



Intraexaminer and Interexaminer Reproducibility of the Downing Test for Sacroiliac Joint Evaluation of Symptomatic and Asymptomatic Individuals

Pedro O.P. Lima PT, PhD,^{a,b} Wenya P.X. Melo PT,^b Márcio A. Bezerra PT, PhD,^{a,b} Gabriel P.L. Almeida PT, PhD,^{a,b} Ana Carla L. Nunes PT, PhD,^a and Rodrigo R. Oliveira PT, PhD^{a,b}

ABSTRACT

Objective: This study aimed to assess the intraexaminer and interexaminer reproducibility of the Downing test in sacroiliac joint evaluation in symptomatic and asymptomatic individuals.

Methods: A reliability study was conducted with a test–retest design in 54 college students of both sexes. To assess the intraexaminer reproducibility, each participant was evaluated twice by the same examiner with a 7-day interval, and to assess the interexaminer reproducibility, each participant was evaluated by 2 examiners.

Results: Of the 54 participants included in the study, 18 (33.3%) were asymptomatic and 36 (66.7%) were symptomatic; a total of 108 sacroiliac joints were evaluated. Sacroiliac joint diagnosis based on the Downing test presented low intraexaminer reproducibility in all participants ($\kappa = 0.12$, 95% confidence interval [CI] 0.03-0.22), in asymptomatic individuals ($\kappa = 0.18$, 95% CI 0.02-0.34), and in symptomatic individuals ($\kappa = 0.28$, 95% CI 0.17-0.39). The interexaminer reproducibility also was low in all participants ($\kappa = 0.18$, 95% CI 0.09-0.27), in asymptomatic individuals ($\kappa = 0.22$, 95% CI 0.15-0.37), and in symptomatic individuals ($\kappa = 0.16$, 95% CI 0.05-0.27). The standard error of the measurement values were not lower than smallest detectable change values considering a CI of 95% for all participants.

Conclusion: For this group of asymptomatic and symptomatic participants, the reproducibility of the Downing test was poor. The clinical utility of this test used in isolation is not supported by the present study. (*J Chiropr Med* 2019;18:163-170)

Key Indexing Terms: *Physical Examination; Range of Motion, Articular; Hip Joint; Sacroiliac Joint; Low Back Pain*

INTRODUCTION

Low back pain is a burden on public health, with approximately 83 million years of life lost every year.¹ (Years of life lost are calculated from the number of deaths multiplied by a standard life expectancy at the age at which death occurs.) In highly developed countries, low back pain is one of the major musculoskeletal sources of disability affecting quality of life and family and social relationships and impairs the ability to work.² The sacroiliac joint (SIJ) is

a common source of pain in the 10% to 27% of patients who present low back pain.³

The gold standard to diagnose pain generated by the SIJ is the invasive technique of infiltration with local anesthetics; when there is a decrease in complaints after infiltration indicates an SIJ injury. This technique requires specific and costly procedures that are not commonly available in clinical practice. As well, SIJ mobility is not evaluated with this procedure, thus it may not be useful to measure articular function.⁴

A few studies were conducted to identify SIJ dysfunction and several clinical procedures have been used indirectly to assess different constructs, such as motion tests to assess mobility, pain provocation tests to stress SIJ structures, and palpation tests for pelvic position.^{5,6} The key pain provocation tests are distraction, compression, thigh thrust, sacral thrust, and Gaenslen test. Current evidence supports that a combination of 3 or more positive pain provocation SIJ tests have sensitivity and specificity of 91% and 78%, respectively.^{7,8} Diagnostic accuracy of commonly used palpation tests for SIJ position or motion

^a Department of Physical Therapy, Federal University of Ceará, Fortaleza, Ceará, CE, Brazil.

^b League of Sport Physical Therapy, Federal University of Ceará, Fortaleza, Ceará, CE, Brazil.

Corresponding author: Pedro O. P. Lima, PT, PhD, Major Weyne Street, 1440 Fortaleza, CE, 60430-110, Brazil. Tel.: +55 85 33668632. (e-mail: pedrolima@ufc.br).

Paper submitted September 19, 2017; in revised form November 22, 2018; accepted November 27, 2018.

1556-3707

© 2019 National University of Health Sciences.

<https://doi.org/10.1016/j.jcm.2018.11.007>

was also evaluated, and the sensitivity/specificity values of all tests were poor, namely Gillet test, standing and sitting flexion, posterior superior iliac spine position, and prone knee flexion tests. This result was an expected finding given that the reference standard was related to SIJ pain and not dysfunction.⁹⁻¹¹ Another procedure used to assess SIJ mobility is the Downing test, which was designed to provoke tension forces in ligaments to rotate the iliac bone (anterior or posterior) and to cause temporary misalignment of the lower limbs (ie, apparent leg length discrepancy). According to theories originating from osteopathic medicine, lower limb misalignment can be generated by inadequate ligament tension, which may cause asymmetry in the SIJ, spine, and pelvis, inducing a biomechanical disorder.¹²⁻¹⁴ Although the Downing test is used in clinical practice and was described since 1940, its reliability and agreement remain unknown so far.¹⁵

Therefore, this study aimed to verify the intraexaminer and interexaminer reproducibility of the Downing test in SIJ evaluation in symptomatic and asymptomatic individuals. Our hypothesis was that this test does not have sufficient reliability to support its solo use in clinical practice, similar to the other tests for SIJ mobility.

METHODS

Design and Participants

A reliability study was conducted with a test–retest design and we reported the findings according to Guidelines for Reporting Reliability and Agreement Studies. Data were collected between January and March 2013 in the Human Movement Analysis Lab of the Federal University of Ceará. This study was approved by the human research ethics committees of the Federal University of Ceará with protocol number 96/787.

The sample of participants was selected from college students of both sexes. The required sample size was estimated based upon the recommendations of specific guidelines on clinometric properties that suggest a sample of at least 50 participants for reproducibility studies.¹⁶ We included participants with and without symptoms of low back pain or SIJ pain. We excluded participants who had acute low back pain (ie, pain onset within last 6 weeks) who were pregnant, had previous abdominal wall or spinal surgery, had a body mass index above 25, were menstruating during the tests, and had suspected or confirmed serious pathologies (ie, nerve root compromise, fractures, cancer, infectious diseases of the spine, cauda equina syndrome, and widespread neurologic disorders). Participants who missed consecutive tests were also excluded.

Procedures

Measurements were performed by 2 physiotherapists to test the interexaminer reproducibility of the Downing test.

The test–retest interval was approximately 10 minutes, and the order of examiners was randomized. To avoid exchanging information between the examiners, none of them were present during the measures performed by the other examiner. For the study of intraexaminer reproducibility, one of the examiners conducted 2 measurements of the Downing test on 2 different occasions, with an interval of 7 days to reduce recall. Examiners A and B were physiotherapists with extensive experience using the Downing test, and they received additional training to optimize the standardization of the tests. We have adopted the same clinical, environmental, and temporal conditions to avoid external influences or internal errors during the period of data collection.

The application of the Downing test followed a pre-established sequence (Fig 1)^{15,17}:

1. Normalization of SIJ by performing bridge exercise;
2. Determination of reference markings on both legs;
3. Examination of the apparently shorter limb using the medial malleolus as a reference;
4. Perform the hip abduction with internal rotation maneuver on the apparently shorter limb (“limb shortening”) and check the effect of the maneuver;
5. Normalization of SIJ, as in item 1;
6. Perform the hip adduction with external rotation maneuver on the apparently longer limb (“limb lengthening”) and check the effect of the maneuver;
7. Normalization of SIJ, lengthening of the previously shortened limb, and check the effect of the maneuver;
8. Normalization of SIJ, shortening of the previously lengthening limb, and check the effect of the maneuver; and
9. Normalization of SIJ.

More details about the test’s description can be found in the [appendix](#).

The reference markings were used for interpretation of the test. To verify the effect of the limb “shortening” or “lengthening” maneuvers, the previously performed marking on the contralateral limb was used as a reference. Distances between the reference markings were measured with a digital caliper (Lee) by both examiners. The distance between lines 1 and 2 represents the degree of limb “shortening”; if it is equal to zero, the ipsilateral SIJ is no mobility to posterior rotation. Distance between lines 2 and 3 represents the degree of limb “lengthening”; if it is equal to zero, the ipsilateral SIJ is no mobility to anterior rotation. An SIJ diagnosis was determined by each examiner based on distance measures.

Pain intensity was measured with the Pain Numerical Rating Scale, and disability was measured with the Roland Morris Disability Questionnaire.^{18,19} These outcome measures were administered on the first day, after to the Downing test, and the sample was divided into symptomatic (low back pain or SIJ pain) and asymptomatic individuals. Consequently, the examiners were blinded to the group of each individual.

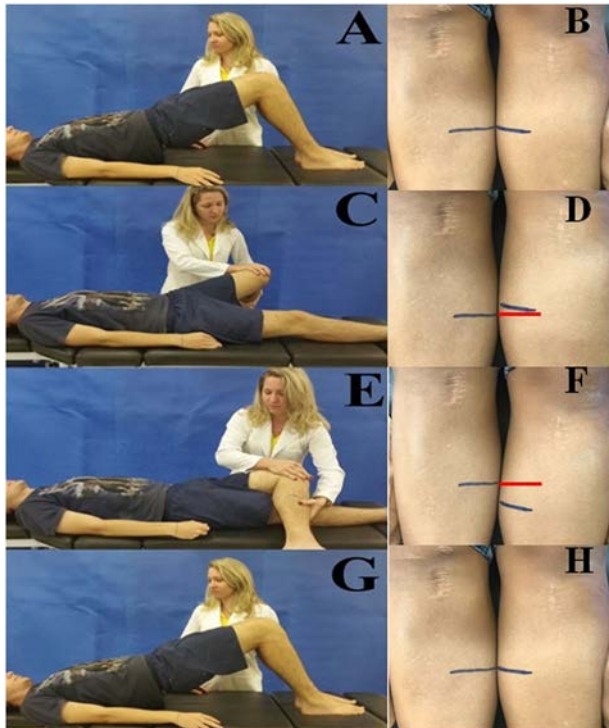


Fig 1. Representation of Downing test steps. (A) Normalization of the sacroiliac joint, (B) determination of reference markings, (C) “limb shortening” induced by hip abduction with internal rotation, (D) effect of limb shortening when compared with contralateral limb marking, (E) “limb lengthening” induced by hip adduction with external rotation, (F) effect of limb lengthening when compared with contralateral limb marking, (G) normalization of the sacroiliac joint, and (H) reference markings for test interpretation.

Statistical Analysis

The sample characteristics were presented using descriptive statistics. We used the intraclass correlation coefficient ($ICC_{2,1}$) to describe the reliability of continuous data (distances between the reference markings) and coefficient of kappa (κ) for categorical data (SIJ diagnosis). The ICC values lower than 0.4 can be classified as poor,

between 0.4 and 0.7 can be classified as satisfactory, and over 0.7 can be classified as excellent.²⁰ The κ values less than 0 mean there is no reliability, and from 0 to 0.19 indicate poor reliability, 0.20 to 0.39 fair reliability, 0.40 to 0.59 moderate reliability, 0.60 to 0.79 good reliability, and 0.8 to 1 excellent reliability.²¹

We used 3 measures of agreement: the Bland-Altman plots, the standard error of the measurement (SEM), and the smallest detectable change (SDC).²² The SEM was calculated by multiplying the standard deviation of the mean differences between the 2 measurements by the square root of 1 minus ICC (standard deviation differences $\times \sqrt{1 - ICC}$), and the SDC was calculated using the formula $SDC = 1.96 \times \sqrt{2} \times SEM$. The SEM reflects the absolute error of the instrument, and the SDC reflects the smallest within-person change in a score that can be interpreted as a real change, above the measurement error 1 of an individual.²⁰ Data were analyzed with Statistical Package for Social Sciences software version 17.0 (SPSS Inc, Chicago, Illinois) with $\alpha = 0.05$.

RESULTS

Seventy-eight participants were invited to participate in this study, but 24 were excluded owing to eligibility criteria. The sample consisted of 54 participants: 49 (90.7%) were women and 36 (66.7%) were symptomatic; and 108 lower limbs were evaluated in total. The symptomatic group presented low levels of pain and disability. The baseline characteristics of the participants are presented in Table 1.

Sacroiliac joint mobility based on distances between lines 1 and 2 (posterior rotation) and lines 2 and 3 (anterior rotation) showed poor intraexaminer and interexaminer reliability for all participants, asymptomatic and symptomatic individuals (Table 2). Sacroiliac joint diagnosis based on the Downing test presented poor intraexaminer reliability values for all participants and the asymptomatic group and considerable values for the symptomatic group. Interexaminer reliability values were

Table 1. Characteristics of the Sample

Variables	All (n = 108)	Asymptomatic (n = 36)	Symptomatic (n = 72)
N (%)	54 (100.0)	18 (33.3)	36 (66.7)
Age (y)	21.26 \pm 3.47	21.11 \pm 4.10	21.33 \pm 3.18
Weight (kg)	62.39 \pm 9.18	61.66 \pm 11.08	62.75 \pm 8.22
Height (m)	1.64 \pm 0.06	1.64 \pm 0.05	1.64 \pm 0.06
BMI (kg/m ²)	22.97 \pm 2.58	22.64 \pm 3.03	23.13 \pm 2.3
Pain intensity (points)	2.98 \pm 3.14	0	4.47 \pm 2.8
Disability (points)	3.33 \pm 3.95	0.61 \pm 0.59	4.69 \pm 4.1

NOTE. Continuous data are shown in mean \pm standard deviation. BMI, body mass index.

Table 2. Intraexaminer and Interexaminer Reliability of the Downing Test Based on Distances Between Reference Markings Measured by Digital Caliper

Reliability	All (n = 108)	Asymptomatic (n = 36)	Symptomatic (n = 72)
Intraexaminer reliability			
Distance between lines 1 and 2	0.19 (0.01-0.36)	0.15 (0.01-0.57)	0.19 (0.01-0.48)
Distance between lines 2 and 3	0.36 (0.19-0.52)	0.20 (0.01-0.61)	0.44 (0.13-0.67)
Interexaminer reliability			
Distance between lines 1 and 2	0.30 (0.10-0.47)	0.32 (0.02-0.58)	0.29 (0.05-0.50)
Distance between lines 2 and 3	0.37 (0.19-0.53)	0.41 (0.11 - 0.64)	0.35 (0.13-0.53)

NOTE. Intraclass correlation coefficient (95% confidence interval).

also poor for all participants and the symptomatic group, and considerable values were found for the asymptomatic group (Table 3).

The limits of agreement for the distance between lines 1 and 2 (limb “shortening” maneuver) for the intraexaminer agreement ranged from -6.56 to 7.00 mm and for the interexaminer agreement ranged from -4.46 to 7.49 mm. The limits of agreement for the distance between lines 2 and 3 (limb “lengthening” maneuver) for the intraexaminer agreement ranged from -9.49 to 8.85 mm and for the interexaminer agreement ranged from -6.07 to 8.96 mm, as shown by the Bland-Altman plots (Fig 2). The SEM values were not lower than SDC values considering a CI of 95% for all participants in both maneuvers of the Downing test (Table 4).

DISCUSSION

The present study was the first to evaluate the intraexaminer and interexaminer reproducibility of the Downing test, a procedure that has not yet been investigated in the scientific literature. This study showed poor intraexaminer and interexaminer reproducibility of the Downing test in both asymptomatic and symptomatic individuals. The process of diagnosis and treatment of individuals with low back pain can include evaluation of the sacroiliac joint. For individuals to recover normal mobility, the pain should be reduced and the function increased.²³ Assessment methods of joint mobility should show that the data were accurately evaluated so that the measure can be useful in research and the clinical setting.²⁴

Reproducible clinical measures are essential in diagnosis, but also in clear interdisciplinary communication and in obtaining accurate results. A clinically useful tool must be reliable and allow that different examiners obtain similar results.^{25,26}

Many studies have evaluated the reproducibility of pain provocation tests and mobility tests of the SIJ; but the intraexaminer and interexaminer reproducibility frequently ranged from weak to moderate in both cases.^{11,25,27-31} In a recent study, the reproducibility of diagnostic tests for SIJ, disc, and facet pain was evaluated and the values found were low, which compromises the accuracy diagnostic in clinical trials.³²

The Downing test is widely used in clinical practice for the diagnosis of mobility impairment in SIJ; however, there were no studies conducted on its reproducibility. Unlike the Gillet test and standing flexion test, the Downing test does not require palpation of specific anatomic structures as reference point, which could contribute to a more reliable diagnostic method.^{11,15} However, the Downing test did not reach the minimum of 80% agreement required to be considered reproducible in this study. This was true for both symptomatic and asymptomatic individuals.

Interpretation of assessor-dependent tests is complex and involves several factors, such as the strength and range of motion performed by the examiner during the “shortening” and “lengthening” maneuvers of the Downing test, the accuracy in the reference markings, and the patient’s relaxation. In general, the reproducibility of pain provocation tests tend to outweigh mobility tests because the presence of pain is a clearer outcome for the examiner.²⁹

Table 3. Intraexaminer and Interexaminer Reliability of the Downing Test Based on SIJ Diagnosis

Reliability	All (n = 108)	Asymptomatic (n = 36)	Symptomatic (n = 72)
Intraexaminer reliability			
	0.12 (0.03-0.22)	0.18 (0.02-0.34)	0.28 (0.17-0.39)
Interexaminer reliability			
	0.18 (0.09-0.27)	0.22 (0.15-0.37)	0.16 (0.05-0.27)

κ’s coefficient (95% confidence interval).
 SIJ, sacroiliac joint.

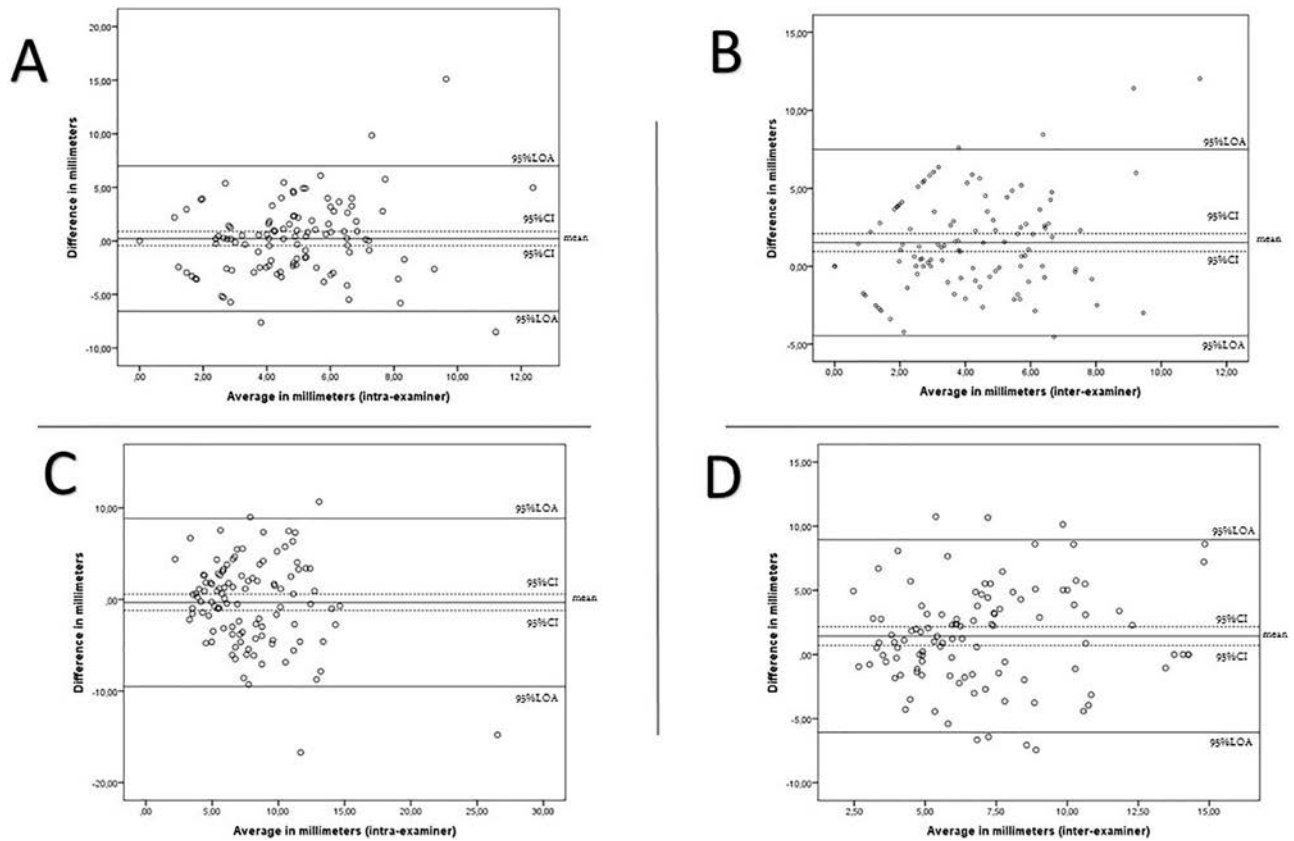


Fig 2. Bland-Altman plots of Downing test for all participants ($n = 108$). (A) Distance between lines 1 and 2 (limb “shortening” maneuver) for intraexaminer agreement, (B) distance between lines 1 and 2 (limb shortening maneuver) for interexaminer agreement, (C) distance between lines 2 and 3 (limb “lengthening” maneuver) for intraexaminer agreement, and (D) distance between lines 2 and 3 (limb lengthening maneuver) for interexaminer agreement. CI, confidence interval; LOA, limits of agreement.

The combination of pain provocation tests has been shown to increase the diagnostic probability. However, the poor reliability of diagnostic parameters (from physical examination to test execution) has been demonstrated in several studies.^{32,33} The problem is that there is no widely accepted reference standard for SIJ dysfunction. Studies showed low values of sensitivity and specificity of the Gillet test, standing flexion test, and motion demand spring

test and yet found a high prevalence of positive tests in asymptomatic individuals.⁸

Limitations

Despite the standardization of the test steps and training of the evaluators, it is possible that the results reflect differences of execution between examiners owing to

Table 4. SEM and SDC of the Downing Test Based on Distances Between Reference Markings Measured by Digital Caliper

All (n = 108)	SEM	SDC (95% CI)
Intraexaminer agreement		
Distance between lines 1 and 2 (mm)	3.12	8.63 (2.52-14.75)
Distance between lines 2 and 3 (mm)	3.74	10.38 (3.04-17.71)
Interexaminer agreement		
Distance between lines 1 and 2 (mm)	2.55	7.07 (2.07-12.07)
Distance between lines 2 and 3 (mm)	3.04	8.43 (2.47-14.39)

CI, confidence interval; SDC, smallest detectable change; SEM, standard error of the measurement.

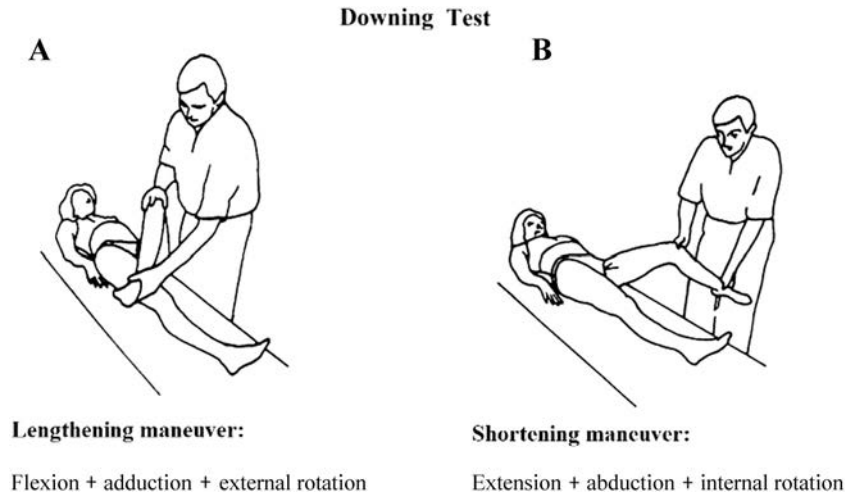


Fig 3. Downing test (Courtesy of Klaus Hösele).

differences in skill, experience, and manual practice. Also, participants' level of disability was low and with moderate pain intensity, which can be a limitation and should be carefully interpreted. Other limitations include that the study was on a limited group of participants and that this test was evaluated in isolation and not in combination with any other tests or procedures.

CONCLUSION

For this group of asymptomatic and symptomatic participants, the reproducibility of the Downing test was poor. The clinical utility of this test used in isolation is not supported by the present study.

FUNDING SOURCES AND CONFLICTS OF INTEREST

No funding sources or conflicts of interest were reported for this study.

CONTRIBUTORSHIP INFORMATION

Concept development (provided idea for the research): P.O.P.L., W.P.X.M., R.R.O.

Design (planned the methods to generate the results): P.O.P.L., W.P.X.M., R.R.O.

Supervision (provided oversight, responsible for organization and implementation, writing of the manuscript): P.O.P.L., W.P.X.M., R.R.O.

Data collection/processing (responsible for experiments, patient management, organization, or reporting data): P.O.P.L., W.P.X.M., R.R.O.

Analysis/interpretation (responsible for statistical analysis, evaluation, and presentation of the results): P.O.P.L., W.P.X.M., M.A.B., G.P.L.A., R.R.O.

Literature search (performed the literature search): P.O.P.L., W.P.X.M., M.A.B., R.R.O.

Writing (responsible for writing a substantive part of the manuscript): P.O.P.L., W.P.X.M., M.A.B., G.P.L.A., A.C.L.N., R.R.O.

Critical review (revised manuscript for intellectual content, this does not relate to spelling and grammar checking): P.O.P.L., W.P.X.M., M.A.B., G.P.L.A., A.C.L.N., R.R.O.

APPENDIX

The patient is supine, comfortably. The examiner ensures that the participant is lying in a perfectly symmetrical position. He or she releases the muscular tension of the lower limbs with small movements of internal and external rotation of the legs and asks the participant to take off his hips from the table and then let himself or herself fall gently. Then the examiner checks the respective height of the internal malleolus, which allows him or her to appreciate a possible difference in length of the legs.

He will then execute the maneuvers.

Anteriorization Maneuver

The examiner is placed on the other side to leg to be tested, that is, on the right for the left leg and vice versa. Passively, he performs the leg adduction, then maximum external rotation to perceive resistance to movement (Fig 3A). Then the examiner releases the adduction by maintaining the external rotation and releases the external rotation at the last moment just before resting the leg in extension

(Fig 3B). The examiner verifies the lengthening of the lower limb, which can be 5 to 20 mm depending on the ligament laxity of the participant. The maneuver is made under the same conditions on the other side.

If everything is normal after these 2 maneuvers, the internal malleoli are in the same situation as at the beginning. If we note a difference, that is, a malleolus higher than the other, we can conclude to a posterior ilium, the high side, which therefore has a resistance to lengthening (ie, anteriorization).

Posteriorization Maneuver

This time the examiner is on the same side and wears the thigh in maximum abduction, then in internal rotation until feeling the resistance. Then, he releases the abduction while maintaining the internal rotation, which is relaxed just before resting the leg in extension. The examiner checks the obtained shortening (from 5 to 20 mm). Then he does the maneuver on the other side and compares.

If a malleolus is lower than it was before, we can conclude for an anterior ilium of the low side.

Figure 3 shows positioning for the Downing test.

Practical Applications

- The Downing test has been used to evaluate mobility impairment in SIJ.
- This study found that the intraexaminer and interexaminer reproducibility of the Downing test is poor.

REFERENCES

1. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2013;380(9859):2197-2223.
2. Cher D, Polly D, Berven S. Sacroiliac joint pain: burden of disease. *Med Devices (Auckl)*. 2014;7:73.
3. Manchikanti L, Abdi S, Atluri S, et al. An update of comprehensive evidence-based guidelines for interventional techniques in chronic spinal pain. Part II: guidance and recommendations. *Pain Physician*. 2013;16(2 suppl):S49-283.
4. Szadek KM, van der Wurff P, van Tulder MW, Zuurmond WW, Perez RS. Diagnostic validity of criteria for sacroiliac joint pain: a systematic review. *J Pain*. 2009;10(4):354-368.
5. Cooperstein R, Morschhauser E, Lisi A, Nick TG. Validity of compressive leg checking in measuring artificial leg-length inequality. *J Manipulative Physiol Ther*. 2003;26(9):557-566.
6. Cooperstein R, Hickey M. The reliability of palpating the posterior superior iliac spine: a systematic review. *J Can Chiropr Assoc*. 2016;60(1):36-46.
7. Riddle DL, Freburger JK. Evaluation of the presence of sacroiliac joint region dysfunction using a combination of tests: a multicenter intertester reliability study. *Phys Ther*. 2002;82(8):772-781.
8. Laslett M. Evidence-based diagnosis and treatment of the painful sacroiliac joint. *J Man Manipulative Ther*. 2008;16(3):142-152.
9. Dreyfuss P, Dreyer S, Griffin J, Hoffman J, Walsh N. Positive sacroiliac screening tests in asymptomatic adults. *Spine*. 1994;19(10):1138-1143.
10. Dreyfuss P, Michaelsen M, Pauza K, McLarty J, Bogduk N. The value of medical history and physical examination in diagnosing sacroiliac joint pain. *Spine (Phila Pa 1976)*. 1996;21(22):2594-2602.
11. Aufdemkampe G. Intraexaminer and interexaminer reliability of the Gillet test. *J Manipulative Physiol Ther*. 1999;22(1):4-9.
12. Defrin R, Benyamin SB, Aldubi RD, Pick CG. Conservative correction of leg-length discrepancies of 10mm or less for the relief of chronic low back pain. *Arch Phys Med Rehabil*. 2005;86(11):2075-2080.
13. Grace S, Orrock P, Vaughan B, Blaich R, Coutts R. Understanding clinical reasoning in osteopathy: a qualitative research approach. *Chiropr Man Therap*. 2016;24:6.
14. Thomson OP, Petty NJ, Moore AP. Clinical decision-making and therapeutic approaches in osteopathy – a qualitative grounded theory study. *Man Ther*. 2014;19(1):44-51.
15. Duffield HJ. Manipulation of the sacro-iliac joint. *Ulster Med J*. 1940;9(2):108-114.
16. 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(1):34-42.
17. Bienfait M. *Bases Elementares Técnicas de Terapia Manual e Osteopatia*. São Paulo, Brazil: Summus Editorial; 1997.
18. Nusbaum L, Natour J, Ferraz MB, Goldenberg J. Translation, adaptation and validation of the Roland-Morris questionnaire – Brazil Roland-Morris. *Braz J Med Biol Res*. 2001;34(2):203-210.
19. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine (Phila Pa 1976)*. 2005;30(11):1331-1334.
20. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033-1039.
21. Altman D. *Practical Statistics for Medical Research*. London, England: Chapman and Hall; 1991.
22. Francq BG, Govaerts B. How to regress and predict in a Bland-Altman plot? Review and contribution based on tolerance intervals and correlated-errors-in-variables models. *Stat Med*. 2016;35(14):2328-2358.
23. Elgueta-Cancino E, Schabrun S, Danneels L, Hodges P. A clinical test of lumbopelvic control: development and reliability of a clinical test of dissociation of lumbopelvic and thoracolumbar motion. *Man Ther*. 2014;19(5):418-424.
24. Lima PO, de Oliveira RR, de Moura Filho AG, Raposo MC, Costa LO, Laurentino GE. Reproducibility of the pressure biofeedback unit in measuring transversus abdominis muscle activity in patients with chronic nonspecific low back pain. *J Bodyw Mov Ther*. 2012;16(2):251-257.
25. Gibbons P, Dumper C, Gosling C. Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment. *J Osteopath Med*. 2002;5(2):53-58.
26. Chiarotto A, Ostelo RW, Turk DC, Buchbinder R, Boers M. Core outcome sets for research and clinical practice. *Braz J Phys Ther*. 2017;21(2):77-84.

27. Arab AM, Abdollahi I, Joghataei MT, Golafshani Z, Kazemnejad A. Inter-and intra-examiner reliability of single and composites of selected motion palpation and pain provocation tests for sacroiliac joint. *Man Ther.* 2009;14(2):213-221.
28. Bemis T, Daniel M. Validation of the Long Sitting Test on subjects with iliosacral dysfunction. *J Orthop Sports Phys Ther.* 1987;8(7):336-345.
29. Laslett M, Aprill CN, McDonald B, Young SB. Diagnosis of sacroiliac joint pain: validity of individual provocation tests and composites of tests. *Man Ther.* 2005;10(3):207-218.
30. Schneider M, Homonai R, Moreland B, Delitto A. Interexaminer reliability of the prone leg length analysis procedure. *J Manipulative Physiol Ther.* 2007;30(7):514-521.
31. Vincent-Smith B, Gibbons P. Inter-examiner and intra-examiner reliability of the standing flexion test. *Man Ther.* 1999;4(2):87-93.
32. van Tilburg CW, Groeneweg JG, Stronks DL, Huygen FJ. Inter-rater reliability of diagnostic criteria for sacroiliac joint-, disc- and facet joint pain. *J Back Musculoskelet Rehabil.* 2017;30(3):551-557.
33. Szadek KM, van der Wurff P, van Tulder MW, Zuurmond WW, Perez RS. Diagnostic validity of criteria for sacroiliac joint pain: a systematic review. *J Pain.* 2009;10(4):354-368.