

## Limited Intraobserver and Interobserver Reliability for the Common Measures of Hip Joint Congruency Used in Dysplasia

A. Noelle Larson MD, Brien Rabenhorst MD,  
Adriana De La Rocha BS, MS,  
Daniel J. Sucato MD, MS

Received: 14 April 2011 / Accepted: 5 October 2011 / Published online: 29 November 2011  
© The Association of Bone and Joint Surgeons® 2011

### Abstract

**Background** A congruent hip frequently is cited in the literature as a prerequisite for performing a reconstructive pelvic osteotomy for hip dysplasia. The designation as congruent is important as incongruent joints generally are regarded as requiring a salvage-type procedure. Several measures of congruency are described in the literature, but it is unclear whether these measures are reliable.

**Questions/purposes** We therefore evaluated the intraobserver and interobserver reliabilities of three measures of hip congruency in patients with mild to severe dysplasia.

**Methods** We selected radiographs of 30 skeletally mature patients with symptomatic hip dysplasia who were potential candidates for a periacetabular osteotomy. All radiographs were rated by six reviewers on two occasions. Reviewers were shown a series of AP and von Rosen views of either the symptomatic hip or the contralateral side and asked to rate congruency using two previously published measures – the classifications of Yasunaga et al. and Okano et al. The reviewers also were asked to determine whether the hip was congruent based on their subjective opinion (yes/no). The radiographs were reviewed and measurements were repeated 1 month later. Kappa analysis was performed to determine intraobserver and interrater reliabilities between and among individual raters, respectively.

**Results** Using the classifications of Yasunaga et al. and Okano et al., we found intraobserver reliability kappa values of 0.43 and 0.37, respectively and interobserver reliability values of 0.18 and 0.25, respectively. The intraobserver reliability of a subjective opinion of yes/no response was 0.74 while the interobserver reliability was 0.21.

**Conclusions** When evaluating a spectrum of hip dysplasia, commonly used measures of hip congruency have low intraobserver and interobserver reclassifications of Yasunaga et al. and Okano et al. Interobserver reliability for all three methods was low. New radiographic parameters to consistently identify congruency will be helpful for evaluating the preoperative indications for reconstructive osteotomy.

**Clinical Relevance** Hip congruency has long been cited as a requirement for a reconstructive pelvic osteotomy, but based on the above findings, there is no clear agreement regarding what comprises a congruent joint. Thus, better criteria are needed to assess preoperative hip congruency to determine the role of congruency in patient outcomes after surgery for hip dysplasia.

---

The institution of one or more of the authors (DJS) has received funding from DePuy and Medtronic.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request. Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained. This study was performed at the Texas Scottish Rite Hospital for Children.

---

A. N. Larson  
Department of Orthopedic Surgery, Mayo Clinic,  
Rochester, MN, USA

B. Rabenhorst  
Department of Orthopaedic Surgery and Rehabilitation,  
Texas Tech University Health Sciences Center,  
Lubbock, TX, USA

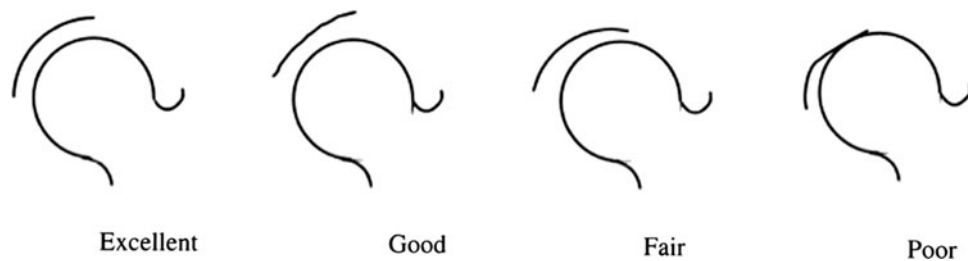
A. De La Rocha, D. J. Sucato (✉)  
Pediatric Orthopedics, Texas Scottish Rite  
Hospital for Children, 2222 Welborn St,  
Dallas, TX 75230, USA  
e-mail: dan.sucato@tsrh.org

**Introduction**

When considering a reconstructive osteotomy for treatment of hip dysplasia, several factors must be considered, including the degree of dysplasia, the presence of osteoarthritis, and hip congruency [13]. Hip congruency refers to the relationship of the femoral head contour to the acetabulum. In a congruent joint, the arc of the femoral head matches the arc of the acetabulum with a consistent joint space throughout. Some authors suggest a reconstructive pelvic osteotomy, such as a Salter or Pemberton osteotomy, can be performed only if adequate congruency exists between the acetabulum and the femoral head [11, 12, 14], using either a neutral hip radiographic view or a von Rosen (abduction and internal rotation view) [16]. If the joint is incongruent, a salvage osteotomy such as a Shelf or a Chiari may be required to obtain additional acetabular coverage. The periacetabular osteotomy (PAO) used in the adolescent or adult hip is a procedure used to achieve the necessary amount of femoral head coverage and may challenge some of the traditional concepts of preoperative joint congruency.

Previous studies have evaluated intrarater and interobserver reliabilities of radiographic measures of dysplasia and osteoarthritis [1, 3, 8, 9]. There are several reported classifications of hip congruency [6, 7, 9, 17]. Two classifications of congruency are commonly used [2, 5, 9, 17, 20]. Yasunaga et al. proposed a four-part classification for congruency: excellent if the curvature of the acetabulum and femoral head is identical or nearly identical and the

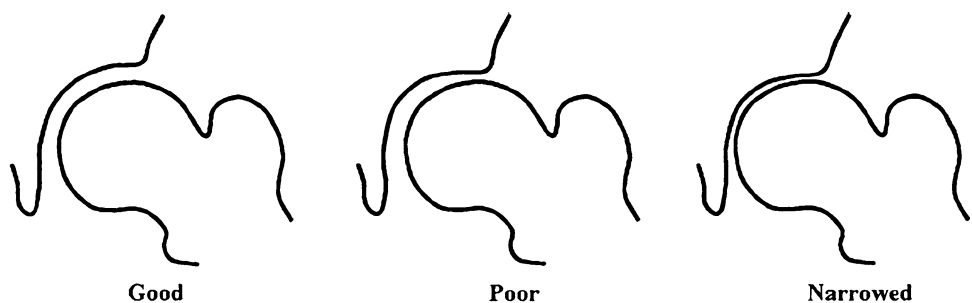
joint space is entirely maintained, good if the joint space is adequately maintained, fair if partial narrowing of the joint space has occurred, and poor if the joint space is obliterated (Fig. 1) [20]. Okano et al. modified this to a three-part classification: good if the joint space width at the narrowest point is 50% or greater than the joint space width at the widest point, poor if the joint space width at the narrowest point is less than 50% of the joint space width at the widest point, and narrowed if no more than 2 mm of joint space remains in any area of the hip (Fig. 2) [9, 10]. Also, Clohisy et al. described congruency based on the subjective assessment of conformity between the acetabulum and femoral head using all radiographic views but with an emphasis on the AP view [2, 3]. They defined a congruent hip as one that has matching arcs of the femoral head and the acetabulum [3]. In our practice group, however, we have noted differences in opinion regarding what constitutes a congruent joint, particularly in cases of severe dysplasia. A different assessment of congruency could substantially affect the approach to treatment. Also, as newer literature reports the outcomes of PAO based on preoperative congruency, it is important to have reliable, reproducible measures of congruency to render the results of these studies meaningful [9, 19]. As these criteria are general, for the purposes of this study, we defined a third rating based solely on the rater’s clinical practice, whether the joint would be considered congruent or not (the subjective classification). The reliability of these measures, however, has not been well studied [3, 9]. Clohisy et al.



**Fig. 1** The classification of Yasunaga et al. is comprised of four possible ratings for congruency. (Published with permission from Yasunaga Y, Takahashi K, Ochi M, Ikuta Y, Hisatome T, Nakashiro J,

Yamamoto S. Rotational acetabular osteotomy in patients forty-six years of age or older: comparison with younger patients. *J Bone Joint Surg Am.* 2003;85:266–272.)

**Fig. 2** The classification of Okano et al. has three possible ratings assessing hip congruency. (Published with permission from Okano K, Enomoto H, Osaki M, Shindo H. Joint congruency as an indication for rotational acetabular osteotomy. *Clin Orthop Relat Res.* 2009;467:894–900.)



reported an interrater kappa coefficient of 0.29 for a yes/no congruency rating [3]. However, Okano et al. reported a high interrater kappa coefficient of 0.92 using their three-part classification of congruency [9].

We therefore performed this study to determine the intrarater and interrater reliabilities of these three measures of hip congruency.

## Patients and Methods

We retrospectively reviewed an institutional database containing records of 158 adolescents with hip dysplasia who were potential candidates for PAO. All patients had symptomatic hip dysplasia. Radiographs of 30 hips were selected meeting the following criteria. Inclusion criteria were symptomatic hip dysplasia, skeletal maturity (closure of the triradiate cartilage and proximal femoral physis), and high-quality AP and von Rosen view radiographs (abduction internal rotation view). Representative radiographs were chosen independently by one author (ANL) who was not involved in the review process. All radiographs were obtained using a standardized imaging protocol, all patient information was removed, and each hip was assigned a study number. Various dysplastic hips were selected as to encompass a wide spectrum of disease severity and joint congruency. Several otherwise normal unaffected hips on the contralateral side of the dysplastic hips also were selected to complete the assortment. Several hips had severe dysplasia with severe subluxation, which reduced at least partially on the von Rosen view. Underlying diagnoses included developmental hip dysplasia, cerebral palsy, dysplasia associated with orthopaedic syndromes, and Legg-Calvé-Perthes disease. We excluded hips with frank dislocation and hips in skeletally immature patients. A total of 30 hips were included in the study. Institutional review board approval was obtained for all aspects of this study.

The observers had no specific training using the classification systems of Yasunaga et al. [20] and Okano et al. [9] beyond what is readily available in the literature, and simply were given published diagrams as a guideline. We used the resources typically available to an orthopaedic surgeon and did not seek any specialized training or contact with Dr Okano or Dr Yasunaga. The AP and von Rosen views for each hip were reviewed by four experienced pediatric orthopaedic surgeons (DJS, JGB, DAP, KER) and two pediatric orthopaedic fellows (SG, MK). All four attending surgeons commonly treat hip dysplasia. Fellows were included to see if there would be better or worse agreement among less experienced surgeons. The observers were blinded regarding patient physical examination, history, symptoms, and identity. Observers were provided with the diagrams (Figs. 1 and 2) for the classification

systems of Yasunaga et al. [20] and Okano et al. [9], respectively, and were asked to provide a yes or no response regarding whether the hip was congruent based on the criteria that they would use in a clinical setting. The observers classified the congruency of each hip using the three methods and submitted confidential responses. To obtain intraobserver reliability measures, this procedure was repeated 1 month later with the viewing order of the films shuffled.

Before beginning the study, a biostatistician (RHB) was consulted to determine the appropriate number of cases to include in the study and assist in study design. Based on the number of reviewers, number of rating criteria, and heterogeneous patient population, a sample size of 30 cases was recommended. This historically has been used for other comparison studies at our institution. A formal power analysis was not completed for this study because the goal was not to test a hypothesis, but to estimate differences between the various classification systems.

Kappa values were used to measure intraobserver and interobserver reliabilities with 1.0 indicating perfect agreement, and 0 indicating a level of agreement expected by chance alone [4]. A simple kappa was used to measure agreement for the subjective rating. For the classifications of Yasunaga et al. [20] and Okano et al. [9], which are ordered variables, we used a weighted kappa as a refinement [15]. This differentiates disagreement by proximity of the response. Thus, if the ratings disagree by only one level compared with two levels, this is reflected in a higher (and improved) weighted kappa score. Intrarater reliability was calculated for each observer, comparing results from the first reading with the results from the second reading 1 month later. Also, a combined intrarater reliability was calculated, assessing the intrarater agreement in the group of attending staff and in the group of fellows. Interrater reliability was measured between the staff and the fellows for the first and second readings. No subanalysis was performed based on the degree of dysplasia. Instead, we sought to test the validity of these hip congruency measures over a wide spectrum of hip dysplasia. Level of agreement between the different congruency rating methods was assessed using a z-test.

## Results

A total of 30 hips were reviewed, representing a spectrum of disease from the unaffected sides to severely dysplastic hips. The intrarater reliability ranged from 0.30 to 0.71 for the criteria of Yasunaga et al. (Table 1). An intrarater reliability for the criteria of Yasunaga et al. could not be performed for one staff surgeon, as the rater did not select all choices. The reliability using the classification of

Okano et al. ranged from 0.15 to 0.41 (Table 1). Looking at each reviewer independently using the subjective yes/no criteria, we found that the intrarater reliability ranged from  $-0.03$  to  $0.84$ . The combined intrarater reliability was  $0.74$  for the subjective criteria,  $0.43$  for the Yasunaga et al. classification and  $0.37$  for the Okano et al. classification (Table 2). Comparing the combined intrarater reliability between the staff and the fellows, we found that the staff and fellows had low intrarater reliability for the classifications of Yasunaga et al. and Okano et al. Overall, only the subjective yes/no method yielded high combined intrarater reliability. The z-test was used to determine which rating system had superior intrarater reliability. The subjective yes/no method had statistically improved kappa scores for combined intrarater reliability compared with the classifications of Okano et al. and Yasunaga et al. ( $p < 0.001$  and  $p < 0.001$ , respectively).

Using the classifications of Yasunaga et al., Okano et al., and subjective criteria, the interrater reliability ranged from  $-0.02$  to  $0.46$  (Table 3). Of all the comparisons, the fellows had the best agreement on the first evaluation of the radiographs using the classification of Okano et al. and subjective criteria. All other interrater analyses revealed a kappa value less than  $0.4$ . The subjective opinion of congruency between the staff was particularly poor, with results being no different than chance. We evaluated the interrater reliability for both trials to see if agreement

differed between the first and second reviews; however, no improvement was seen. There was no difference in results when comparing agreement among the fellows and agreement among the staff. Using the z-test to determine which rating system had superior interrater reliability, we found that the subjective criteria and the classification of Okano et al. had statistically better matching rates than the classification of Yasunaga et al. ( $p < 0.001$  and  $p < 0.001$ , respectively).

## Discussion

Opinions differ in clinical practice regarding what constitutes a congruent hip, particularly the joint in a severely dysplastic hip. This discrepancy may result in markedly different recommendations for the type of surgical treatment. Also, as the literature reports the outcomes of PAO based on preoperative congruency, it is important to have reliable, reproducible measures of congruency to render the results of these studies meaningful [9, 19]. We have noted differences in opinion regarding what constitutes a congruent joint, particularly in cases of severe dysplasia. A different assessment of congruency could substantially affect the approach to treatment. Of the various reported classifications of hip congruency [6, 7, 9, 17], we selected three that are commonly used, are easily measured in a clinical setting, do not require any specific imaging software, and, thus, are of potential practical value. We performed this study to measure the intraobserver and interobserver reliabilities of these commonly used measures of congruency.

Our investigation has several limitations. First, we evaluated only radiographs. The observers had no knowledge of the patient's history or physical examination findings and could not contextualize the radiographs. Second, measurements of interrater reliability from the second reading may have limited validity owing to practice effects and potential bias, although reviewers were asked to avoid discussing the study with other practitioners between the two readings. Third, although the radiographs were taken at one center we noted some variability in positioning

**Table 1.** Intrarater reliability

Reviewer	Yasunaga et al.* Kappa value	Okano et al.* Kappa value	Yes/no (subjective) Kappa value
Fellow 1	0.30	0.34	0.43
Fellow 2	0.36	0.38	0.72
Staff 1	n/a <sup>^</sup>	0.15	$-0.03$
Staff 2	0.38	0.40	0.47
Staff 3	0.64	0.41	0.84
Staff 4	0.71	0.34	0.64

Values  $< 0$  reflects worse than chance alone; \* weighted kappa; <sup>^</sup>rater did not select all choices, so unable to perform intrarater reliability.

**Table 2.** Combined intrarater reliability

Reviewers	Yasunaga et al.* Kappa value (95% CI)	Okano et al.* Kappa value (95% CI)	Yes/no (subjective) Kappa value (95% CI)
All reviewers combined	0.43 (0.34–0.53)	0.37 (0.25–0.49)	0.74 (0.64–0.84)
Fellows combined	0.34 (0.17–0.51)	0.38 (0.18–0.58)	0.63 (0.43–0.83)
Staff combined	0.48 (0.36–0.60)	0.37 (0.22–0.51)	0.78 (0.65–0.90)

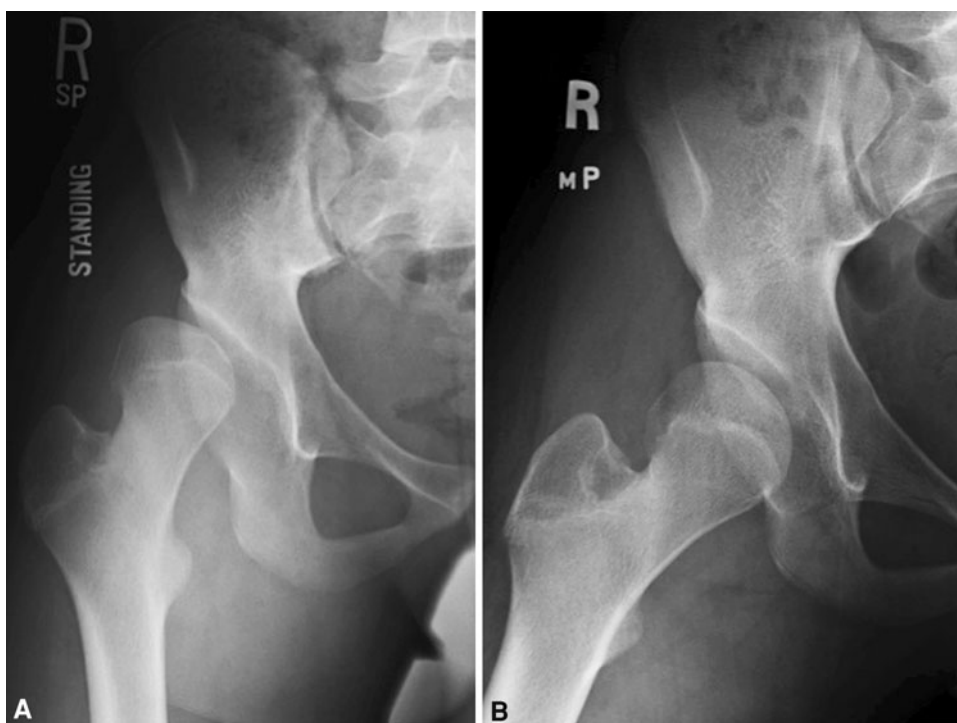
Subjective criteria had higher kappa values than the those of Yasunaga et al. and Okano et al. ( $p < 0.001$ ) for all reviewers combined; \* weighted kappa.

**Table 3.** Interrater reliability

Reviewers	Number of observers	Yasunaga et al.* Kappa value (95% CI)	Okano et al.* Kappa value (95% CI)	Yes/no (subjective) Kappa value (95% CI)
Fellows only – first reading	2 fellows	0.24 (0–0.48)	0.41 (0.10–0.72)	0.46 (0.11–0.81)
Fellows only – second reading	2 fellows	0.33 (0.09–0.57)	0.33 (0.04–0.62)	0.25 (–0.12–0.62)
Staff only – first reading	4 staff	0.16 (0.06–0.26)	0.26 (0.12–0.40)	0.02 (–0.20–0.24)
Staff only – second reading	4 staff	0.25 (0.15–0.35)	0.16 (0.00–0.32)	–0.02 (–0.26–0.22)
Combined – first reading	2 fellows 4 staff	0.18 (0.12–0.24)	0.25 (0.15–0.35)	0.21 (0.09–0.33)
Combined – second reading	2 fellows 4 staff	0.22 (0.14–0.30)	0.24 (0.14–0.34)	0.13 (0.01–0.25)

There were no statistically significant differences in kappa values between the three methods; \* weighted kappa.

**Fig. 3A–B** These are representative (A) AP and (B) von Rosen views of a hip with severe dysplasia. Subjectively, four raters judged this hip to be incongruent, and two thought the hip was congruent. Interestingly, all six raters believed this represented a poorly congruent hip with the Okano criteria. With the Yasunaga criteria, there was one poor rating, two fair, and three good.

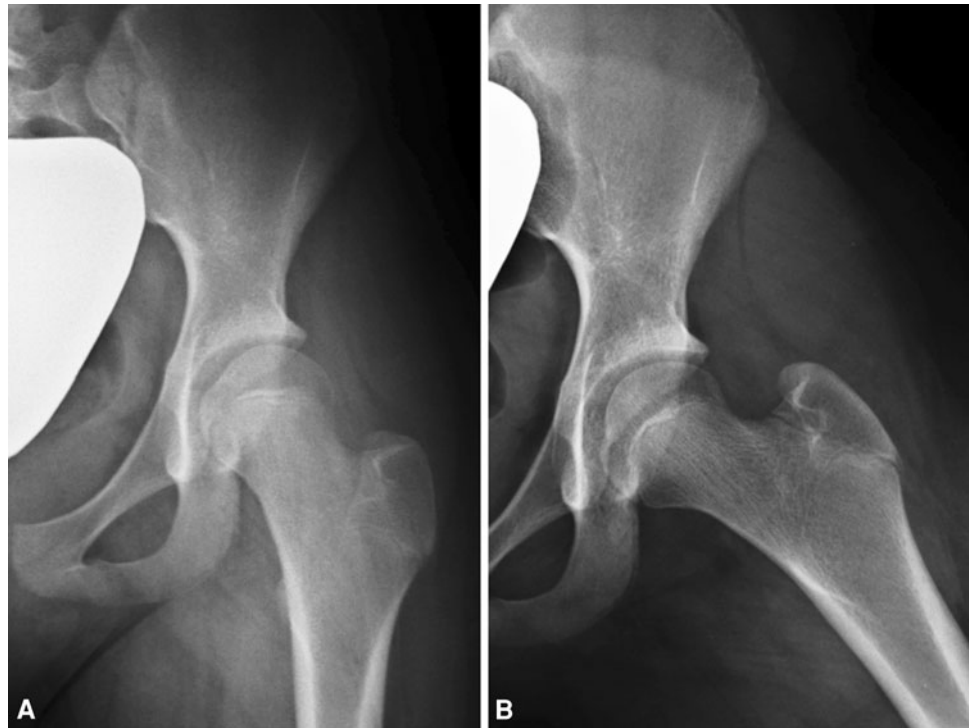


of the affected leg. In particular, positioning for the von Rosen view can be limited by the patient's symptoms and restricted ROM. For this reason, all observers were given the AP and von Rosen views with which to rate the hip congruency. Fourth, some of radiographs were of hips of patients with severe dysplasia (Fig. 3). These three congruency measures may have better intrarater and interrater reliabilities if applied to a patient population with more subtle findings of hip dysplasia. In some cases, improved agreement was noted in patients with minimal dysplasia (Fig. 4). Finally, our observers had no specific training in using the classification systems of Yasunaga et al. and Okano et al. It is possible that with training the reliability would have been higher. Nevertheless, we provided the reviewers with the information currently available in the literature.

We presumed there would be good intrarater reliability for the three methods. We found low combined intraobserver reliability for the classifications of Okano et al. and Yasunaga et al. When evaluating the reviewers independently, two attending surgeons had much higher intrarater reliability for the classification of Yasunaga et al. For the most part, our reviewers had difficulty duplicating their results 1 month apart for either classification system. To our knowledge, there are no previous published studies regarding intraobserver reliability of the classification systems of Okano et al. and Yasunaga et al. The combined intraobserver reliability for the subjective criteria was high at 0.74. Even if raters do not agree among themselves on a congruent hip, they consistently recognize what they personally consider to be a congruent joint. Clohisy et al. reported on intrarater reliability of a subjective measure of congruency [3].



**Fig. 4A–B** These are representative (A) AP and (B) von Rosen views of a hip with mild dysplasia. All four staff agreed that this hip was excellent, good, and congruent. The two fellows rated the hip as good or poor with the criteria of Yasunaga et al. and Okano et al., but agreed subjectively that the hip was congruent.



They found a combined intrarater reliability of 0.50. Thus, subjective opinion appears to produce the highest intrarater reliability when compared with other measures.

We also presumed there would be low interrater reliability for measurements of hip congruency. We found low interobserver reliability for all three methods when used to measure congruency in hips with a spectrum of hip dysplasia. This was true for the subgroup of pediatric orthopaedic fellows and the attending orthopaedic surgeons. Clohisy et al. had similar findings in their study on interobserver reliability for various hip measures [3]. They rated congruency using a subjective yes/no criteria with an AP view of the pelvis [3]. This method is similar to our subjective criteria. They found the kappa coefficient for the congruency rating was poor at 0.29. We had similar results with an interrater reliability of 0.21 using subjective criteria for congruency. As our method is based only on the qualitative judgment whether the arc of the acetabulum matches the arc of the femoral head, it is understandable that there would be differences in opinion, resulting in a low kappa score. Okano et al. and Yasunaga et al. provided more detailed descriptions of congruency using three- and four-part classification systems [9, 10, 17–20]. In our literature review, we did not find any previous reports of the interrater reliability for the classification of Yasunaga et al. As part of a larger study, Okano et al. rated 20 hips using their method and reported an excellent interrater kappa value of 0.92 [9]. To our knowledge, this has not been reproduced in other studies. Our overall interrater kappa

using the classification of Okano et al. was 0.25, reflecting low agreement in our patient population. In contrast to our reviewers, Okano et al. [10] likely are familiar with their classification system and are better able to produce similar results, and thus high kappa scores. Alternatively, the hips in our series might have had more severe dysplasia than those in the series by Okano et al., rendering their classification system less reliable among our reviewers.

There has been increasing interest in the role of hip congruency as a surgical indication and as a prognostic factor for results after acetabular osteotomy. Traditionally, congruency has been considered a prerequisite for reconstructive osteotomy. Our observations suggest practitioners may have their own subjective understanding of what constitutes a congruent hip. Only subjective opinion was a reproducible measure of congruency for the individual surgeon, with good intrarater reliability. However, other commonly used measures of congruency have low intraobserver reliability, and all three methods have low interobserver reliability. Additional studies with more specific guidelines are needed to validate the current measures of congruency. Alternatively, a new radiologic rating of congruency with greater reproducibility among practitioners may aid in refining operative indications and understanding postoperative outcomes for osteotomies in the context of severe hip dysplasia.

**Acknowledgments** We thank John G. Birch MD, David A. Podeszwa MD, Karl E. Rathjen MD, Sumeet Garg MD, and Michael Kwon MD for serving as reviewers in this study. We also thank

Tara Kristof, Cindy Daniel, and Sarah Lassen for assistance with manuscript preparation, and Richard H. Browne PhD, for assistance in study design and biostatistics.

## References

- Broughton NS, Brougham DI, Cole WG, Menelaus MB. Reliability of radiological measurements in the assessment of the child's hip. *J Bone Joint Surg Br.* 1989;71:6–8.
- Clohisey JC, Carlisle JC, Beaulé PE, Kim YJ, Trousdale RT, Sierra RJ, Leunig M, Schoenecker PL, Millis MB. A systematic approach to the plain radiographic evaluation of the young adult hip. *J Bone Joint Surg Am.* 2008;90(suppl 4):47–66.
- Clohisey JC, Carlisle JC, Trousdale R, Kim YJ, Beaulé PE, Morgan P, Steger-May K, Schoenecker PL, Millis M. Radiographic evaluation of the hip has limited reliability. *Clin Orthop Relat Res.* 2009;467:666–675.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37–46.
- Cunningham T, Jessel R, Zurakowski D, Millis MB, Kim YJ. Delayed gadolinium-enhanced magnetic resonance imaging of cartilage to predict early failure of Bernese periacetabular osteotomy for hip dysplasia. *J Bone Joint Surg Am.* 2006;88:1540–1548.
- Li PL, Ganz R. Morphologic features of congenital acetabular dysplasia: one in six is retroverted. *Clin Orthop Relat Res.* 2003;416:245–253.
- Miller JA Jr, Brand RA, Andrews JG. Quantifying osteoarthrotic hip incongruence: an approach to optimizing osteotomies. *Clin Orthop Relat Res.* 1997;343:124–134.
- Nelitz M, Guenther KP, Gunkel S, Puhl W. Reliability of radiological measurements in the assessment of hip dysplasia in adults. *Br J Radiol.* 1999;72:331–334.
- Okano K, Enomoto H, Osaki M, Shindo H. Joint congruency as an indication for rotational acetabular osteotomy. *Clin Orthop Relat Res.* 2009;467:894–900.
- Okano K, Yamada K, Takahashi K, Enomoto H, Osaki M, Shindo H. Joint congruency in abduction before surgery as an indication for rotational acetabular osteotomy in early hip osteoarthritis. *Int Orthop.* 2010;34:27–32.
- Salter RB. Role of innominate osteotomy in the treatment of congenital dislocation and subluxation of the hip in the older child. *J Bone Joint Surg Am.* 1966;48:1413–1439.
- Salter RB, Dubos JP. The first fifteen year's personal experience with innominate osteotomy in the treatment of congenital dislocation and subluxation of the hip. *Clin Orthop Relat Res.* 1974;98:72–103.
- Siebenrock KA, Scholl E, Lottenbach M, Ganz R. Bernese periacetabular osteotomy. *Clin Orthop Relat Res.* 1999;363:9–20.
- Utterback JD, MacEwen GD. Comparison of pelvic osteotomies for the surgical correction of the congenital hip. *Clin Orthop Relat Res.* 1974;98:104–110.
- Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005;37:360–363.
- von Rosen. Diagnosis and treatment of congenital dislocation of the hip joint in the new-born. *J Bone Joint Surg Br.* 1962;44:284–291.
- Yasunaga Y, Ikuta Y, Kanazawa T, Takahashi K, Hisatome T. The state of the articular cartilage at the time of surgery as an indication for rotational acetabular osteotomy. *J Bone Joint Surg Br.* 2001;83:1001–1004.
- Yasunaga Y, Ikuta Y, Shigenobu T, Nakamura S, Yamamoto S, Nakashiro J. Rotational acetabular osteotomy for hip dysplasia: spontaneous medial enlargement of the acetabulum. *Acta Orthop Scand.* 2001;72:8–12.
- Yasunaga Y, Ochi M, Terayama H, Tanaka R, Yamasaki T, Ishii Y. Rotational acetabular osteotomy for advanced osteoarthritis secondary to dysplasia of the hip. *J Bone Joint Surg Am.* 2006;88:1915–1919.
- Yasunaga Y, Takahashi K, Ochi M, Ikuta Y, Hisatome T, Nakashiro J, Yamamoto S. Rotational acetabular osteotomy in patients forty-six years of age or older: comparison with younger patients. *J Bone Joint Surg Am.* 2003;85:266–272.