

Good inter- and intraobserver reliability for assessment of the slip angle in 77 hip radiographs of children with a slipped capital femoral epiphysis

Bengt HERNGREN^{1,2}, Mikael LINDELL², and Gunnar HÄGGLUND¹

¹ Department of Orthopedics, Lund University, Lund, Sweden; ² Department of Orthopedics, Ryhov County Hospital, Jönköping, Sweden

Correspondence: bengt.herngren@med.lu.se

Submitted 2017-07-30. Accepted 2017-10-30.

Background and purpose — The decision on and the outcome of treatment for a slipped capital femoral epiphysis (SCFE) depend on the severity of the slip. In 2015, web-based registration was introduced into the Swedish Pediatric Orthopedic Quality (SPOQ) register. To determine whether the inclusion of commonly used methods in Sweden for radiographic measurement of SCFE (the calcar femorale [CF] method and the Billing method) is justified, we measured the inter- and intraobserver reliability of these 2 measurements. We also evaluated the internationally more commonly used head-shaft angle (HSA) method.

Material and methods — 4 observers with different levels of experience with radiographic measurements analyzed 77 routine preoperative hip radiographs of children with SCFE. Inter- and intraobserver reliability was evaluated.

Results — The interobserver reliability analysis for the 4 observers showed for CF an ICC of 0.99 (CI 0.97–0.99) and for Billing an ICC of 0.99 (CI 0.98–0.99). The interobserver reliability analysis for 2 observers showed for HSA an ICC of 0.98 (CI 0.97–0.99).

Intraobserver reliability (2 observers) showed a mean difference below 1° for all 3 methods and with a 95% limit of agreement not exceeding ±6.8°.

Interpretation — We found good reliability for both intra- and interobserver measurements of all 3 methods used for the assessment of the slip angle on routine preoperative lateral hip radiographs.

ann et al. 2006). SCFE is caused by the displacement between the epiphysis and the metaphysis of the proximal femur. The epiphysis remains in the acetabulum while the femur usually rotates outward and in extension (Jerre 1995, Loder 2001). The recommended method of treatment (Souder et al. 2014, Loder 2017) and the outcome (Kocher et al. 2004, Larson et al. 2010, Terjesen and Wensaas 2017) depend on the severity of the slip. The reliability of the methods used to measure the slip angle is therefore important.

Begun in 2015, the Swedish Pediatric Orthopedic Quality register (SPOQ, www.spoq.se) is now a web-based registration tool for 5 pediatric orthopedic conditions. For SCFE, the surgeon is requested to register certain variables including the preoperative slip angle. 2 methods for measurement of the slip angle in SCFE are used in the SPOQ register: the calcar femorale (CF) method (Hansson et al. 1988) and the Billing method (Billing et al. 2002). The aim of this study was to determine whether the reliability of the CF and Billing methods justifies their use in the register. We hypothesized that the intraobserver and interobserver reliability for these methods would warrant their use in SPOQ. We also included a comparison between the CF and the internationally more commonly used head-shaft angle (HSA) method (Southwick 1967).

Material and methods

Conventional radiographs from 94 consecutively registered children with SCFE included in the SPOQ register during 2013 and 2014 were assessed. The radiographs used were routine preoperative examinations from all Swedish hospitals

Slipped capital femoral epiphysis (SCFE) is the most common hip disorder in children aged 9–15 years (Loder 1996, Lehm-

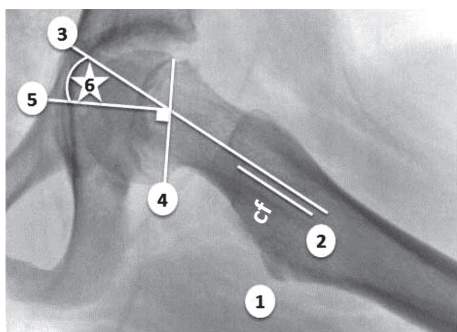


Figure 1. Slip angle measured using the calcar femorale method in a Lauenstein view.

1. Identify the calcar femorale (cf) and the lesser trochanter ①.
2. From the level of the lesser trochanter draw a line ② three cm in a proximal direction parallel to the calcar femorale.
3. Extend a line ③ parallel to line ② up through the femoral neck.
4. Define a line ④ through the physal anterior and posterior margins.
5. Draw a line ⑤ perpendicular to line ④.
6. Slip angle ⑥.

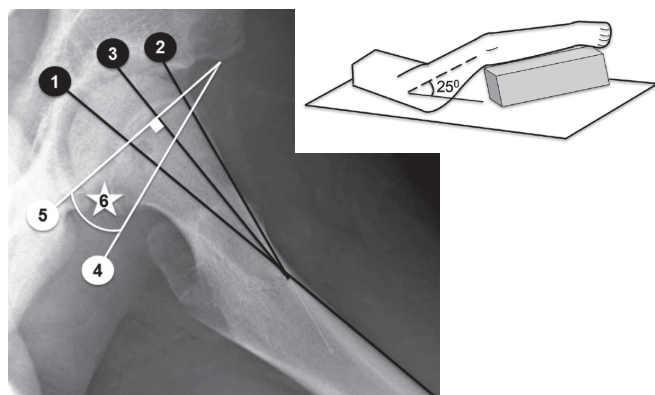


Figure 2. Slip angle measured using the Billing method in Billing lateral view with the patient positioned according to the figure to the right.

1. Draw a line ① along the anterior cortex of the proximal femur. Extend the line up through the femoral head and neck.
2. Draw a line ② along the anterior border of the femoral neck.
3. Draw the bisector ③ to lines ① and ②.
4. Define a line ④ through the physal anterior and posterior margins.
5. Draw a line ⑤ perpendicular to line ③.
6. The slip angle ⑥ is the angle between lines ④ and ⑤.

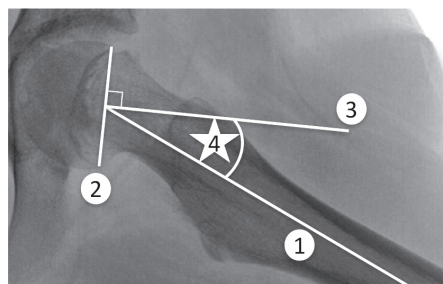


Figure 3. Slip angle measured using the lateral head–shaft angle method in a Lauenstein view.

1. Draw a line ①, parallel with the proximal femoral shaft, further up into the femoral neck.
2. Define a line ② through the physal anterior and posterior margins.
3. Draw a line ③ perpendicular to line ②.
4. Lateral head–shaft angle ④.

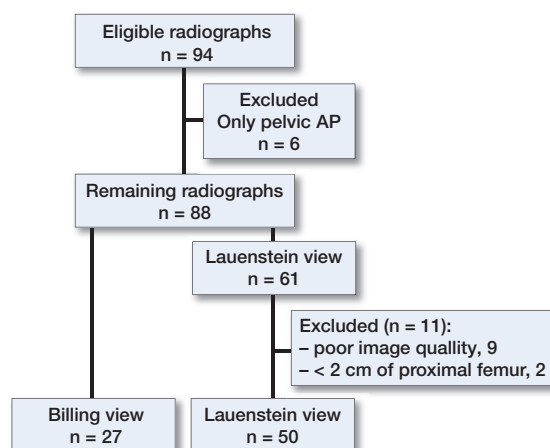


Figure 4. Study material.

that treated children with SCFE during this period. 1 fourth-year resident in orthopedics (ML – Observer 1), 1 specialist in orthopedics (BH – Observer 2), and 2 specialists in pediatric radiology (Observers 3 and 4) acted as observers.

To obtain both presumptive normal hips and hips with SCFE in the study material, we chose the right hip for assessment for every second radiograph in the consecutive list irrespective of whether it was a hip with SCFE or a normal hip. The web-based instructions, available through the SPOQ, were used (Figures 1 and 2) together with a similar instruction on how to measure the lateral HSA (Figure 3).

For the Lauenstein view (both hips), horizontal rotational alignment with an obturator index between 0.7 and 1.8 (Tönnis 1976) and at least 2 cm of the proximal femur below the lesser trochanter had to be included (Lehmann et al. 2013). For the Billing lateral view, radiographs were accepted if the lesser trochanter was not protruding posteriorly or anteriorly. A correct rotational alignment was emphasized by the developer of this method to be a crucial factor (Billing et al. 2002). According to these criteria, 50 radiographs in the Lauenstein view and 27 in the Billing lateral view were included in the analysis (Figure 4).

Interobserver reliability

The orthopedic resident and the orthopedic specialist (Observers 1 and 2), respectively, used their standard picture archiving and communication system (PACS). The orthopedic specialist was experienced in the use of both the CF and Billing methods. The orthopedic resident had no previous experience in any method. The pediatric radiologists (Observers 3 and 4) assessed all radiographs using their standard PACS. They were both experienced in using the Billing method but not the CF or HSA methods.

The radiographs were all given a unique number in a list that did not follow any alphabetical order or pattern according to age, date, or sex. The observers were blinded to the measurements made by the other observers, radiographic reports,

information from medical records, or their own previous measurements. They were allowed to use their own preferred screen settings. Standardization of the measurements was performed before the study by thorough discussion and interpretation of the instructions. The orthopedic resident measured 20 different pelvic radiographs under supervision of Observer 2.

We used these instructions for a single measurement for all 4 readers for the Billing and CF methods.

Intraobserver reliability

Observers 1 and 2 also measured the lateral HSA and assessed the radiographs twice each following the Billing method, the CF method, and the HSA. An interval of at least 6 weeks was used between the repeated measurements for analysis of intraobserver reliability.

Statistics

We assumed a t-distribution for a sample size of < 50 and a normal distribution for a sample size of ≥ 50 . The effect size was set to 3° with 90% power and with a confidence level of 99%. The expected standard deviation was derived from a similar study (Carney and Liljenquist 2005).

Intraobserver variation for each of the measurements was assessed using the mean difference, with its 95% limits of agreement (Bland and Altman 1986, Lehmann et al. 2013). For the purpose of graphic presentation, we plotted the differences against the mean measurements (Bland–Altman plots).

Interobserver variation for 2 observers measuring HSA was assessed using the intraclass correlation coefficient (ICC) and 95% confidence interval (CI) with 2-way random and absolute agreement for single measures. The first measurements were used for both observers (McGraw and Wong 1996). For the 4 observers measuring Billing and CF, interobserver reliability was evaluated using the intraclass correlation coefficient (ICC) and CI with 2-way random and absolute agreement for average measures. The first measurements were used for all observers (McGraw and Wong 1996, Hermanson et al. 2017).

When comparing the HSA with the CF method, we used the first measurements for both methods. For statistical analysis, the variability was described using the Bland–Altman method, with its 95% limits of agreement (Sedgwick 2013).

IBM SPSS Statistics for Windows version 24 (IBM Corp, Armonk, NY, USA) was used for the statistical analyses.

Ethics, funding, and potential conflicts of interest

Ethical approval was authorized by the Regional Ethical Review Board in Lund, Sweden (registration number 2013/87). Informed consent was obtained from all participants and from one parent or guardian.

Funding was received from the Swedish Association of Local Authorities and Regions (SKL), and the Futurum Academy for Health and Care, Jönköping County Council, Jönköping.

The authors declare no conflicts of interest.

Table 1. Intraobserver variation between first and second measurements of slip angle ($^\circ$)

Intraobserver	Subjects	Difference mean (SD)	95% limits of agreement
Observer 1 – Billing	27	–0.8 (2.9)	–6.5 to 5.0
Observer 2 – Billing	27	–0.2 (1.9)	–4.0 to 3.5
Observer 1 – CF	50	0.0 (2.6)	–5.1 to 5.1
Observer 2 – CF	50	0.2 (1.4)	–2.3 to 3.0
Observer 1 – HSA	50	0.1 (1.1)	–2.1 to 2.3
Observer 2 – HSA	50	0.4 (3.3)	–6.0 to 6.8

Table 2. Difference ($^\circ$) between HSA and CF measurement

HSA – CF	Subjects	Difference mean (SD)	95% limits of agreement
Observer 1	50	5.9 (4.8)	–3.5 to 15.3
Observer 2	50	3.4 (4.8)	–6.1 to 12.9

Results

The mean slip angles for the different methods used were: 23° (3° to 59°) for the Billing method, 23° (-8° to 81°) for the CF method, and 26° (-7° to 89°) for the HSA method.

The interobserver reliability analysis for 4 observers showed for CF an ICC of 0.99 (CI 0.97–0.99) and for Billing an ICC of 0.99 (CI 0.98–0.99). The interobserver reliability analysis for 2 observers showed for HSA an ICC of 0.98 (CI 0.97–0.99).

Intraobserver reliability analysis for 2 observers showed a mean difference between the first and second measurement of less than one degree for all three methods. The 95% limits of agreement ranged between -6.5° and 6.8° (Table 1). Bland–Altman plots for HSA and CF visualize the proximity achieved between the first and second measurements (Figures 5 and 6).

The mean difference between the first measurements of HSA and CF was below 6° for 2 observers (Table 2).

Discussion

We found good inter- and intraobserver reliability for all 3 methods for assessing the slip angle on routine preoperative hip radiographs.

The HSA method showed an acceptable inter- and intraobserver reliability. The HSA method produced on average a higher value for the slip angle compared with the CF method. The 95% limit of agreement between the 2 methods also showed a rather wide range of 19° for both observers. On the other hand, the HSA method for observer 1 together with the CF method for observer 2 showed the highest intraobserver

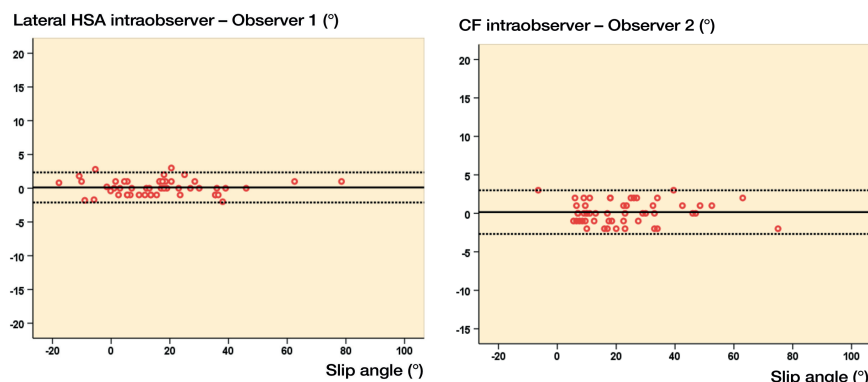


Figure 5. Intraobserver variation (°) for HSA – Observer 1 (left panel). The solid line represents the mean value and the dotted lines show the limits for 2 standard deviations above and below the mean value.

Figure 6. Intraobserver variation (°) for CF – Observer 2 (right panel). The solid line represents the mean value and the dotted lines show the limits for 2 standard deviations above and below the mean value.

reliability. These aspects should all be considered when comparing reports using either of these methods.

In Sweden, the Billing method (Billing et al. 2002) and the CF method (Hansson et al. 1988) are frequently used even though the accuracy of the Billing method for the measurement of a severe slip has been questioned (Loder 2001). An advantage of the CF method is that the CF remains in an unchanged position after remodeling and is identifiable even in adulthood (Harty 1957, Griffin 1982, Hansson et al. 1988); this provides a method for detecting SCFE after growth plate closure (Hansson et al. 1988).

Variability in the radiographic technique can affect the measurement of the slip angle on the Lauenstein view (Jerre 1950, Loder 2001, Carney and Liljenquist 2005). Multiplanar computerized tomography is probably the most reproducible method to assess the slip angle in SCFE (Cohen et al. 1986, Gelberman et al. 1986, Guzzanti and Falciglia 1991, Monazzam et al. 2013) but this technique is not currently an established routine examination in Swedish hospitals for children suspected to have SCFE.

Loder et al. (1999) used Lauenstein radiographs of 48 hips with SCFE (38 children), and 4 observers measured the lateral HSA. They reported no influence of observer experience, no statistically significant difference between the observers and an interobserver variability of ± 12 degrees. Carney and Liljenquist (2005) used 3 observers to test the variability of the lateral HSA using Lauenstein radiographs of 108 hips (55 with SCFE and 53 normal). They reported an intraobserver variability for the HSA of ± 5.9 degrees and concluded that a single observer should document at least a 12-degree change between 2 radiographs to ensure a true change. We found an inter- and intraobserver variability that was comparable with these previous results.

In our study, 11 Lauenstein radiographs did not meet the technical image criteria. Other investigators have also described an inability to obtain reproducible radiographs because of variability in limb position caused by osseous deformities through the physis and/or children experiencing pain (Cohen et al. 1986). Jones et al. (2017) showed by comparing Lauenstein views with 361 simulated models from CT scans that a small

error in positioning could cause a greater than 10° error in the reported lateral HSA.

Clear instructions for the measurement procedure can probably compensate for differences in professional experience provided that the radiographic technique is of acceptable quality. As a consequence of our study, the 4 observers together prepared an updated instruction for all Swedish hospitals on how to achieve a correct Lauenstein view: the hips should be in maximal abduction, the knees flexed to 90° , the plantar aspects of the feet placed together with the lateral aspects of the feet resting against the table, absence of significant asymmetry in the appearance of the obturator foramina (Tönnis 1976), the central beam through the most cranial part of the pubic symphysis, and with a minimum of 5 cm of the femur below the lesser trochanter included in the radiograph.

Our findings indicate that, independent of the experience of the observer, the inter- and intraobserver reliability values for the methods in this study are acceptable for routine use in a national quality register for SCFE. We will consider the inclusion of the HSA as an alternative measurement method for the SPOQ register.

Limitations

The severity of the slips in our study was less than that previously reported in similar studies (Loder et al. 1999, Carney and Liljenquist 2005, Lehmann et al. 2013) and this may have influenced our results for both the intra- and interobserver variability.

We could not blind the radiographs to personal identity numbers because of the need for secure storage of patient information. To compensate, the radiographs were all given a unique number in a list that did not follow any alphabetical order or pattern according to age, date, or sex. We also used a minimum of 6 weeks between the radiographic assessments. Fewer Billing lateral views than Lauenstein views (27 and 50, respectively) were included in this study.

Acknowledgements: Bo Rolander, statistician at the Futurum Academy for Health and Care, Jönköping County Council, Sweden. Håkan Bostrom and Hanna Hebelka Bolminger, specialists in pediatric radiology at Drottning Silvias Barn- och Ungdomssjukhus, Göteborg, Sweden.

Study design: BH and ML; data collection: BH and ML; data analysis: BH, ML, and GH; manuscript preparation: BH, ML, and GH.

Acta thanks Anders Wensaas and other anonymous reviewers for help with peer review of this study.

- Billing L, Bogren H G, Wallin J. Reliable X-ray diagnosis of slipped capital femoral epiphysis by combining the conventional and a new simplified geometrical method. *Pediatric Radiol* 2002; 32 (6): 423-30.
- Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1 (8476): 307-10.
- Carney B T, Liljenquist J. Measurement variability of the lateral head–shaft angle in slipped capital femoral epiphysis. *J Surg Orthop Adv* 2005; 14 (4): 165-7.
- Cohen M S, Gelberman R H, Griffin P P, Kasser J R, Emans J B, Millis M B. Slipped capital femoral epiphysis: Assessment of epiphyseal displacement and angulation. *J Pediatr Orthop* 1986; 6 (3): 259-64.
- Gelberman R H, Cohen M S, Shaw B A, Kasser J R, Griffin P P, Wilkinson R H. The association of femoral retroversion with slipped capital femoral epiphysis. *J Bone Joint Surg Am* 1986; 68 (7): 1000-7.
- Griffin J B. The calcar femorale redefined. *Clin Orthop Relat Res* 1982; (164): 211-14.
- Guzzanti V, Falciglia F. Slipped capital femoral epiphysis: Comparison of a roentgenographic method and computed tomography in determining slip severity. *J Pediatr Orthop* 1991; 11 (1): 6-12.
- Hansson L I, Häggglund G, Ordeberg G, Sandstrom S. The calcar femorale as a landmark in hip physiolysis. *Acta Orthop Scand* 1988; 59 (2): 134-8.
- Harty M. The calcar femorale and the femoral neck. *J Bone Joint Surg Am* 1957; 39 (3): 625-30.
- Hermanson M, Häggglund G, Riad J, Rodby-Bousquet E. Inter- and intra-rater reliability of the head–shaft angle in children with cerebral palsy. *J Child Orthop* 2017; 11 (4): 256-62.
- Jerre T. A study in slipped upper femoral epiphysis: With special reference to the late functional and roentgenological results and to the value of closed reduction [thesis]. *Acta Orthop Scand* 1950; Suppl 6.
- Jerre R. Physiolysis of the hip: Epidemiology, diagnosis and long-term follow-up [thesis]. Institute of Surgical Sciences, Gothenburg 1995.
- Jones C E, Cooper A P, Doucette J, Buchan L L, Wilson D R, Mulpuri K, d'Entremont A G. Southwick angle measurements and SCFE slip severity classifications are affected by frog-lateral positioning. *Skeletal Radiol* 2017 Aug 24 [Epub ahead of print].
- Kocher M S, Bishop J A, Weed B, Hresko M T, Millis M B, Kim Y J, Kasser J R. Delay in diagnosis of slipped capital femoral epiphysis. *Pediatrics* 2004; 113 (4): e322-5.
- Larson A N, Yu E M, Melton L J, Peterson H A, Stans A A. Incidence of slipped capital femoral epiphysis: A population-based study. *J Pediatr Orthop B* 2010; 19 (1): 9-12.
- Lehmann C L, Arons R R, Loder R T, Vitale M G. The epidemiology of slipped capital femoral epiphysis: an update. *J Pediatr Orthop* 2006; 26 (3): 286-90.
- Lehmann T G, Vetti N, Laborie L B, Engesaeter I O, Engesaeter L B, Rosendahl K. Intra- and inter-observer repeatability of radiographic measurements for previously slipped capital femoral epiphysis at skeletal maturity. *Acta Radiol* 2013; 54: 587-91.
- Loder R T. The demographics of slipped capital femoral epiphysis: An international multicenter study. *Clin Orthop Relat Res* 1996; (322): 8-27.
- Loder R T. Effect of femur position on the angular measurement of slipped capital femoral epiphysis. *J Pediatr Orthop* 2001; 21 (4): 488-94.
- Loder R T. Slipped capital femoral epiphysis: A spectrum of surgical care and changes over time. *J Child Orthop* 2017; 11 (2): 154-9.
- Loder R T, Blakemore L C, Fairley F A, Laidlaw A T. Measurement variability of slipped capital femoral epiphysis. *Journal of Orthopaedic Surgery* 1999; 7 (1): 33-42.
- McGraw K O, Wong S P. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996; 1 (1): 30-46.
- Monazzam S, Dwek J R, Hosalkar H S. Multiplanar CT assessment of femoral head displacement in slipped capital femoral epiphysis. *Pediatric Radiol* 2013; 43 (12): 1599-605.
- Sedgwick P. Limits of agreement (Bland–Altman method). *BMJ* 2013; 346: f1630.
- Souder C D, Bomar J D, Wenger D R. The role of capital realignment versus in situ stabilization for the treatment of slipped capital femoral epiphysis. *J Pediatr Orthop* 2014; 34 (8): 791-8.
- Southwick W O. Osteotomy through the lesser trochanter for slipped capital femoral epiphysis. *J Bone Joint Surg Am* 1967; 49 (5): 807-35.
- Terjesen T, Wensaas A. Prognostic factors for long-term outcome of chronic slipped capital femoral epiphysis treated with fixation in situ. *J Child Orthop* 2017; 11 (2): 114-9.
- Tönnis D. Normal values of the hip joint for the evaluation of X-rays in children and adults. *Clin Orthop Relat Res* 1976; (119): 39-47.