of primary osteoarthritis of the hip, http://www.nfp53.ch/e_module.cfm?Projects. Command = details&get = 13; 2008. Available at[accessed 11.09.09].
20. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
21. Balady GJ, Berra KA, Golding LA. General principles of exercise prescription. In: Franklin BA, Whaley MH, Howley ET, Balady GJ, Eds. ACSM's guidelines for exercise testing and prescription. Lippincott Williams & Wilkins; 2000:137–64.

22. Borms J, Van Roy P, Santens JP, Haentjens A. Optimal duration of static stretching exercises for improvement of coxo-femoral flexibility. J Sports Sci 1987;5:39–47.
23. Cibere J, Thorne A, Bellamy N, *et al*. Reliability of the hip examination in osteoarthritis: effect of standardization. Arthritis Rheum 2008;59: 373–81.
24. Gabbe BJ, Bennell KL, Wajswelnder H, Finch CF. Reliability of common lower extremity musculoskeletal screening tests. Phys Ther Sport 2004;90–7.
25. Bland JM, Altman DG. Measurement error and correlation coefficients. BMJ 1996;313:41–2.
26. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.
27. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101–10.
28. Bland JM, Altman DG. Measurement error. BMJ 1996;313:744.
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.