

Inter-observer and intra-observer reliability in the radiographic diagnosis of avascular necrosis of the femoral head following reconstructive hip surgery in children with cerebral palsy

Kim Hesketh¹ · Wudbhav Sankar² · Benjamin Joseph³ · Unni Narayanan⁴ · Kishore Mulpuri^{1,5}

Received: 29 January 2015 / Accepted: 26 February 2016 / Published online: 14 March 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Purpose The incidence of avascular necrosis (AVN) following reconstructive hip surgery in cerebral palsy (CP) ranges from 0 to 69 % in the current literature. The purpose of this study was to determine the inter- and intra-observer reliability of radiographically diagnosing AVN in children with CP after hip surgery.

Methods A retrospective review of 65 children with CP who had reconstructive hip surgery between 2009 and 2012 at BC Children's Hospital was completed. Anterior–posterior and lateral radiographs were presented to four pediatric orthopaedic surgeons over two rounds. Surgeons were asked to review the set of unidentified radiographs and comment 'yes' or 'no' for the presence of AVN. Two weeks later the same set of radiographs was sent in a different order and the surgeons were again asked to comment on AVN. Inter- and intra-observer reliability was determined using kappa statistics.

Results The intra-observer reliability ranged from 0.65 to 0.88 with an average score of 0.76. Inter-observer reliability showed greater variability, ranging from 0.41 to 0.77 with an average score of 0.56 across all surgeons.

Conclusions Although the intra-rater reliability produced a strength of “good” and the inter-rater reliability a strength of “moderate” agreement, the variability within these scores is clinically important as it demonstrates the difficulty in identifying AVN. This may explain the variability in AVN that is reported in the literature. The need for further education and research in the diagnosis of AVN in children with CP who have undergone reconstructive hip surgery is clinically necessary.

Keywords Cerebral palsy · Avascular necrosis · AVN · Inter-observer and intra-observer reliability · Reconstructive hip surgery for children with CP

Introduction

Hip displacement is the lateral migration of the femoral head from the acetabulum and is one of the most common musculoskeletal problems in children with cerebral palsy (CP) [1–4]. It can result in pain, reduced range of motion (ROM), skin breakdown, difficulty with perineal hygiene, and difficulty with seating and positioning, and can contribute to the development of scoliosis and pelvic obliquity [3, 7, 9, 10].

Reconstructive surgery for displaced hips has become the standard of care for children with CP [2, 17, 21]. This typically involves a varus derotational osteotomy (VDRO) of the femur and muscle releases of the adductors and hip flexors. An open reduction and pelvic osteotomy may also be required depending upon the degree of displacement. Goals of the surgical intervention include maintaining or attaining a flexible, aligned, and painless hip [24].

Avascular necrosis (AVN) of the femoral head is a potential complication when undertaking hip surgery in

✉ Kishore Mulpuri
kulpuri@cw.bc.ca

¹ Department of Orthopaedic Surgery, BC Children's Hospital, 1D.66, 4480 Oak St, Vancouver, BC V6H 3V4, Canada

² The Children's Hospital of Philadelphia, Philadelphia, PA, USA

³ Paediatric Orthopaedic Service, Manipal, Karnataka, India

⁴ Hospital for Sick Children, Toronto, ON, Canada

⁵ Department of Orthopedics, University of British Columbia, Vancouver, BC, Canada

children with CP. The exact pathophysiology is unknown, but may be that the blood supply to the hip is disrupted, leading to cellular necrosis. For children with CP following reconstructive surgery, it is hypothesized that AVN may occur from disturbance in the blood supply to the femoral head during soft tissue releases in the groin, injury to vessels that ascend the femoral head during the VDRO, or excessive pressure on the femoral head when there is insufficient shortening [16, 18]. There is no detailed description of AVN in this population, making it difficult to classify. However, it is well documented in developmental dysplasia of the hip (DDH). Kalamachi and MacEwen is the most common classification in this population [12]. It was first described in 1980 and modified by Kruczynski in 1996 [15]. The severity of AVN in this population can range from mild, involving a hypoplastic epiphysis of the femoral head, to severe, with damage to the growth plate resulting in growth disturbances and misalignment of the proximal femur [19]. The clinical effects of AVN are also not well documented in the CP population. However, idiopathic AVN has been reported to cause pain, limited ROM, and an altered gait pattern [13]. In children with developmental dysplasia of the hip, Roposch et al. reported there is a large variation in the clinical presentation of AVN. Severe forms lead to hip pain and premature debilitating osteoarthritis [19]. Mild forms may have minimal deformity and dysfunction [19]. Pain and the associated clinical relevance of AVN may be underestimated in children with CP due to the inherent difficulty in quantifying pain and symptoms for this challenging patient population.

The incidence of AVN after hip surgery in children with CP varies in the literature from 0 to 69 % [2, 4, 6, 7, 11, 14, 16, 18, 22, 23]. The disparity in the reported rates of AVN may be due to a number of reasons. First, the differences may be related to the variability in risk for AVN development across series. Second, inadequate follow-up of patients may contribute to under- or over-reported AVN. Finally, and most relevant to our study, the disparity in reported rates of AVN may be related to differences in diagnosing and identifying AVN radiographically. To date, there is no known evidence documenting the reliability amongst orthopaedic surgeons in identifying and diagnosing AVN in children with CP following hip surgery. The purpose of this study was to determine the inter- and intra-observer reliability of pediatric orthopaedic surgeons radiographically diagnosing AVN in children with CP after hip surgery.

Materials and methods

A retrospective review of children, ages 2–18 years, with CP who had reconstructive hip surgery for hip displacement between 2009 and 2012 at BC Children's Hospital

was completed. Patients were excluded if there was no diagnosis of CP or they had undergone salvage surgery. A sample of 65 anterior–posterior and frog-lateral radiographs was selected for the study.

Radiographs were taken by a single radiology department in a standardized fashion and collected at post-surgical follow-up visits. Anterior–posterior and frog-lateral radiographs for each case were presented in random order to four pediatric orthopaedic surgeons over two rounds. The radiographs were saved from an electronic database then pasted and randomized in a Word document for the surgeons to evaluate. In the first round surgeons were asked to independently review the set of unidentified radiographs and comment 'yes' or 'no' for the presence of AVN. Two weeks later the same set of radiographs was sent in a different order and the surgeons were again asked to comment 'yes' or 'no' for the presence of AVN. Because the diagnosis of AVN is a highly subjective process and there is no standardized method for objective diagnosis in this population, no defined criteria to assess for AVN was provided. Diagnosis was left to the subjective interpretation of each surgeon, to reflect the current practice for AVN diagnosis. Each surgeon was experienced with the pediatric hip.

The inter- and intra-observer reliability was determined using kappa statistics as described in Altman [1], indicating that agreement <0.2 is poor, 0.21–0.40 is fair, 0.41–0.60 is moderate, 0.61–0.8 is good, and 0.81–1.00 is very good.

Results

Sixty-five radiographs of both left and right hips were evaluated to give a total of 130 hips. Since each hip was individually classified by the presence or absence of AVN, reliability should not be affected by side other than by chance or bias. Thus, each hip can be considered unique and reliability values are reported across all 130 hips. A secondary analysis was performed to examine right and left hips separately.

Total AVN diagnosed

To quantify how many hips in total had AVN, those in which at least two surgeons identified AVN were considered positive. In trial one, 23 hips (18 %; 10 right, 13 left) were identified as having AVN, while in trial two, 21 hips (16 %; 8 right, 13 left) were identified. Of the right hips, seven were identified in both trials, with three and one hips independently identified in trials one and two, respectively. Of the left hips, 11 were identified in both trials with two hips independently identified in each. Between the two trials, a total of 11 right hips and 15 left hips in 25 different patients were diagnosed with AVN (Table 1). There were

Table 1 AVN-positive hips—number of surgeons reporting

Subject no.	No. of surgeons reporting AVN			
	Right hip		Left hip	
	Trial 1	Trial 2	Trial 1	Trial 2
1	–	–	4	4
2	–	–	3	3
3	4	3	3	3
11	–	–	2 ^a	1
15	4	4	–	–
17	2 ^a	1	–	–
19	–	–	2	2
20	–	–	4	4
21	–	–	2	3
22	–	–	4	4
23	–	–	4	4
27	–	–	1	3 ^a
28	4	4	–	–
29	2 ^a	1	–	–
30	4	3	–	–
33	3	3	–	–
34	–	–	2	3
35	1	2 ^a	–	–
47	–	–	0	3 ^a
48	4	4	–	–
49	4	4	–	–
52	–	–	2 ^a	1
53	–	–	4	3
54	–	–	2	2
65	2 ^a	1	–	–

^a Indicates hip met AVN criteria for only the specified trial; *dashes* indicate hip was classified as AVN-negative

eight hips in which all four surgeons positively identified AVN, and eight hips in which only two surgeons identified AVN. Example radiographs of these optimal and suboptimal agreements are shown in Fig. 1a, b, respectively. In comparison, there were 81 hips (41 right, 40 left) in which all four surgeons reported no AVN in either trial.

Intra-observer reliability

The intra-observer reliability showed variability among the surgeons, ranging from 0.65 to 0.88 with an overall mean score of 0.76 (Table 2). Across all hips, the intra-observer reliability of all four surgeons was relatively high. Surgeon 4 had very good intra-observer reliability (0.88), and surgeons 1, 2, and 3 had good agreement (0.80, 0.72, and 0.65, respectively). Looking at the right and left hips separately, the intra-observer reliability was more variable. Surgeon 4 demonstrated very good intra-observer reliability for both right and left. Surgeon 1 had very good agreement for the

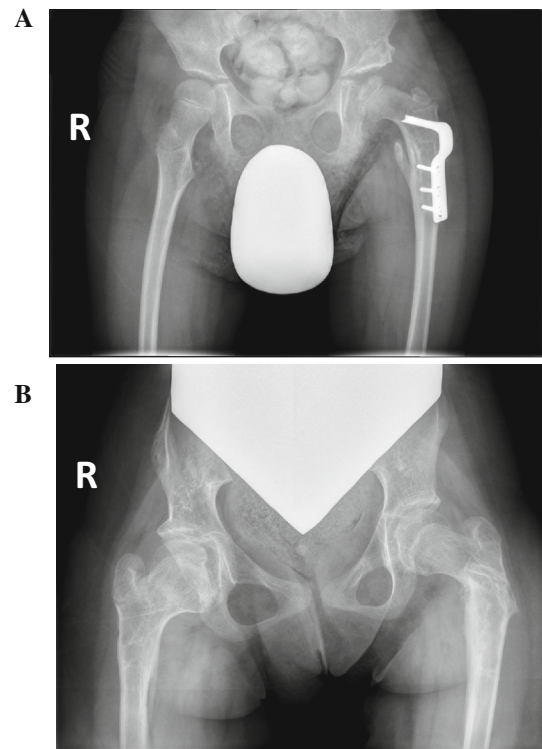


Fig. 1 Optimal and suboptimal agreement between paediatric orthopaedic surgeons on the presence of avascular necrosis (AVN) of the femoral head in children with cerebral palsy. Anteroposterior radiographs of the hips were identified to have AVN by **a** 4/4 surgeons (*left hip*), and **b** 2/4 surgeons (*right hip*)

Table 2 Intra-observer reliability—kappa values and interpretation

	Kappa values			
	Surgeon 1	Surgeon 2	Surgeon 3	Surgeon 4
All hips	0.80 (G)	0.72 (G)	0.65 (G)	0.88 (V)
Right hip	0.84 (V)	0.69 (G)	0.69 (G)	0.82 (V)
Left hip	0.77 (G)	0.75 (G)	0.60 (M)	0.93 (V)

Altman [1] kappa values: V, very good; G, good; M, moderate

right hip and good agreement for the left. Surgeon 2 demonstrated good agreement for both sides, while surgeon 3 had good agreement for the right and moderate agreement for the left (Table 2).

Inter-observer reliability

The inter-observer reliability was lower and showed greater variability than intra-observer reliability, ranging from 0.41 to 0.77 (Table 3) with an overall mean score of 0.56 (moderate). Surgeon pairs 1–2, 2–3, 2–4 and 3–4 scored moderate across all hips with means of 0.50, 0.47, 0.43, and 0.57, respectively. Surgeon pairs 1–3 and 1–4 scored good across all hips with means of 0.69 and 0.72,

Table 3 Inter-observer reliability—kappa values and interpretation

Surgeon pair	Kappa values		
	Trial 1	Trial 2	Mean
1–2	0.46 (M)	0.55 (M)	0.50 (M)
1–3	0.63 (G)	0.74 (G)	0.69 (G)
1–4	0.77 (G)	0.67 (G)	0.72 (G)
2–3	0.49 (M)	0.46 (M)	0.47 (M)
2–4	0.45 (M)	0.41 (M)	0.43 (M)
3–4	0.62 (G)	0.53 (M)	0.57 (M)

Altman [1] kappa values: *G* good, *M* moderate

Table 4 Inter-observer reliability by hip—kappa values and interpretation

Surgeon pair	Kappa values		
	Trial 1	Trial 2	Mean
Right hip			
1–2	0.48 (M)	0.48 (M)	0.48 (M)
1–3	0.74 (G)	0.63 (G)	0.69 (G)
1–4	0.84 (V)	0.82 (V)	0.83 (V)
2–3	0.5 (M)	0.47 (M)	0.49 (M)
2–4	0.37 (F)	0.48 (M)	0.43 (M)
3–4	0.72 (G)	0.63 (G)	0.67 (G)
Left hip			
1–2	0.44 (M)	0.59 (M)	0.52 (M)
1–3	0.53 (M)	0.83 (V)	0.68 (G)
1–4	0.72 (G)	0.57 (M)	0.65 (G)
2–3	0.49 (M)	0.45 (M)	0.47 (M)
2–4	0.54 (M)	0.36 (F)	0.45 (M)
3–4	0.53 (M)	0.43 (M)	0.48 (M)

Altman [1] kappa values: *V* very good, *G* good, *M* moderate, *F* fair

respectively. Looking at right and left hips separately, inter-observer reliability had a larger variation but a higher mean in the right hip compared to the left (Table 4). Overall means for surgeon pairs for the right and left hip were 0.60 and 0.54, respectively, both a score of moderate.

Discussion

The frequency of AVN in children with CP after reconstructive hip surgery varies in the literature from 0 to 69 % [2, 4, 6, 7, 11, 14, 16, 18, 22, 23]. The diagnosis of AVN is a highly subjective process dependent upon the individual surgeon and there is no standardized, rigorous method for objective diagnosis. One possible reason for the incidence variability in the literature may be the ability to identify and report AVN. The purpose of this study was to

determine the inter- and intra-observer reliability of orthopaedic surgeons for the radiographic diagnosis of AVN in children with CP after hip surgery. Although the intra-observer reliability produced a strength of good agreement and the inter-observer reliability produced a strength of moderate agreement, the variability within these scores is clinically important as it demonstrates the difficulty in identifying AVN. These results may account for much of the disparity in the literature in the reported frequency of AVN in children with CP after reconstructive hip surgery.

A study completed by Davidson et al. investigating the reliability of diagnosing AVN in children with slipped capital femoral epiphysis (SCFE) reported very high inter- and intra-observer reliability [5]. The discrepancy when compared to the identification of AVN in children with CP indicates there may be a specific difficulty in diagnosing AVN within the CP population.

There are a few limitations with this study. No standard definition or classification scheme of AVN was given to the surgeons. This reflects how AVN is reported in the current literature; there are no clear guidelines used when identifying and reporting AVN after hip surgery in children with CP. Our study evaluated the reliability amongst experienced surgeons only. Within this small sample of surgeons, those with the most clinical experience had better intra- and inter-observer reliability. Including fellows and residents, who are often involved in identifying post-operative complications, may have led to even greater variability in the identification of AVN. This study was limited to the radiographic diagnosis of AVN and did not include the use of magnetic resonance imaging (MRI). Using the Ficat classification scheme of idiopathic AVN, Schmitt-Sody et al. found an inter-observer mean reliability coefficient of 0.36 for radiographs and 0.37 for MRI [20]. Ficat reported that early signs of AVN are not visible on radiographs [8]. Although MRI offers a more detailed view of the hip, these examinations are not used on a routine basis with our patient population and are therefore less clinically relevant.

To identify risk factors and decrease the incidence of AVN, studies evaluating the outcome of reconstructive hip surgery in children with CP need to investigate and report the incidence of AVN. A classification system to identify and quantify the severity of AVN in this population is also necessary. The results of this study suggest that the reliability of radiographic diagnosis of AVN in children with CP after reconstructive hip surgery is “moderate”. These results may explain the current variability in the frequency of AVN reported in the literature. In order to improve the reliability of the diagnosis of AVN there needs to be a clear definition and classification system used. The need for further education and research in the diagnosis of AVN in

children with CP who have undergone reconstructive hip surgery is clinically necessary.

Compliance with ethical standards

No funding was received for this study

Conflict of interest Kim Hesketh, Benjamin Joseph, Unni Narayanan and Kishore Mulpuri declare no conflict of interest. Wudbhav Sankar has received royalties for an edited text from the publisher Lippincott Wolters Kluwer.

Ethical approval 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.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Altman D (1991) *Practical Statistics for Medical Research*. Chapman and Hall, London
- Canavese F, Gomez H, Kaelin A, Ceroni D, de Coulon G (2013) Percutaneous pelvic osteotomy and intertrochanteric varus shortening osteotomy in nonambulatory GMFCS level IV and V cerebral palsy patients: preliminary report on 30 operated hips. *J Pediatr Orthop B* 22:1–7
- Connelly A, Flett P, Graham HK, Oates J (2009) Hip surveillance in Tasmanian children with cerebral palsy. *J Paediatr Child Health* 45:437–443
- Davids JR, Gibson TW, Pugh LI, Hardin JW (2013) Proximal femoral geometry before and after varus rotational osteotomy in children with cerebral palsy and neuromuscular hip dysplasia. *J Pediatr Orthop* 33:182–189
- Davidson D, Ghag A, Beauchamp RD, Tredwell SJ, Mulpuri K (2010) Inter-observer and intra-observer agreement in the radiographic diagnosis of avascular necrosis of the femoral head following slipped capital femoral epiphysis. *J Child Orthop* 4:327–330
- Dhawale AA, Karatas AF, Holmes L, Rogers KJ, Dabney KW, Miller F (2013) Long-term outcome of reconstruction of the hip in young children with cerebral palsy. *Bone Joint J* 95-B:259–265
- Dobson F, Boyd RN, Parrott J, Nattrass GR, Graham HK (2002) Hip surveillance in children with cerebral palsy. Impact on the surgical management of spastic hip disease. *J Bone Joint Surg Br* 84:720–726
- Ficat RP (1985) Idiopathic bone necrosis of the femoral head. Early diagnosis and treatment. *J Bone Joint Surg Br* 67:3–9
- Hagglund G, Andersson S, Duppe H, Lauge-Pedersen H, Nordmark E, Westbom L (2005) Prevention of dislocation of the hip in children with cerebral palsy. The first ten years of a population-based prevention programme. *J Bone Joint Surg Br* 87:95–101
- Hagglund G, Lauge-Pedersen H, Wagner P (2007) Characteristics of children with hip displacement in cerebral palsy. *BMC Musculoskelet Disord* 8:101
- Huh K, Rethlefsen SA, Wren TA, Kay RM (2011) Surgical management of hip subluxation and dislocation in children with cerebral palsy: isolated VDRO or combined surgery? *J Pediatr Orthop* 31:858–863
- Kalamchi A, MacEwen GD (1980) Avascular necrosis following treatment of congenital dislocation of the hip. *J Bone Joint Surg Am* 62:876–888
- Khalife R, Ghanem I, El Hage S, Dagher F, Kharrat K (2010) Risk of recurrent dislocation and avascular necrosis after proximal femoral varus osteotomy in children with cerebral palsy. *J Pediatr Orthop B* 19:32–37
- Koch A, Jozwiak M, Idzior M, Molinska-Glura M, Szulc A (2015) Avascular necrosis as a complication of the treatment of dislocation of the hip in children with cerebral palsy. *Bone Joint J* 97-B:270–276
- Kruczynski J (1996) Avascular necrosis of the proximal femur in developmental dislocation of the hip. Incidence, risk factors, sequelae and MR imaging for diagnosis and prognosis. *Acta Orthop Scand Suppl* 268:1–48
- McNerney NP, Mubarak SJ, Wenger DR (2000) One-stage correction of the dysplastic hip in cerebral palsy with the San Diego acetabuloplasty: results and complications in 104 hips. *J Pediatr Orthop* 20:93–103
- Miller F, Girardi H, Lipton G, Ponzio R, Klaumann M, Dabney KW (1997) Reconstruction of the dysplastic spastic hip with periliac pelvic and femoral osteotomy followed by immediate mobilization. *J Pediatr Orthop* 17:592–602
- Mubarak SJ, Valencia FG, Wenger DR (1992) One-stage correction of the spastic dislocated hip. Use of pericapsular acetabuloplasty to improve coverage. *J Bone Joint Surg Am* 74:1347–1357
- Roposch A, Wedge JH, Riedl G (2012) Reliability of Bucholz and Ogden classification for osteonecrosis secondary to developmental dysplasia of the hip. *Clin Orthop Relat Res* 470:3499–3505
- Schmitt-Sody M, Kirchoff C, Mayer W, Goebel M, Jansson V (2008) Avascular necrosis of the femoral head: inter- and intraobserver variations of Ficat and ARCO classifications. *Int Orthop* 32:283–287
- Shore BJ, Yu X, Desai S, Selber P, Wolfe R, Graham HK (2012) Adductor surgery to prevent hip displacement in children with cerebral palsy: the predictive role of the Gross Motor Function Classification System. *J Bone Joint Surg Am* 94:326–334
- Song HR, Carroll NC (1998) Femoral varus derotation osteotomy with or without acetabuloplasty for unstable hips in cerebral palsy. *J Pediatr Orthop* 18:62–68
- Stasikelis PJ, Ridgeway SR, Pugh LI, Allen BL Jr (2001) Epiphyseal changes after proximal femoral osteotomy. *J Pediatr Orthop B* 10:25–29
- Zarrinkalam R, Rice J, Brook P, Russo RN (2011) Hip displacement and overall function in severe cerebral palsy. *J Pediatr Rehabil Med* 4:197–203