Inter-observer and intra-observer agreement in the radiographic diagnosis of avascular necrosis of the femoral head following slipped capital femoral epiphysis

Darin Davidson · Arvinder Ghag ·
Richard D. Beauchamp · Stephen J. Tredwell ·
Kishore Mulpuri

Received: 7 April 2010 / Accepted: 4 June 2010 / Published online: 24 June 2010
© EPOS 2010

Abstract

Introduction Slipped capital femoral epiphysis (SCFE) is a common pediatric hip disorder. Avascular necrosis (AVN) of the femoral head is a devastating complication of SCFE. The frequency of this complication reported in the literature has been variable. It was the objective of this study to estimate the inter- and intra-observer agreement between two experienced pediatric orthopaedic surgeons for the radiographic diagnosis of AVN following SCFE.

Methods A retrospective review of all cases of SCFE treated at our center between 1995 and 2005 was performed. All cases of AVN and a random sample of 19 of the remaining cases were selected for study. The most recent anteroposterior and lateral radiographs were presented to two experienced pediatric orthopaedic surgeons in a random order. Inter-observer reliability was determined by calculating the kappa statistic to assess for clinical agreement. Each observer repeated this process two weeks after the initial review.

Results There were a total of 103 cases of SCFE, of which four were diagnosed with AVN. The inter-observer

agreement in the first trial was 0.79. The intra-observer agreement for the first observer was 0.9 and for the second observer, it was 0.88.

Conclusion The agreement, both inter- and intra-observer, for the radiographic diagnosis of AVN amongst adolescents with previous SCFE is very high. The results of this study suggest that the reported discrepancy of AVN in the literature following SCFE is not likely due to the lack of inter- and intra-observer agreement.

Keywords Avascular necrosis ·
Slipped capital femoral epiphysis

Introduction

Slipped capital femoral epiphysis (SCFE) is a common pediatric hip condition, characterized by anterior and superior displacement, through the physis, of the proximal femur relative to the proximal femoral epiphysis. One of the most devastating potential complications of this condition is avascular necrosis (AVN) of the femoral head. Though the risk of AVN following a stable SCFE, as defined by the ability to weight-bear, has been reported to be very low, the risk associated with an unstable SCFE, as defined by an inability to weight-bear, has been reported to be 47% [1]. There have been other reports in the literature which have reported a lower risk of AVN following an unstable SCFE. Studies by Rattey et al. [2] and Kennedy et al. [3] both encountered a 15% incidence of AVN following an unstable SCFE. The cause of this discrepancy is unclear and, as a result, the risk of AVN in the population of adolescents with an unstable SCFE remains elusive. Given the significant morbidity which is frequently associated with this

D. Davidson · A. Ghag · R. D. Beauchamp ·
S. J. Tredwell · K. Mulpuri
Department of Orthopaedics, University of British Columbia,
3114-910 W. 10th Avenue, Vancouver, BC V5Z 4E3, Canada

R. D. Beauchamp · S. J. Tredwell
Department of Pediatric Orthopaedics, British Columbia
Children's Hospital, A200-4480 Oak Street,
Vancouver, BC V6H 3V4, Canada

K. Mulpuri (✉)
Department of Pediatric Orthopaedics,
British Columbia Children's Hospital,
A208A-4480 Oak Street, Vancouver,
BC V6H 3V4, Canada
e-mail: kmulpuri@cw.bc.ca

complication, it would be of benefit to further define the risk of AVN.

There are several potential explanations for discrepancies in the risk of a complication reported in the literature. There is the potential for measurement error, the use of different definitions of the complication between studies, the lack of agreement in the diagnosis of the complication, or differences in the populations studied. The cause of the discrepancy between the risk of AVN following an unstable SCFE reported in the literature remains unclear. Further understanding of the potential causes of the differing reported risks of AVN following an unstable SCFE may help in the determination of the true risk and, as such, lead to an improved understanding of the association between SCFE and AVN. In order to investigate the potential for the lack of agreement in the diagnosis of AVN as a contributor to the discrepancy in the literature, it was the purpose of this study to estimate the inter-observer and intra-observer agreement in the radiographic diagnosis of AVN following SCFE. It was hoped that such evidence would contribute towards an increased understanding of the discrepancy in the frequency of AVN following unstable SCFE between studies in the literature.

Methods

This study consisted of an inter-observer and intra-observer agreement study of the radiographic diagnosis of AVN following SCFE. The study population was determined on the basis of a retrospective review of all cases of SCFE treated at our center between 1995 and 2005. From this population, all cases which were diagnosed as having AVN were included in the study. The diagnosis of AVN was made, in each case in this study, on the basis of radiographs. Though this may have resulted in a lower number of cases of AVN than may have actually occurred, this would not have affected the agreement regarding the radiographic diagnosis of AVN, which was the objective of the study. Agreement of either bone scan or magnetic resonance imaging (MRI) diagnosis of AVN was beyond the scope of this study. The sole criterion for the definition of AVN was diagnosis of the complication in the clinical records. In order to provide controls for the estimation of agreement, a random sample of 19 of the remaining cases was taken using a random number generator.

The most recently acquired anteroposterior and lateral radiographs for each patient included in the study were digitized and presented in a random order, using a random number generator, to two experienced, fellowship-trained pediatric orthopaedic surgeons. No defined criteria were used by the observers to assess the presence of AVN on the radiographs included in the study. The cases of AVN

included in the study were defined on the basis of the diagnosis having been made in the clinical records. The radiographs which were studied were obtained at a mean of 19.6 months after treatment of the SCFE. Each surgeon independently assessed the radiographs and recorded which ones he believed to be consistent with AVN. The procedure was repeated two weeks later. The radiographs were presented in a blinded, random order and were interpreted independently by the observers. The observers were not informed of how many cases of AVN were included in the sample, nor were they given any details regarding the patients or classification of the SCFE.

Inter-observer and intra-observer agreement were calculated using the kappa statistic, which determines the agreement beyond that which would be expected due to chance [4]. The kappa statistic was determined for each observer comparing between the first and second observation in order to estimate the intra-observer agreement and between the two observers for the first trial in order to estimate the inter-observer agreement. The interpretation of the kappa values was performed according to the criteria of Landis and Koch [4], who recommended that a value greater than 0.8 be considered ‘almost perfect’, between 0.6 and 0.8 ‘substantial’, between 0.4 and 0.6 ‘moderate’, between 0.2 and 0.4 ‘fair’, between 0 and 0.2 ‘slight’, and between 0 and -1 ‘poor’.

Results

Over the time horizon of the study, there were a total of 103 cases of SCFE treated at our center. Of these, four were diagnosed with AVN of the femoral head (Fig. 1) and the remaining 99 were not diagnosed with AVN (Fig. 2). Each case of AVN demonstrated typical findings of femoral head fragmentation and collapse. The first observer recorded eight cases in the first trial and seven cases in the second trial. The intra-observer agreement kappa value was 0.9. The second observer recorded six cases of AVN in the first trial and five in the second trial. The intra-observer agreement kappa value was 0.88. The inter-observer agreement was determined in the first trial and the kappa value was 0.79.

Discussion

The objective of this study was to estimate the inter-observer and intra-observer agreement in the radiographic diagnosis of AVN following SCFE. The results of this study indicated that, according to the criteria of Landis and Koch [4], the intra-observer agreement for both observers was ‘almost perfect’ and the inter-observer agreement was

Fig. 1 **a** Anteroposterior and **b** frog-lateral radiographs of a patient with a previous slipped capital femoral epiphysis (SCFE) who was diagnosed with avascular necrosis (AVN) of the femoral head

Fig. 2 **a** Anteroposterior and **b** frog-lateral radiographs of a patient with a previous SCFE who was not diagnosed with AVN of the femoral head

‘substantial’. The importance of this data relates to the interpretation of the differences in the reported risk of AVN following unstable SCFE as reported to be 47% by

Loder et al. [1] compared to the 15% risk reported in other studies [2, 3]. On the basis of the results of this study, it would appear to be unlikely that this discrepancy would have occurred solely on the basis of disagreement between observers regarding the radiographic diagnosis of AVN. It was beyond the scope of the present study to investigate other potential explanations for the difference, such as different patient populations, measurement error, different levels of experience amongst the observers, or different time points for obtaining the radiographs.

A recent study investigated inter- and intra-observer agreement using observers of differing specialty and training level, of the Ficat [5] classification for suspected cases of AVN of the femoral head [6]. The inter-observer agreement using radiographs was 0.36 and for MRI, it was 0.37. The intra-observer agreement on radiographs was 0.53 and for MRI, it was 0.50. Though the agreement results in this study demonstrated greater inter- and intra-observer agreement, it is difficult to compare the results of the study due to differences in the patient population.

The potential limitations of this study include the use of radiographs, and not higher level imaging, for the assessment of the agreement of the diagnosis of AVN. Though MRI and bone scan may be utilized in the diagnosis of AVN, the agreement parameters with these modalities require separate study. It is interesting to note, however, that inter- and intra-observer agreement in Ficat [5] classification between radiographs and MRI was not substantially different and, in fact, there was improved agreement using radiographs [6]. The relatively few number of cases of AVN relative to the total number of cases presented to the observers may have been a source of bias. As both observers indicated a greater number of cases of AVN than were diagnosed, there is the possibility that the results were biased due to the large number of control cases. It is also possible, however, that this finding indicates the difficulty which may exist in the definition and diagnosis of AVN on the basis of radiographs. This may indicate the need for standardization of the definition, and method of diagnosis, for future studies. In addition, only the agreement between experienced observers was investigated in this study. It is possible that less experienced clinicians, such as residents or fellows, may not demonstrate as high a level of both inter-observer and intra-observer agreement as was encountered in this study. A final limitation of the results of this study relates to the patient population. Due to the fact that only cases of SCFE were included, these results may not apply to other patient populations. Despite only four cases of AVN having been diagnosed clinically, the observers indicated between five and eight cases. This discrepancy, though interesting, does not impact on the inter- and intra-observer agreement, since agreement studies do not necessarily make comparisons to a gold

standard diagnostic test [4]. This finding does, however, indicate the difficulty in the diagnosis of AVN in the clinical setting.

Conclusion

The results of this study provide evidence that there is a high level of both inter-observer and intra-observer agreement amongst experienced pediatric orthopaedic surgeons in the radiographic diagnosis of avascular necrosis (AVN) following slipped capital femoral epiphysis (SCFE). The discrepancy in the reported risk of AVN following unstable SCFE is not likely related to disagreement; however, it may have been related to different time points in the course of AVN at which the radiographs were obtained, measurement error, or differences in the baseline population.

Acknowledgments No financial support was received for this project.

References

1. Loder RT, Richards BS, Shapiro PS, Reznick LR, Aronson DD (1993) Acute slipped capital femoral epiphysis: the importance of physeal stability. *J Bone Joint Surg Am* 75(8):1134–1140
2. Rattey T, Piehl F, Wright JG (1996) Acute slipped capital femoral epiphysis. Review of outcomes and rates of avascular necrosis. *J Bone Joint Surg Am* 78(3):398–402
3. Kennedy JG, Hresko MT, Kasser JR, Shrock KB, Zurakowski D, Waters PM et al (2001) Osteonecrosis of the femoral head associated with slipped capital femoral epiphysis. *J Pediatr Orthop* 21(2):189–193
4. Landis JR, Koch GG (1977) An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 33(2):363–374
5. Ficat RP (1985) Idiopathic bone necrosis of the femoral head. Early diagnosis and treatment. *J Bone Joint Surg Br* 67(1):3–9
6. Schmitt-Sody M, Kirchhoff 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(3):283–287