Legg-Calve-Perthes disease. Part I: Classification of radiographs with use of the modified lateral pillar and Stulberg classifications.

Herring JA, Kim HT, Browne R.

Texas Scottish Rite Hospital for Children, 2222 Welborn Street, Dallas, TX 75129, USA. tony.herring@tsrh.org

BACKGROUND: Accurate and reliable radiographic classifications of the relative severity and outcome of Legg-Calve-Perthes disease are essential in the study of that disease. As part of a prospective multicenter study, we sought to define more clearly the lateral pillar classification of severity and the Stulberg classification of outcome; we sought especially to define the borderlines between classification groups. METHODS: We performed interobserver and intraobserver trials of the lateral pillar and Stulberg classifications using sets of twenty radiographs chosen from a prospective study of 345 hips. To establish reliable definitions of the lateral pillar classification, we added a new, intermediate group termed the B/C border group, which includes femoral heads with a thin or poorly ossified lateral pillar and those with a loss of exactly 50% of the original height of the lateral pillar. The resulting classification consists of four groups: A, B, B/C border, and C. In our application of the classification system of Stulberg et al., we defined a class-II femoral head as round and fitting within 2 mm of a circle on both anteroposterior and frog-leg lateral radiographs. We defined a Stulberg class-III femoral head as out of round by more than 2 mm on either view and a Stulberg class-IV femoral head as one with at least 1 cm of flattening of the weight-bearing articular surface. To assess interobserver and intraobserver agreement, we performed two trials of each classification with six orthopaedic surgeons reviewing twenty radiographs or pairs of radiographs. RESULTS: In the first trial of the lateral pillar classification, there was 81% agreement per radiograph and the average weighted kappa was 0.71. In the second trial, there was 85% agreement per radiograph and the weighted kappa averaged 0.79. Intraobserver reliability testing showed a 77% match between Trials 1 and 2, an average weighted kappa of 0.81, and an average generalizability coefficient of 0.91. In Trial 1 of the Stulberg classification, there was 91% agreement per radiograph and an average weighted kappa of 0.82. In Trial 2, there was 92% agreement per radiograph and an average weighted kappa of 0.82. Intraobserver reliability testing showed an 89% match between Trials 1 and 2, an average weighted kappa value of 0.88, and an average generalizability coefficient of 0.92. CONCLUSIONS: The interobserver and intraobserver trials of these classifications produced kappa values and generalizability coefficients in the excellent range. The modified lateral pillar classification and the redefined
Stulberg classification are sufficiently reliable and accurate for use in studies of Legg-Calve-Perthes disease.