

Original Research

The Reliability of Intraoperative Hip Arthrography in Cerebral Palsy Hip Reconstruction

Colyn Watkins, MD¹; Laura Lins, MD²; Trisha Miller, MS¹; Travis Matheney, MD¹; Brian Snyder, MD¹; Kemble Wang, MD³; Kerr Graham, MD⁴; Benjamin Shore, MD¹

¹Boston Children's Hospital, Boston, MA; ²University of Wisconsin-Madison Department of Orthopedic Surgery, Madison, WI; ³Melbourne Orthopaedic Surgeons, East Melbourne, Australia; ⁴The Royal Children's Hospital, Melbourne, Australia

Correspondence: Colyn Watkins, MD, Boston Children's Hospital, 300 Longwood Ave., Boston, MA 02115.
E-mail: Colyn.Watkins@childrens.harvard.edu

Received: September 30, 2022; Accepted: February 10, 2023; Published: May 1, 2023

DOI: 10.55275/JPOSNA-2023-595

Abstract

Background: This study investigates surgeon interpretation of intraoperative hip arthrography after femoral osteotomy to guide decision-making regarding addition of concomitant pelvic osteotomy (PO) for Cerebral Palsy (CP) hip reconstruction.

Methods: Ten orthopaedic surgeons reviewed preoperative radiographic images of 21 children with CP, undergoing hip reconstruction. Surgeons recorded if PO was indicated after 1) reviewing the preoperative radiographs. Raters who exhibited poor reliability (kappas <0.60) were excluded from arthrogram analysis. Agreement in decision-making was evaluated by estimating the concordance across six raters with respect to changing their recommendation after reviewing the arthrogram images by estimating Fleiss's kappa coefficient.

Results: Intra-rater reliability for the three surgeons who conducted repeat evaluations was substantial to almost perfect (k=0.67, 0.86, 1.00). Pairwise inter-rater reliability for radiographic decision was moderate to almost perfect with kappa values ranging from 0.56 to 0.87. In 20 of 32 hip evaluations, at least 5 of 6 raters recommended PO after reviewing the radiograph, 64% of the cohort (122/192). However, upon reviewing the arthrogram images, only 45% (86/192) of reviewers recommended a PO. Attending surgeons changed their decisions away from PO in 35% of cases overall while fellows changed their decisions in only 7% of cases. Further radiographic characterization of the labrum position demonstrated substantial agreement (kappa=0.73; 95% CI=0.48-0.98) and almost perfect agreement regarding presence of dye pool (kappa=0.85; 95% CI=0.50-1.00).

Conclusions: Interpretation of intraoperative hip arthrography to guide addition of PO in CP hip reconstruction is reliable with moderate agreement according to experience level and with substantial retest reliability.

Level of Evidence: Level IV

Key Concepts

- Arthrograms can be helpful to determine hip stability during hip reconstruction in cerebral palsy patients.
- The reliability of the interpretation of arthrograms is an important area of study.
- Decision-making in neuromuscular hip reconstruction is not standardized across surgeons or institutions.

Introduction

The vast majority of children with cerebral palsy (CP) are born with anatomically normal hips without dysplasia or dislocation.¹ However, we understand that while the brain injury in those with CP is non-progressive, the musculoskeletal deformities tend to progress at a rate commensurate with the child's functional limitations.² Hip displacement is the second most common orthopaedic manifestation of children with CP.² The risk of hip displacement in all children with CP has been estimated at approximately 33% in large, population-based studies.^{3,4} In non-ambulatory children, the risk of hip displacement can be as high as 60-70%.⁵ Hip subluxation and dislocation in children with CP hip dysplasia can lead to difficulty with activities of daily living, hygiene, and pain.^{1,6,7}

The goals of orthopaedic treatment for hip displacement in children with CP hip dysplasia is to obtain a concentrically reduced hip with pain-free, symmetric range of motion. Reconstructive surgical treatment typically involves a combination of soft tissue releases, proximal femoral varus derotation osteotomy (VDRO), and pelvic osteotomy (PO).⁸ However, the decision of when to add a PO is in part, subjective without clear guidelines or surgical indications within the literature.^{2,8}

Intraoperative hip arthrography has been well-described for decision-making in developmental hip dysplasia.⁹ However, hip arthrography has only been described in a small sample of children with CP undergoing CP hip reconstruction¹⁰ and indications regarding how and when to interpret CP arthrograms are missing from the

literature. Furthermore, reliability and application of intraoperative arthrography in CP hip reconstruction has not been previously studied. Therefore, the aims of this study were 1) to characterize how intraoperative hip arthrography after VDRO affects decision-making regarding performing a PO in children with CP hip dysplasia; 2) report the intra- and inter-rater reliability associated with these interpretations and 3) identify specific intraoperative arthrographic parameters which could be used in future study/classification to help standardize decision-making regarding CP hip reconstruction. We hypothesized that the interpretation of intraoperative hip arthrography would have moderate retest reliability and that the additional arthrogram information would guide surgeons when to not perform a concomitant PO in the setting of CP hip reconstruction.

Methods

A survey was developed using the Research Electronic Data Capture system (REDCap Inc, Nashville, TN) that contained de-identified preoperative and intraoperative radiographic images (including 32 arthrograms) of 21 children with CP who underwent hip reconstruction surgery at one tertiary care institution. Patients were selected if they had intraoperative hip arthrogram performed at the time of their hip reconstruction surgery. Patient characteristics were not provided to the scoring physicians at the time of survey administration. Plain radiographic characteristics are summarized in Table 1 and were presented to the examiners at the time of survey participation. Ten pediatric orthopaedic

Table 1. Summary

Patient Radiographic characteristics (N=21)	Freq.	(%)
Radiographic characteristics		
Left Acetabular Index (<i>mean ± SD</i>)	27.5	± 7.1
Right Acetabular Index (<i>mean ± SD</i>)	29.4	± 6.2
Left Migration Progression (<i>mean ± SD</i>)	51.3	± 18.6
Right Migration Progression (<i>mean ± SD</i>)	55.9	± 19.2

SD, standard deviation

surgeons from two institutions with variable experience participated in the survey. Five pediatric orthopaedic fellows from two institutions (represented novice learners with limited experience in hip arthrography) and five pediatric orthopaedic fellowship-trained surgeons (represented intermediate to advanced practitioners with variable clinical experience in interpreting CP hip arthrography) participated in the survey. The survey assessed the anticipated surgical plans after 1) reviewing the preoperative radiographs and 2) after viewing the intraoperative arthrograms which were performed after the VDRO. Representative arthrogram images were selected from the senior author’s intraoperative images. A series of three images were provided for review, which included an AP pelvis, a frog lateral X-ray, and a flexed and adducted view. Arthrograms were performed after the femoral osteotomy by placing an 18-gauge spinal needle along the femoral neck into the hip joint through the lateral incision. A mixture of 1:1 normal saline and Optiray (Ioversol, Guerbet, NJ) (5-10cc) was injected into the joint to obtain the diagnostic arthrogram images. In the case of unilateral procedures, arthrogram dye was injected into the contralateral “normal” hip for comparison. After viewing the preoperative radiographs, survey participants were asked whether they would perform a PO in addition to performing appropriate soft tissue lengthening and VDRO for each hip. The same questions were repeated after reviewing the intraoperative arthrogram images. The five-fellowship trained attending orthopaedic surgeons repeated the

survey 6 months apart to assess intra-rater reliability. Intra-rater reliability was not assessed amongst the fellows, as it was felt that their learning curve was significant, and the 6-month repeat assessment would not be accurate based on their educational development.

Intra-rater agreement for preoperative radiographs was assessed for five surgeons by estimating kappa coefficients with 95% confidence intervals. Pairwise agreement in decision-making for preoperative radiographs was evaluated by estimating kappa coefficients along with 95% confidence intervals between each of the 10 raters. Four raters in total who exhibited poor intra-rater (n=2) and inter-rater (n=4) reliability (kappas <0.60) for preoperative radiographs were excluded from arthrogram analysis. These raters included two fellows and two experienced pediatric orthopaedic surgeons, highlighting the concept that decision-making regarding when to perform a pelvic osteotomy on plain radiographs is unreliable. Because the goal of this study was to assess how intraoperative arthrographic images aid in the decision-making of performing a pelvic osteotomy, these four reviewers were excluded because they displayed poor reliability with plain radiographic evaluation at baseline and it would be indeterminable if the addition of additional arthrogram images was helpful to them. The pairwise comparisons for all 10 raters for radiographic and arthrographic assessment can be found in Appendix 1. Agreement in decision-making using the intraoperative arthrogram was evaluated by estimating the concordance across six raters

with respect to changing their recommendation after reviewing the arthrogram images. Confidence intervals were calculated for each concordance estimate based on 1000 bootstrapped samples. In addition, concordance across attending-level surgeons (n=3) was estimated in contrast to fellow-level surgeons (n=3), and differences by experience were determined by disjoint confidence intervals.

Lastly, inter-rater reliability of important radiographic characteristics of the arthrogram was compared between two attending surgeons of variable clinical experience but with a preference for using intraoperative arthrography (10 years and 1 year of experience, respectively). These two surgeons also identified a series of arthrographic features that were felt to be important for the decision of performing a PO. Important radiographic variables included: Position of the labrum and presence of medial dye pool on stress testing. Position of the labrum was classified into three positions: 1) Upsloping – where the labrum was seen to be blunted and turned up, 2) Flat – where the labrum was seen to be horizontal to Hilgenreiner’s line, and 3) Down sloping – where the labrum was seen to be sharp and followed the contour of the femoral head in a down sloping position (Figure 1A-C). This classification is guided by Heinrich et al.¹¹ Medial dye pooling was assessed by simulating hip instability, by placing the hip in 90 degrees of hip flexion and adduction past midline with posterior directed force to simulate the natural position of sitting in a wheelchair with adduction. Significant medial dye pool was defined as medial dye pooling greater than 50% of the contralateral “normal” hip or in the setting of two abnormal hips, dye pooling greater than one third of the femoral head diameter on the affected hip. Inter-rater agreement between these two surgeons was compared with respect to these radiographic treatment decisions (X-ray and arthrogram), presence of dye pool on arthrogram, and position of the labrum on arthrogram (upsloping, horizontal, or down sloping labrum). Association between dye pool and labrum assessments and arthrogram-based decisions was evaluated using chi-squared tests.

Interpretations for concordance and kappa statistics were based on the scale from Landis and Koch: <0.10, less than chance agreement; 0.10-0.20, slight agreement; 0.21-0.40 fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; and ≥ 0.81 , almost perfect agreement.¹²

Results

Mean age at surgery was 8.3 years (+/- 4 years). The majority (76%) of patients were classified as GMFCS IV or V. A majority (67%) of patients were non-verbal. Most patients were male (67%). Intra-rater agreement on responses 6 months apart was substantial for the three attending surgeons for decisions based on preoperative X-rays (kappa = 0.77, 0.90, 1.00; 95% CI= 0.34-1.00) and final decision based on arthrogram (kappa = 0.67, 0.86, 1.00; 95% CI= 0.20-1.00) (Table 2). Pairwise inter-rater reliability for decisions made based on the radiograph across all six surgeons was fair to almost perfect with kappa values ranging from 0.56 to 0.87 (Table 3).

In 20 of 32 hip evaluations, at least 5 of 6 raters recommended PO in 64% (122/192) of radiographic reviews (Table 4). Upon reviewing the arthrogram, 45% (86/192) of reviews resulted in a decision for PO, a decrease in the recommendation for PO by 21% (40/192).

Among all raters, when we looked at decision-making, we found that those hips that demonstrated a neutral or upsloping labrum had a recommendation for PO in 85% of responses, while a downsloping labrum had a recommendation of PO in 22% of responses. For those cases where a positive dye pool was present, 89% of responses recommended PO.

When stratified by experience level, attending surgeons changed their decisions after interpreting the arthrogram to avoiding a pelvic osteotomy in 35% of cases overall, while fellows changed their decisions in only 7% of cases (Table 5). Across all raters there was fair concordance with the addition of the intraoperative arthrogram (0.44; 95% CI = 0.28-0.61). Stratifying the results by

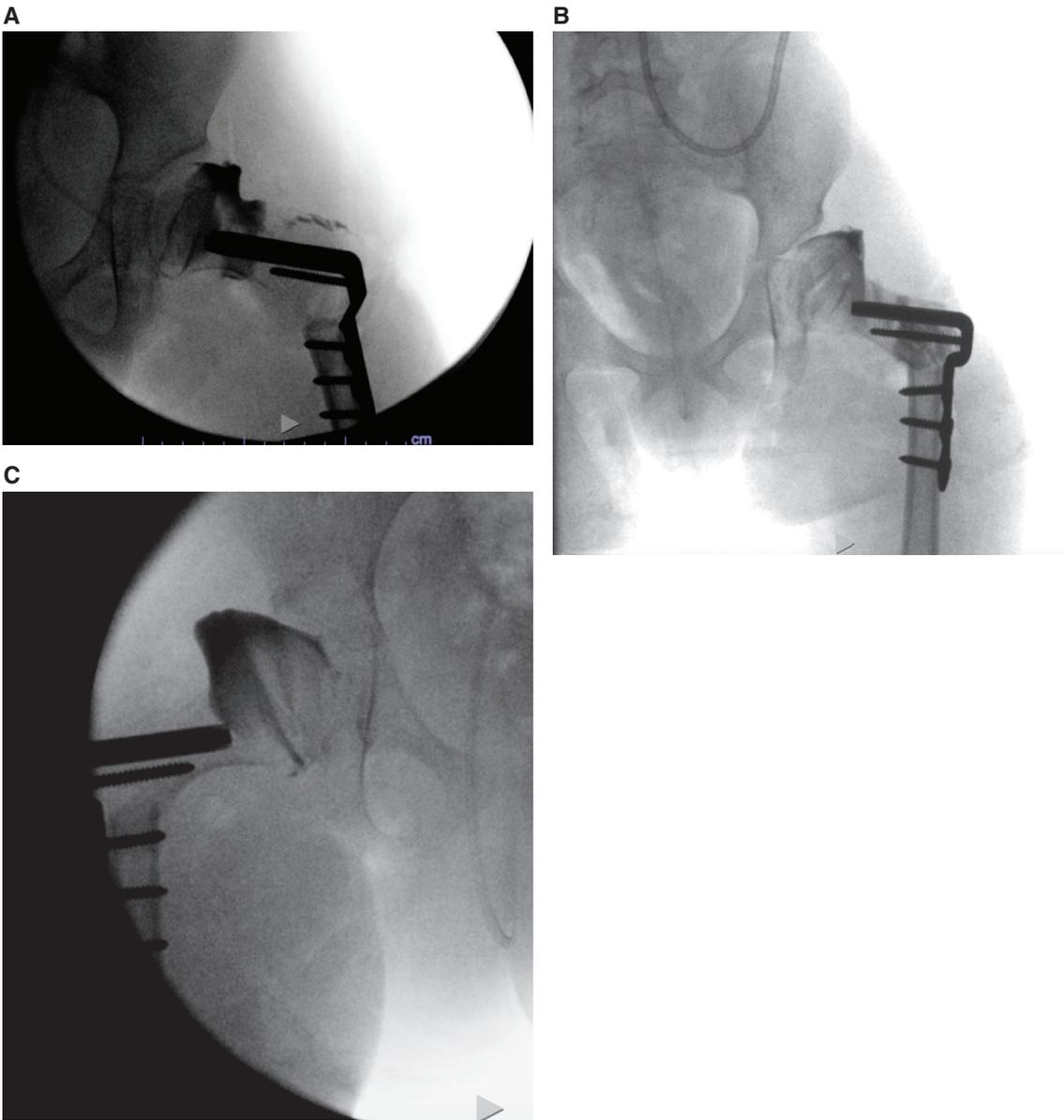


Figure 1. (A) “Downsloping Labrum” (B) “Horizontal Labrum” (C) “Upsloping Labrum”

Table 2. Intra-Rater Reliability Summary

Rater	X-ray		Arthrogram	
	k	(95% CI)	k	(95% CI)
Attending 1	0.90	[0.47, 1.00]	0.67	[0.20, 1.00]
Attending 2	1.00	[0.57, 1.00]	1.00	[0.49, 1.00]
Attending 3	0.77	[0.34, 1.00]	0.86	[0.36, 1.00]

k, kappa coefficient; CI, confidence interval

experience level resulted in attending surgeons making the most changes after arthrogram interpretation with moderate concordance compared to fellows who made fewer changes and demonstrated substantial concordance (concordance=0.46; 95% CI=0.21-0.67 vs. 0.79; 95% CI= 0.62-0.96) (Table 5).

Finally, inter-rater agreement between the two attending surgeons with the greatest experience with intraoperative hip arthrography demonstrated substantial (kappa=0.75;

Table 3. Pairwise Inter-Rater Reliability Kappa Coefficients for All Raters by Experience Level

Rater	Fellows			Attending		
	Fellow 1	Fellow 2	Fellow 3	Attending 1	Attending 2	Attending 3
Fellow 1		0.80	0.87	0.87	0.87	0.78
Fellow 2	0.80		0.67	0.67	0.67	0.56
Fellow 3	0.87	0.67		0.87	0.87	0.65
Attending 1	0.87	0.67	0.87		0.87	0.65
Attending 2	0.87	0.67	0.87	0.87		0.65
Attending 3	0.78	0.56	0.65	0.65	0.65	

95% CI=0.40-1.00) agreement on plain radiographs and almost perfect (kappa=0.84; 95% CI=0.49-1.00) agreement when deciding on performing a PO after arthrography interpretation (Table 6). Further radiographic characterization of the labrum position between the two raters demonstrated substantial agreement (kappa=0.73; 95% CI=0.48-0.98) (Table 6). Agreement between the two raters regarding presence of dye pool was almost perfect (kappa=0.85; 95% CI=0.50-1.00) (Table 6). Both reviewers changed their decision nine times when interpreting the arthrogram and in all nine cases the decision was to not perform a PO.

Discussion

Decision-making related to hip reconstruction for CP related hip dysplasia is challenging.^{2,8} The decision for surgery is often made after reviewing the preoperative physical examination, static radiographs, and three-dimensional imaging when available.^{2,8} While this approach is recognized as the standard of care regarding CP hip reconstruction, we believe that the addition of intraoperative hip arthrography after VDRO may provide supplemental information which has the potential to standardize decision-making and reduce practice variation regarding CP hip reconstruction. In this study, we demonstrate moderate intra-rater and inter-rater reliability for decision-making after arthrogram interpretation among surgeons and present a series of discrete arthrographic features which may play a role in

the future of standardizing decision-making regarding CP hip reconstruction.

The literature is unclear as to which children with CP hip dysplasia should undergo combined pelvic and femoral procedures rather than isolated femoral procedures.¹³ The accepted decision-making algorithm is to broadly start with soft tissue release, add a VDRO, and then decide if a PO is required.² Al-Ghadir et al.¹² reviewed cases of combined femoral and PO surgery versus isolated femoral osteotomy and reported a 25% reoperation rate for femoral osteotomy alone and a 0% reoperation rate for combined pelvic and femoral osteotomies. The authors concluded that results of isolated femoral surgery were inferior to a combined approach. Song et al.¹⁴ reported similar results when comparing combined femoral and pelvic osteotomies to femoral osteotomy alone with a 24% re-subluxation rate with VDRO alone and 13% with a combined approach. However, Chang et al.¹⁵ published a 16% rate of re-subluxation with isolated femoral osteotomy and concluded that femoral osteotomy alone can provide a stable and reduced joint. Based on the variable results in the literature, there is a need to standardize decision-making regarding the inclusion of PO in CP hip reconstruction. Intraoperative hip arthrography has been previously applied with success in other hip diseases to improve outcomes, and we hypothesized that arthrography could also be applied reliably in CP hip reconstruction.

Table 4. Decisions for Osteotomy for All Hips

Hip	X-ray		Arthrogram		Change in		Any	
	Decision for pelvic osteotomy	(%)	Decision for pelvic osteotomy	(%)	Decision for pelvic osteotomy	(%)	Decision for pelvic osteotomy	(%)
Hip 1	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 2	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 3	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 4	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 5	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 6	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 7	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Hip 8	1	(17%)	0	(0%)	1	(17%)	1	(17%)
Hip 9	1	(17%)	0	(0%)	1	(17%)	1	(17%)
Hip 10	1	(17%)	1	(17%)	2	(33%)	2	(33%)
Hip 11	1	(17%)	1	(17%)	2	(33%)	2	(33%)
Hip 12	3	(50%)	3	(50%)	0	(0%)	3	(50%)
Hip 13	5	(83%)	3	(50%)	2	(33%)	5	(83%)
Hip 14	5	(83%)	3	(50%)	2	(33%)	5	(83%)
Hip 15	5	(83%)	1	(17%)	4	(67%)	5	(83%)
Hip 16	5	(83%)	3	(50%)	2	(33%)	5	(83%)
Hip 17	5	(83%)	1	(17%)	4	(67%)	5	(83%)
Hip 18	6	(100%)	4	(67%)	2	(33%)	6	(100%)
Hip 19	6	(100%)	3	(50%)	3	(50%)	6	(100%)
Hip 20	6	(100%)	6	(100%)	0	(0%)	6	(100%)
Hip 21	6	(100%)	6	(100%)	0	(0%)	6	(100%)
Hip 22	6	(100%)	6	(100%)	0	(0%)	6	(100%)
Hip 23	6	(100%)	5	(83%)	1	(17%)	6	(100%)
Hip 24	6	(100%)	5	(83%)	1	(17%)	6	(100%)
Hip 25	6	(100%)	5	(83%)	1	(17%)	6	(100%)
Hip 26	6	(100%)	5	(83%)	1	(17%)	6	(100%)
Hip 27	6	(100%)	6	(100%)	0	(0%)	6	(100%)
Hip 28	6	(100%)	5	(83%)	1	(17%)	6	(100%)
Hip 29	6	(100%)	6	(100%)	0	(0%)	6	(100%)
Hip 30	6	(100%)	2	(33%)	4	(67%)	6	(100%)

Table 4. Continued

	X-ray		Arthrogram		Change in		Any	
Hip 31	6	(100%)	3	(50%)	3	(50%)	6	(100%)
Hip 32	6	(100%)	3	(50%)	3	(50%)	6	(100%)
All hips	122	(64%)	86	(45%)	40	(21%)	124	(65%)

Table 5. Summary of Change in Decision After Arthrogram Review

Agreement in decision-making					
Rater	No. raters	% Changed	<i>k</i>	[95% CI]	Interpretation
Attending 1	1	34%	0.30	[0.00-0.64]	Fair
Attending 2	1	31%	0.37	[0.02-0.71]	Fair
Attending 3	1	38%	0.24	[0.00-0.59]	Fair
Fellow 1	1	9%	0.79	[0.44-1.00]	Substantial
Fellow 2	1	6%	0.87	[0.52-1.00]	Almost perfect
Fellow 3	1	6%	0.87	[0.53-1.00]	Almost perfect
Concordance across rater groups					
Rater	No. raters	% Changed	<i>S</i>	[95% CI]	Interpretation
All surgeons	6	21%	0.44	[0.28-0.61]	Fair
Attending surgeons	3	34%	0.46	[0.21-0.67]	Moderate
Fellow surgeons	3	7%	0.79	[0.62-0.96]	Substantial

k, kappa coefficient; *S*, concordance statistic; *CI*, confidence interval

Table 6. Agreement Between Attending With 10 Years of Arthrography Experience and an Attending With 1 Year of Arthrography Experience

Measurement	<i>k</i>	[95% CI]	Interpretation
Radiograph	0.75	[0.40, 1.09]	Substantial
Arthrogram	0.84	[0.49, 1.20]	Almost perfect
Presence of dye pool	0.85	[0.50, 1.19]	Almost perfect
Position of the labrum	0.73	[0.48, 0.98]	Substantial

k, kappa coefficient; *S*, concordance statistic; *CI*, confidence interval

The results from the current study suggest that surgeons can use intraoperative hip arthrogram data to aid in decision-making regarding the use of a concomitant

PO, potentially making this decision more objective and standardized. We supported our hypothesis, with the finding that attending surgeons modified their

preoperative plans approximately one-third (34%) of the time after reviewing the intraoperative arthrogram images. The majority of the treatment changes resulted in *not performing* a PO. Huh et al.¹⁰ examined their institutional results with VDRO alone or in combination with PO and used an intraoperative arthrogram to determine the need for PO. They demonstrated that even in patients with migration indices >50%, VDRO and stable arthrogram provided adequate stability at a minimum of 2-year follow-up.¹⁰ Despite these promising results, this was a small cohort study and further long-term data is needed to validate the use of intraoperative hip arthrography in the setting of CP hip reconstruction. Upon review, the literature is incomplete regarding which features are important when using intraoperative arthrography when considering CP hip reconstruction. While the additional arthrographic information in this study resulted in a decrease frequency of recommendation for PO, the clinical significance of this decision is unknown and further medium- and long-term study is necessary to validate these intraoperative arthrographic interpretations.

Reliability is the foundation for any radiographic classification system. The reliability of intraoperative arthrography for CP hip reconstruction has not been previously published. Prior to testing the reliability of intraoperative arthrographic interpretation, we needed to test the agreement of our surgeons regarding the interpretation of preoperative plain radiographs. In this study, we found that four of our 10 surgeons demonstrated poor intra-rater reliability as it pertained to plain radiograph decision-making, further strengthening the argument that additional information is needed to make decisions regarding PO in the setting of CP hip reconstruction. All three attending surgeons achieved moderate concordance for PO decision after intraoperative hip arthrogram interpretation (0.46, 0.21-0.67). Interestingly, we saw that the fellows demonstrated much higher concordance (0.79, 0.62-0.96); however, it is important to recognize that the concordance was high because a change in treatment only occurred 7% of the time compared to 34% of the time for attending

surgeons. This finding demonstrates that fellows may not be comfortable interpreting the additional data provided from intraoperative arthrography.

Amongst attending surgeons, there was substantial intra-rater agreement (both X-ray and arthrogram) suggesting that treatment decisions can be standardized based on imaging characteristics rather than subjective surgeon preference. When we investigated the agreement further between two attending surgeons of variable practice experience, but with a clinical preference for using intraoperative hip arthrography, the agreement improved significantly to almost perfect agreement. This data suggests that the use of intraoperative arthrography can provide reliable and reproducible clinical information. While these initial results are promising, further prospective study and characterization through multicenter study are needed to validate the clinical decision-making and demonstrate the longevity of not performing a PO in the setting of CP hip reconstruction.

Intraoperative hip arthrography is a widely accepted modality to help guide the treatment of developmental hip dysplasia but has yet to be accepted in CP hip reconstruction. Heinrich et al.¹¹ described the features of instability in the CP hip, focusing on the lateral acetabular cartilage anlage. Their study described the upsloping or down sloping of the lateral acetabular anlage, which is only apparent after contrast dye is injected into the hip joint. Increasingly upsloping lateral anlage was associated with increased acetabular index and increasing hip instability. Similar to Heinrich et al.,¹¹ we found that often an upsloping labrum was associated with hip instability; however, not all cases of instability demonstrated this labral configuration and the addition of a dynamic assessment is also important.

This study is not without limitations. Due to the nature of the survey, we were only able to report the opinions of the participating surgeons and not what they actually would do clinically. We recognize that the interpretation of intraoperative arthrography is complex and dynamic; however, we have tried to distill elements of the arthrographic interpretation in an effort to aid clinicians

who perform CP hip reconstruction. We recognize that the interpretation of intraoperative hip arthrography involves an element of dynamic assessment which is missing from a written report of this nature; therefore, our compromise was to include static arthrographic images which simulated stability testing (hip flexed, adducted with posterior directed force). We emphasize that this study does not validate whether our clinical decisions regarding the use of PO in the setting of CP hip reconstruction are correct. In order to determine if the decisions to perform PO are clinically appropriate, medium- and long-term prospective, multicenter follow-up is needed. This study can be viewed as a proof of concept, where we demonstrate the potential utility and reliability of intraoperative hip arthrography as it pertains to decision-making regarding the application of PO in CP hip reconstruction. Ultimately, further clinical study is necessary to validate the decision-making presented in this article. We recognize that additional factors contribute to the decision-making when performing a PO, and some of these were not captured in this study. Factors such as age and functional level were not provided to the surgeons, as the study aims were to focus on the radiographic interpretation and assess the associated reproducibility and reliability of interpreting hip arthrography. Finally, we note that utilizing fellows and attendings from only two institutions may overstate the reliability of the arthrogram interpretation and that only using two surgeons to test the classification scheme further undermines the external validity of our conclusions. However, the goal here was to present a proof of concept for the application of intraoperative hip arthrogram assessment in the setting of CP hip reconstruction. Further investigation involving multicenter research study groups, such as the Cerebral Palsy Hip Outcomes Project (CHOP, clinical trials NCT01987882), can be leveraged to study inter-rater and intra-rater reliability in a larger group of surgeons with variable practice patterns to assess the longevity, clinical outcome, and ultimate success associated with using intraoperative hip arthrography in the decision-making for CP hip reconstruction.

In conclusion, intraoperative hip arthrography after VDRO in CP hip reconstruction may be a reliably interpreted data point to guide surgeons regarding the need to add a concomitant PO. However, further clinical study is needed to investigate the validity of this algorithm. Future multicenter research is needed to test the hypotheses presented here regarding the importance of labrum position and the presence of a medial dye pool when considering the use of PO after VDRO in the setting of CP hip reconstruction. It will take long-term clinical and radiographic studies to determine if alterations in management protocols based on intraoperative hip arthrography do in fact improve outcomes for children with hip dysplasia associated with severe CP. Demonstrating reliability in the interpretation of intraoperative hip arthrography in the setting of CP hip reconstruction alone is not sufficient to change practice but has potential to standardize clinical decisions and decrease existing practice variation.

Disclaimer

The authors report no conflicts of interest related to this manuscript.

References

1. Bagg MR, Farber J, Miller F. Long-term follow-up of hip subluxation in cerebral palsy patients. *J Pediatr Orthop*. 1993;13(1):32-36.
2. Kerr Graham H, Selber P. Musculoskeletal aspects of cerebral palsy. *J Bone Joint Surg Br*. 2003;85(2):157-166.
3. Hagglund G, Alriksson-Schmidt A, Lauge-Pedersen H, et al. Prevention of dislocation of the hip in children with cerebral palsy: 20-year results of a population-based prevention programme. *Bone Joint J*. 2014;96-B(11):1546-1552.
4. Terjesen T. The natural history of hip development in cerebral palsy. *Dev Med Child Neurol*. 2012;54(10):951-957.
5. Hagglund G, Lauge-Pedersen H, Wagner P. Characteristics of children with hip displacement in cerebral palsy. *BMC Musculoskelet Disord*. 2007;8:101.
6. Ramstad K, Terjesen T. Hip pain is more frequent in severe hip displacement: a population-based study of 77 children with cerebral palsy. *J Pediatr Orthop B*. 2016;25(3):217-221.
7. Knapp DR, Jr., Cortes H. Untreated hip dislocation in cerebral palsy. *J Pediatr Orthop*. 2002;22(5):668-671.
8. Flynn JM, Miller F. Management of hip disorders in patients with cerebral palsy. *J Am Acad Orthop Surg*. 2002;10(3):198-209.
9. Tennant SJ, Eastwood DM, Calder P, et al. A protocol for the use of closed reduction in children with developmental dysplasia of the hip incorporating open psoas and adductor releases and a short-leg cast: Mid-term outcomes in 113 hips. *Bone Joint J*. 2016;98-B(11):1548-1553.

10. Huh K, Rethlefsen SA, Wren TA, et al. Surgical management of hip subluxation and dislocation in children with cerebral palsy: isolated VDRO or combined surgery? *J Pediatr Orthop.* 2011;31(8):858-863.
11. Heinrich SD, MacEwen GD, Zembo MM. Hip dysplasia, subluxation, and dislocation in cerebral palsy: an arthrographic analysis. *J Pediatr Orthop.* 1991;11(4):488-493.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
13. Al-Ghadir M, Masquijo JJ, Guerra LA, et al. Combined femoral and pelvic osteotomies versus femoral osteotomy alone in the treatment of hip dysplasia in children with cerebral palsy. *J Pediatr Orthop.* 2009;29(7):779-783.
14. Song HR, Carroll NC. Femoral varus derotation osteotomy with or without acetabuloplasty for unstable hips in cerebral palsy. *J Pediatr Orthop.* 1998;18(1):62-68.
15. Chang FM, May A, Faulk LW, et al. Outcomes of isolated varus derotational osteotomy in children with cerebral palsy hip dysplasia and predictors of resubluxation. *J Pediatr Orthop.* 2018;38(5):274-278.

Appendix

Appendix 1. Pairwise Inter-Rater Reliability Kappa Estimates for All Participants

Rater	Radiograph									
	Experience 1					Experience 2	Experience 3		Experience 4	
	Rater 1	Rater 2	Rater 3	Rater 7	Rater 8	Rater 6	Rater 5	Rater 9	Rater 4	Rater 10
1		0.18	0.24	0.30	0.48	0.30	0.30	0.00	0.35	0.76
2	0.18		0.80	0.87	0.38	0.87	0.87	0.78	0.54	0.25
3	0.24	0.80		0.67	0.44	0.67	0.67	0.56	0.61	0.31
7	0.30	0.87	0.67		0.25	0.87	0.87	0.65	0.42	0.24
8	0.48	0.38	0.44	0.25		0.25	0.37	0.18	0.43	0.71
6	0.30	0.87	0.67	0.87	0.25		0.87	0.65	0.55	0.37
5	0.30	0.87	0.67	0.87	0.37	0.87		0.65	0.55	0.37
9	0.00	0.78	0.56	0.65	0.18	0.65	0.65		0.58	0.06
4	0.35	0.54	0.61	0.42	0.43	0.55	0.55	0.58		0.42
10	0.76	0.25	0.31	0.24	0.71	0.37	0.37	0.06	0.42	