


Effect of clinician's experience and expertise on the inter- and intra-observer reliability of hip migration index in children with cerebral palsy

A STROBE-compliant retrospective study

Necdet Demir, MD^a, Mehmet Demirel, MD^{b,*} , Önder Turna, MD^c, Derya Yıldızlar, MD^d, Önder Demirbaş, MD^c, Yavuz Sağlam, MD^b

Abstract

Few studies have investigated the reliability of Reimers' hip migration percentage (RMP) in children with cerebral palsy (CP). Most studies on the topic reflected rating results of physician with a similar level of experience from the same expertise. This study aimed to determine the effect of clinician's experience and expertise on the intra- and inter-observer reliability of RMP.

In this retrospective observational study, hip radiographs of children with CP were identified. 5 observers with different degrees of working experience from 3 different clinical fields, including orthopedics, radiology, and physical medicine and rehabilitation, performed all RMP measurements. All measurements were repeated 6 weeks later by the same observers. Inter- and intra-observer reliability for RMP measurements were assessed using Intraclass Correlation Coefficient (ICC), calculated from 2 sets of repeated measurements on a subset of 50 hips, with a 6 week apart for each observer.

Fifty hip radiographs of 25 children with CP (10 females and 15 males; mean age: 6 years; age range: 2–8 years) were examined in the current study. No significant differences existed in intra- and inter-observer measurements. Excellent intra-observer reliability was observed between the 2 separate measurements for each observer, with a mean ICC of 0.976 (range: 0.956–0.989; $P < .001$). Among 5 observers, inter-observer reliability was excellent for the 2 separate RMP measurements, with the mean ICC minimally increasing between the 2 measurement periods (mean ICC: 0.928, range: 0.838–0.979 and mean ICC: 0.936, range: 0.861–0.983, respectively) ($P < .001$).

Clinician's experience and expertise may not affect inter- and intra-observer reliability of RMP measurements.

Abbreviations: CP = cerebral palsy, HL = hilgenreiner line, ICC = Intraclass Correlation Coefficient, RMP = Reimers' hip migration percentage.

Keywords: cerebral palsy, clinician' experience, expertise, hip displacement, hip migration index, Reimer's migration percentage, reliability

1. Introduction

Hip displacement in cerebral palsy (CP) represents a frequent and serious problem that significantly affects functional status and quality of life.^[1] Nonetheless, a careful clinical and radiographic assessment can identify the children with CP at risk for hip

displacement and thus prevent future associated complications such as progressive subluxation, dislocation, femoral head deformity, and painful secondary osteoarthritis.^[2–4]

In this sense, the Reimers' hip migration percentage (RMP), which defines the migration percentage of the femoral head to the

Editor: Johannes Mayr.

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

This study was reviewed and approved by our institutional review board (Ref. number 2017/10-10; approval issue date November 27, 2017) and informed consent was obtained from all individual participants included in this study.

The authors have no conflicts of interests to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

^a Department of Orthopaedics and Traumatology, Biruni University, School of Medicine, ^b Department of Orthopaedics and Traumatology, Istanbul University, Istanbul School of Medicine, ^c Department of Radiology, ^d Department of Physical Therapy and Rehabilitation, Biruni University, School of Medicine, Istanbul, Turkey.

* Correspondence: Mehmet Demirel, Department of Orthopaedics and Traumatology, Istanbul University, Istanbul School of Medicine, Istanbul 34104, Turkey (e-mail: dr88.mehmet.demirel@gmail.com).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Demir N, Demirel M, Turna Ö, Yıldızlar D, Demirbaş Ö, Sağlam Y. Effect of clinician's experience and expertise on the inter- and intra-observer reliability of hip migration index in children with cerebral palsy: a STROBE-compliant retrospective study. *Medicine* 2021;100:10(e24538).

Received: 14 May 2020 / Received in final form: 26 November 2020 / Accepted: 11 January 2021

<http://dx.doi.org/10.1097/MD.00000000000024538>

acetabulum, is extensively accepted as the gold-standard measure to identify hip displacement in CP.^[5] Nonetheless, few studies have investigated the reliability of this radiographic measure to date.^[5–8] Moreover, according to our review of the literature, most of those reflected rating results of physician with a similar level of experience from the same expertise, and thus neglecting the effect of clinician's experience and expertise on the reliability.

The aim of this study was to determine the intra- and inter-observer reliability of RMP in children with CP based on 5 observers with different levels of experience in performing the measurement from 3 different clinical specialties. We hypothesized that the clinician's experience and expertise both have significant effects on the reliability of RMP measurements.

2. Materials and methods

2.1. Data collection and setting of the study

This retrospective observational study was conducted on children with CP who were diagnosed between 2010 and 2015 and underwent plain hip radiographs at the departments of Orthopedics and Traumatology in a single tertiary referral center. Inclusion criteria for the study were:

1. a diagnosis of CP;
2. an age of between 2 and 8 years;
3. Gross Motor Functional Classification System level II to IV;
4. complete medical records and radiographic images stored in the hospital Picture Archiving and Communication System (Picture Archiving and Communication System, Centricity Universal Viewer, GE Healthcare, Milwaukee); and
5. being willing to participate in the study.

Exclusion criteria were:

1. children who underwent any reconstructive or salvage surgeries of the hip or pelvis;
2. lost to follow-up;
3. inadequate radiographic imaging;
4. concomitant hip disorders such as congenital coxa vara, developmental dysplasia of the hip, Legg-Calve-Perthes disease, avascular necrosis of the femoral head, proximal femoral focal deficiency;

5. concomitant metabolic bone disorders and skeletal dysplasias; and
6. being unwilling to participate the study.

2.2. Participants

A total of 33 children were evaluated based on the above eligibility criteria. After excluding 8 children (2 underwent reconstructive surgery, 3 were lost to follow-up, 1 had inadequate radiographic imaging, and 2 had concomitant hip disorders), remaining 50 hips of 25 children with CP who met the inclusion criteria were enrolled in the study and invited to a final follow-up examination for radiographic assessment. Parents were informed that medical records could be used for scientific purposes only; thus, written informed consent was obtained at the final visit. Approval from the institutional review board of Biruni University, School of Medicine (Ref. number 2017/10-10; approval issue date November 27, 2017) was obtained before the acquisition of written consent and enrollment of participants.

2.3. Measurement of RMP

Children were assisted to have a supine position and render their pelvis and femurs in a neutral position with patellae facing upward, in order to obtain a standard anteroposterior pelvis radiograph (Fig. 1).^[9] If necessary, the pelvis was elevated by placing a pillow under the knees to avoid anterior pelvic tilt (Fig. 2).

The primary outcomes of the study were the intra- and inter-observer reliability of RMP measurements in children with CP based on 5 observers with different levels of experience from different clinical specialties. Therefore, all radiographs were checked and confirmed to be in compliance with the standard radiographic examination by an independent experienced orthopedic surgeon. Then, 5 independent observers with different degrees of working experience from 3 different clinical fields performed all RMP measurements:

- *Observer 1*: an orthopedic resident with 2 years of experience
- *Observer 2*: an orthopedic surgeon with 8 years of experience
- *Observer 3*: a radiology specialist with 3 years of experience
- *Observer 4*: a radiology specialist with 7 years of experience
- *Observer 5*: a physiatrist with 10 years of experience.



Figure 1. Standard patient positioning used for anteroposterior pelvic radiographs of children with cerebral palsy. The lower extremities should be placed parallel to each other (neutral adduction/abduction) while the patellae facing upwards.



Figure 2. The pelvis should be level and flat. If there is a fixed flexion deformity giving rise to a lordosis, the lower extremities should be elevated to flatten the lumbar lordosis.

Afterward, all measurements were repeated 6 weeks later by the same observers, in order to test the intra-observer reliability (Fig. 3). Additionally, each observer was trained in the standardized measurement protocol, including adequate pelvic radiograph, proper patient positioning, and measurement method.

To minimize measurement errors, the hospital Picture Archiving and Communication System was used for review and measurement of images. To calculate RMP, the Hilgenreiner line (HL), which connects the Y-cartilage of both sides, was first drawn, and then Perkin's line was drawn perpendicular to the HL along the outermost edge of the acetabular roof. Next, the outermost and innermost margins of the femoral head were

determined, and 2 lines crossing these outer and inner lines were drawn by intercepting at an angle of 90° to the HL. Eventually, the diameter of the femoral head staying outside the outer line, and the total diameter of the femoral head between the 2 lines were measured in mm. RMP was estimated as the ratio of the outer acetabular part of the femoral head to the total diameter of the femoral head (Fig. 4).

2.4. Statistical analysis

All statistical analyses were conducted using SPSS software (version 20.0. Armonk, NY: IBM Corp.). A P value $<.05$ was

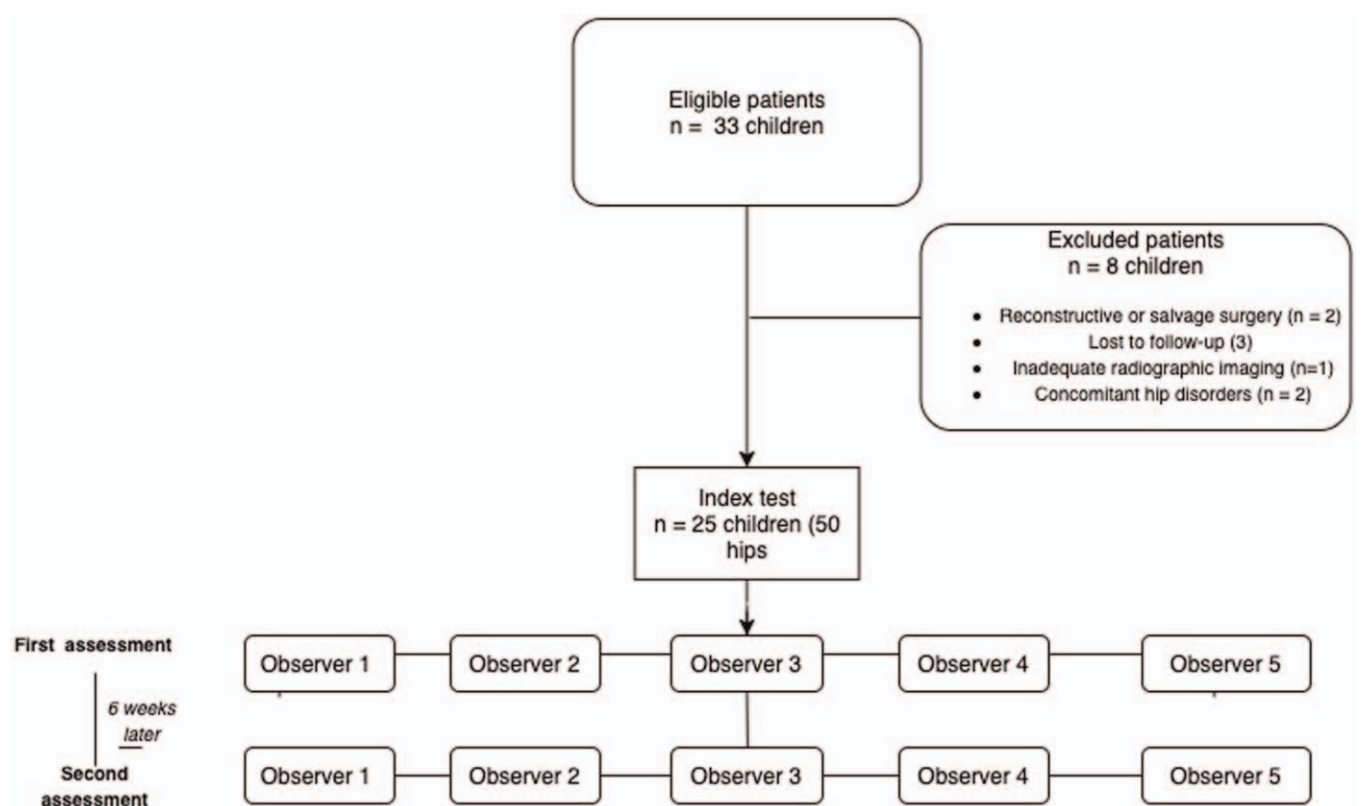


Figure 3. The flow chart of the study design.

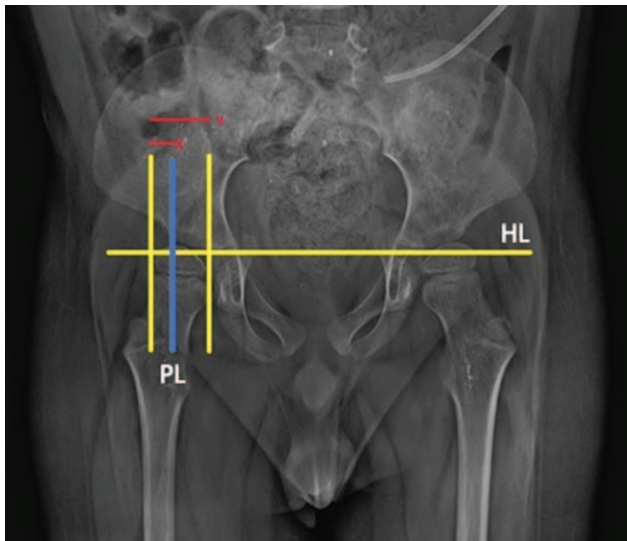


Figure 4. Measurement of Reimers' hip migration percentage (RMP): HL, Hilgenreiner's horizontal line connecting both triradiate cartilages. PL, Perkin's line (blue line) drawn perpendicular to the HL across the lateral margin of the acetabulum. RMP represents the ratio of the femoral head migrated laterally beyond PL. $RMP = X/Y \times 100\%$.

considered significant. For the estimation of inter-observer reliability, 50 hips of 25 children per each observer were sufficient to estimate an intraclass correlation coefficient of 0.80, given a minimum acceptable level of reliability of 0.60.^[10] Comparison of means between 2 observers was undertaken using Mann–Whitney *U* test. Comparison of means among 5 observers was done with non-normally distributed data using the Kruskal–Wallis test. Inter- and intra-observer reliability for RMI measurements were assessed using Intraclass Correlation Coefficient (ICC), calculated from 2 sets of repeated measurements on a subset of 50 hips, with a 6 week apart for each observer. Agreement was considered excellent for $ICC > 0.80$, very good for 0.70 to 0.80, good for 0.60 to 0.70, fair for 0.40 to 0.60, and poor for < 0.40 .

The quality of the current study was assessed using criteria from the STROBE checklist for observational studies.^[11]

3. Results

Fifty hips of 25 children with CP were retrospectively reviewed in the current study. There were 10 females and 15 males. The mean age was 6 years (range: 2–8 years) (Table 1).

Table 1

Baseline information of the study participants.

Number of patients	25
Number of hips	50
Mean age (year)	6
Age range (year)	2–8
Gender	10/15
(Female/Male)	
GMFCS level	Level2 → 1
	Level3 → 9
	Level4 → 3
	Level5 → 12

GMFCS = Gross Motor Functional Classification System.

Table 2. Presents descriptive statistics of RMP measurements on 2 occasions 6 weeks apart. (Figs. 5 and 6) illustrate distributions of intra- and inter-observer measurements on both occasions. No significant differences were observed in intra- and inter-observer measurements on both occasions.

3.1. Intra-observer reliability

Table 3 outlines ICCs of the intra-observer agreement. Excellent intra-observer reliability was observed between the 2 separate measurements of each observer.

3.2. Inter-observer reliability

Table 4 details the results of interobserver reliability assessments. Among 5 observers, inter-observer reliability was excellent for the 2 separate RMP measurements, with the mean ICC minimally increasing between the 2 measurement periods (mean ICC: 0.928, range: 0.838 to 0.979 and mean ICC: 0.936, range: 0.861–0.983, respectively).

When analyzed individually, the orthopedist exhibited excellent inter-observer reliability, with slightly improvement in the ICC between the 2 measurement periods although this reached no statistical significance (0.932–0.944). Likewise, the radiologist demonstrated excellent inter-observer reliability with modestly increase in the ICC (0.871–0.913)

4. Discussion

Our findings obtained by 5 observers with different levels of experience from 3 different expertise have demonstrated excellent inter- and intra-observer reliabilities in measuring RMP. Therefore, evidence from this study failed to support our hypothesis

Table 2

Descriptive statistics of RMP measurements on 2 occasions 6 weeks apart.

RMP values		Observer 1	Observer 2	Observer 3	Observer 4	Observer 5	* <i>P</i> values
First evaluation (week 0)	Mean ± SD (range)	28.34 ± 23.43 0 to 115	28.44 ± 24.12 0 to 118	32.34 ± 20.67 5 to 110	28.21 ± 20.29 0 to 107	30.38 ± 19.48 0 to 93	.85
Second evaluation (week 6)	Mean ± SD (range)	31.24 ± 26.29 0 to 119	29.35 ± 24.67 0 to 112	32.56 ± 22.54 7 to 122	27.46 ± 21.38 0 to 128	30.67 ± 18.78 0 to 93	.81
** <i>P</i> values		.59	.87	.99	.61	.94	

* Non-parametric one-way ANOVA test was used to compare differences among 5 observers; $P < .05$

** Mann–Whitney *U* test was used for comparison of intra-observer measurements; $P < .05$.

RMP = Reimers' Hip Migration Percentage.

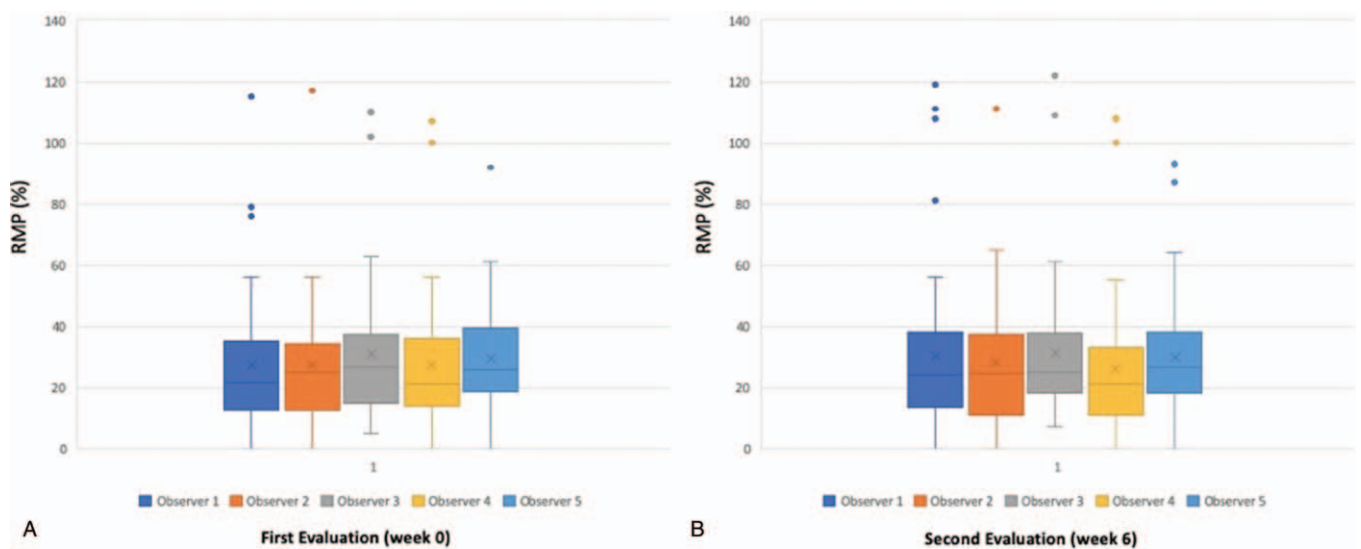


Figure 5. Box and whisker diagrams illustrating the distributions of Reimers' hip migration percentage (RMP) measurements among 5 observers at first and second evaluation. Horizontal bars represent the median, while boxes and whiskers demonstrate the interquartile ranges and data ranges, respectively.

that clinician' experience and expertise have a significant effect on the reliability of RMP measurement.

Parrot et al.^[2] analyzed 25 hip radiographs and concluded that observers with different levels of experience may increase the expected proportion of errors in the measurement of RMP whereas the acceptable variability can be obtained as long as consistent raters are used. In another study by Faraj et al^[6] which set out to assess inter- and intra-observer error of RMP, 2

observers with 6 months of pediatric orthopedic experience performed the measurements with a six-week interval on 44 hips of hip radiographs. The authors observed that only 5% of the intra-observer measurement differences were above 13% (95% confidence interval), but inter-observer measurement error was higher with a 95th upper confidence interval of 21% to 23%. Supporting Parrot et al suggestion, the authors interfered that such inter-measurer errors may be clinically inapplicable since

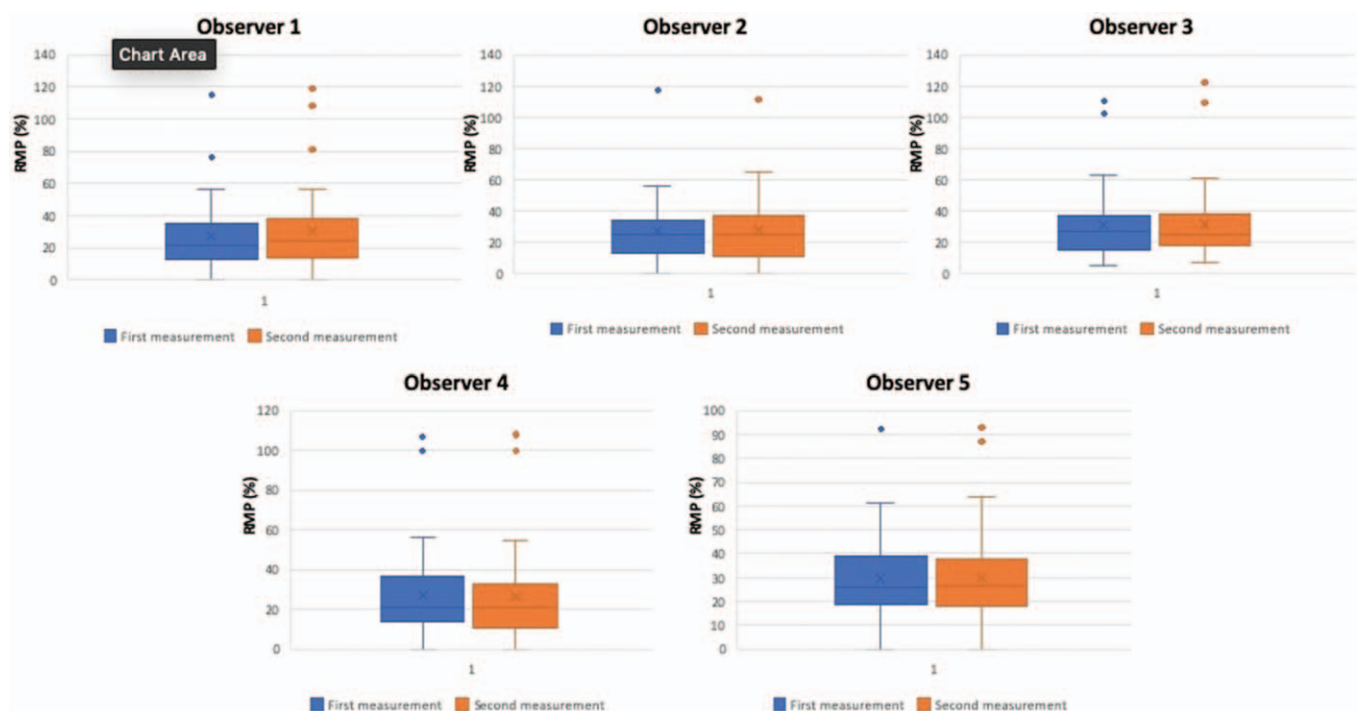


Figure 6. Box and whisker diagrams depicting the distributions of 2 separate Reimers' hip migration percentage (RMP) measurements of each observer. Horizontal bars represent the median, while boxes and whiskers demonstrate the interquartile ranges and data ranges, respectively.

Table 3
Details of intra-observer reliability.

Observers	ICC	95% confidence interval	
		Lower bound	Upper bound
O-1	0.977	0.946	0.991
O-2	0.988	0.972	0.995
O-3	0.974	0.937	0.989
O-4	0.986	0.965	0.994
O-5	0.977	0.944	0.990

ICC = Intraclass correlation coefficient, O-1 = orthopedic resident with 2 years of experience, O-2 = orthopedic surgeon with 8 years of experience, O-3 = radiology specialist with 3 years of experience, O-4 = radiology specialist with 7 years of experience, O-5 = physiatrist with 10 years of experience.

RMP should be measured and compared by the same experienced examiner to assure more consistent measurements. However, our results did not support such a conclusion. Even though our observers had different levels of experience in measuring RMP, this seems to have no significant effect on both intra- and inter-observer reliability in the measurements.

In another study, Cliffe et al^[12] interrogated the reliability of the pelvis radiographs examined in the study of Faraj et al^[6], as the researchers inferred that the orthopedic trainees were likely not to have adequate experience and thus the results may have been distinct if other specialists had completed the measurements. In their study, Cliffe et al analyzed 40 standardized hip radiographs based on a trained pediatric radiographer in measuring RMP and suggested that reliable measurements could be achieved owing to correct patient positioning and suitably experienced radiologist. In our study, similar to Cliffe et al study, 50 standardized hip radiographs were examined, and the intra- and inter-observer reliabilities were found excellent. Additionally, all RMP measurements obtained from the current study were compatible with each other and demonstrated no significant differences. We agree with Cliffe et al that standardized measurement protocol, including adequate pelvic radiographs,

Table 4
Details of inter-observer reliability.

Observer	ICC	95% CI
All observers		
• First measurement	0.928	0.838–0.979
• Second measurement	0.936	0.861–0.983
Orthopedic Physicians (O-1 and O-2)		
• First measurement	0.932	0.854–0.968
• Second measurement	0.944	0.865–0.977
Radiologist (O-3 and O-4)		
• First measurement	0.871	0.791–0.915
• Second measurement	0.913	0.812–0.923
O-2 vs O-3		
• First measurement	0.845	0.782–0.913
• Second measurement	0.958	0.900 to 0.983
O-2 vs O-4		
• First measurement	0.815	0.856–0.948
• Second measurement	0.948	0.874 to 0.978
O-2 vs O-5		
• First measurement	0.912	0.897–0.954
• Second measurement	0.939	0.853–0.975

ICC = Intraclass correlation coefficient, CI = confidence interval.

O-1 = orthopedic resident with 2 years of experience, O-2 = orthopedic surgeon with 8 years of experience, O-3 = radiology specialist with 3 years of experience, O-4 = radiology specialist with 7 years of experience, O-5 = physiatrist with 10 years of experience.

suitable patient positioning, and proper measurement method, is vital in obtaining reliable RMP measurements. However, we consider if suitable conditions are provided, excellent inter- and intra-observer variations can be obtained regardless of the physician's experience and expertise.

RMP has been widely recognized as the gold-standard measure to identify hip displacement in CP. Nevertheless, limited data is available regarding the reliability of this radiographic measure.^[2,5–8] Furthermore, according to our literature review, most of studies on the topic^[2,5,6,13,14] reflected rating results of physician with a similar level of experience from the same expertise, and thus neglecting the effect of clinician's experience and expertise on the reliability. However, differences in levels of rater experience and expertise may contribute to the observed differences in the measurement of RMP. Therefore, unlike most studies on the topic, the current study has taken into this issue account and comprised of observers with different levels of experience from different clinical specialties to evaluate the reliability of RMP.

When interpreting the findings of the present study, some limitations and strengths should be considered. The main strength is the inclusion of 5 observers with varying levels of experience from 3 different clinical specialties to question our hypothesis. To the best of our knowledge, this is the first study to specifically determine the effect of experience and expertise on the reliability of RMP measurements.

One of the major limitations is that we did not address the validity of this radiographic measure and the association of both the validity and the reliability of the measurement with temporal variations in femoral head ossification. Perhaps, with skeletal maturity, substantial differences in pelvic morphology and its radiographic appearance among different pediatric age groups may affect the reliability and validity of RMI measurements as it would not be easy to properly identify the relevant landmarks for measuring RMI. Thus, further studies are needed to clarify this effect. Another limitation is the relatively small number of patients included, though our sample size is comparable with the aforementioned reliability studies.

5. Conclusion

Clinician's experience and expertise may not affect both inter- and intra-observer reliability of RMP measurements in children with CP.

Author contributions

Conceptualization: Derya Yıldızlar.

Data curation: Mehmet Demirel.

Formal analysis: Onder Turna, Derya Yıldızlar, Onder Demirbaş, Yavuz Saglam.

Investigation: Necdet Demir, Mehmet Demirel, Onder Demirbaş, Yavuz Saglam.

Methodology: Necdet Demir, Mehmet Demirel, Onder Turna, Onder Demirbaş.

Writing – original draft: Mehmet Demirel.

Writing – review & editing: Mehmet Demirel.

References

- [1] Lee CW, Kim SG, Na SS. The effects of hippotherapy and a horse riding simulator on the balance of children with cerebral palsy. *J Phys Ther Sci* 2014;26:423–5.

- [2] Parrott J, Boyd RN, Dobson F, et al. Hip displacement in spastic cerebral palsy: repeatability of radiologic measurement. *J Pediatr Orthop* 2002; 22:660–7.
- [3] Hermanson M, Hägglund G, Riad J, et al. Head-shaft angle is a risk factor for hip displacement in children with cerebral palsy. *Acta Orthop* 2015;86:229–32.
- [4] Agarwal K, Chen C, Scher D, et al. Migration percentage and odds of recurrence/subsequent surgery after treatment for hip subluxation in pediatric cerebral palsy: a meta-analysis and systematic review. *J Child Orthop* 2019;13:582–92.
- [5] Craven A, Pym A, Boyd RN. Reliability of radiologic measures of hip displacement in a cohort of preschool-aged children with cerebral palsy. *J Pediatr Orthop* 2014;34:597–602.
- [6] Faraj S, Atherton W, Stott N. Inter-and intra-measurer error in the measurement of Reimers' hip migration percentage. *J Bone Joint Surg Br* 2004;86:434–7.
- [7] Ćobeljić G, Bajin Z, Lešić A, et al. A radiographic and clinical comparison of two soft-tissue procedures for paralytic subluxation of the hip in cerebral palsy. *Int Orthop* 2009;33:503–8.
- [8] Reimers J, Bialik V. Influence of femoral rotation on the radiological coverage of the femoral head in children. *Pediatr Radiol* 1981;10:215–8.
- [9] Kim HT, Kim JI, Yoo CI. Diagnosing childhood acetabular dysplasia using the lateral margin of the sourcil. *J Pediatr Orthop* 2000;20:709–17.
- [10] Walter S, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998;17:101–10.
- [11] Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Internal Med* 2007; 147:573–7.
- [12] Cliffe L, Sharkey D, Charlesworth G, et al. Correct positioning for hip radiographs allows reliable measurement of hip displacement in cerebral palsy. *Dev Med Child Neurol* 2011;53:549–52.
- [13] Pountney T, Mandy A, Gard P. Repeatability and limits of agreement in measurement of hip migration percentage in children with bilateral cerebral palsy. *Physiotherapy* 2003;89:276–81.
- [14] Ha M, Okamoto T, Fukuta T, et al. Preoperative radiologic predictors of successful soft tissue release surgery for hip subluxation among cerebral palsy patients: a STROBE compliant study. *Medicine* 2018;97:e11847.