A Prospective Multicenter Study of Legg-Calvé-Perthes Disease

Functional and Radiographic Outcomes of Nonoperative Treatment at a Mean Follow-up of Twenty Years

A. Noelle Larson, MD, Daniel J. Sucato, MD, John Anthony Herring, MD, Stephen E. Adolfsen, MD, Derek M. Kelly, MD, Jeffrey E. Martus, MD, John F. Lovejoy, MD, Richard Browne, PhD, and Adriana DeLaRocha, MS

Investigation performed at the Texas Scottish Rite Hospital for Children, Dallas, Texas

Background: Long-term studies have indicated good outcomes for most patients with Legg-Calvé-Perthes disease. However, clinical experience suggests that less favorable outcomes are common. We sought to prospectively document pain and function in a cohort of adults who had previously been treated nonoperatively for Legg-Calvé-Perthes disease.

Methods: Patients in our region with Legg-Calvé-Perthes disease were enrolled between 1984 and 1991 as part of a multicenter prospective trial and were treated with hip range-of-motion exercises or bracing. Patients returned for physical examination, radiographs, and completion of outcome measures including the Nonarthritic Hip Score (NAHS) and the Iowa Hip Score (IHS).

Results: Fifty-six patients (fifty-eight hips) were examined at a mean of 20.4 years (range, 16.3 to 24.5 years) after enrollment. The mean NAHS was 79 (range, 35 to 100), and the mean IHS was 74 (range, 43 to 100). Three patients had required hip arthroplasty and one patient had required a pelvic osteotomy. Fourteen (26%) of the remaining hips had no hip osteoarthritis, sixteen (30%) had mild osteoarthritis (Tönnis grade 1), and twenty-four (44%) had moderate or severe osteoarthritic changes on radiographs (grade 2 or 3). Femoroacetabular impingement indicated by physical examination was associated with pain and with poorer outcomes on the IHS and the NAHS ($p = 0.0004$, 0.0014, and 0.0007, respectively). The Stulberg classification was significantly associated with impingement on physical examination ($p = 0.0495$), the NAHS ($p = 0.003$), and the Tönnis grade ($p = 0.012$). Multivariate logistic regression showed that only the Stulberg classification was significantly associated with the NAHS ($p = 0.0032$); the odds ratio for a Stulberg type of I or II compared with IV or V in patients with a fair or poor NAHS was 0.101 (95% confidence interval, 0.018 to 0.573).

Conclusions: Pain, arthritis, and ongoing hip dysfunction are common in patients with Legg-Calvé-Perthes disease that was treated nonoperatively. Hips rated as Stulberg type III or IV more frequently had poor or fair outcomes on the IHS and NAHS (61% and 72% for type III and 77% and 60% for type IV). Patients with a lateral pillar type of B, B/C, or C frequently had pain and radiographic evidence of osteoarthritis. Clinical signs of femoroacetabular impingement were associated with pain and with lower functional scores. This prospective study establishes a modern benchmark for outcomes following nonoperative, weight-bearing treatment of Legg-Calvé-Perthes disease.

Level of Evidence: Prognostic Level II. See Instructions for Authors for a complete description of levels of evidence.

Disclosure: None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

A commentary by Dennis R. Wenger, MD, is linked to the online version of this article at jbjs.org.

Legg-Calvé-Perthes disease likely results from osteonecrosis of the femoral head epiphysis in a skeletally immature patient. The approach to treatment is controversial. Prior to evaluating surgical intervention, it is necessary to have a clear understanding of the disease prognosis following nonoperative treatment.

Previous studies outlining the long-term outcomes of nonoperative treatment of Legg-Calvé-Perthes disease have been retrospective and have often included limited radiographic information from the time of initial presentation[1-6]. Other prospective studies of both operative and nonoperative treatment have not included follow-up past adolescence or early adulthood[7,8]. Good long-term functional outcomes have been reported despite persistent radiographic abnormalities[1,2,4]. However, clinical experience commonly reveals a subset of patients in their second and third decades of life with substantial dysfunction and pain[9].

The etiology of hip pain in young adults with a history of Legg-Calvé-Perthes disease is unclear. Proposed pain generators in these patients include femoroacetabular impingement, instability, labral disease, and early osteoarthritis. New surgical techniques such as femoral osteochondroplasty and labral repair have recently been introduced to treat the sequelae of Legg-Calvé-Perthes disease[10-12]. It is not known whether radiographic and clinical findings of impingement in patients with Legg-Calvé-Perthes disease are associated with patient-reported symptoms and functional outcome measures.

We sought to carefully describe the clinical and radiographic outcomes of nonoperatively treated Legg-Calvé-Perthes disease in a cohort of patients in their third decade of life. This study was designed to test several hypotheses: (1) many adults with a history of Legg-Calvé-Perthes disease would have hip dysfunction and symptoms consistent with femoroacetabular impingement; (2) the prevalence of osteoarthritis in the affected hips would be high following the nonoperative treatment of Legg-Calvé-Perthes disease; and (3) the lateral pillar classification at the fragmentation stage and the Stulberg classification at skeletal maturity would be associated with impingement, the functional outcome scores, and osteoarthritis at the time of the latest follow-up.

Materials and Methods

Patient Population and Study Design

As part of a multicenter study, 337 patients (345 hips) with Legg-Calvé-Perthes disease were enrolled between 1984 and 1991 and followed prospectively. The methods and results of that study have been reported previously[7]. The treatment method was determined by the attending surgeon's enrollment in a prospective study in which only one of five treatment methods was chosen for all patients whom that investigator would manage throughout the study. The five methods were hip range-of-motion therapy, weight-bearing abduction bracing, no treatment, femoral osteotomy, and innominate osteotomy. For the duration of the study, each surgeon agreed to use only one treatment method for all patients meeting the enrollment criteria. Our facility and a center in a neighboring state were both nonoperative treatment sites for the multicenter study. Thus, this paper focuses exclusively on the long-term outcomes of patients in our region who were treated nonoperatively. Patients were enrolled during the initial phases of the Legg-Calvé-Perthes disease, and treatment began at the time of diagnosis. Patients who were younger than six years of age, who had undergone prior treatment, or who had known risk factors for osteonecrosis were excluded.

As part of the multicenter study, 206 hips underwent nonoperative treatment with standardized hip motion exercises (seventy-seven hips) or weight-bearing abduction bracing (129 hips). The choice between range-of-motion exercises and bracing was based on the individual surgeon, who had agreed to use one treatment method for Legg-Calvé-Perthes disease for the duration of the study.

Seventy-four patients (seventy-eight hips) from the original study were enrolled at our center; fifty-nine of these hips had been treated with bracing and nineteen with range-of-motion exercises. Three of these patients had died and two were incarcerated at the time of the present study. Thus, sixty-nine of the patients initially enrolled at our center were potentially available for follow-up. Forty-one (59%) of these patients returned for follow-up, and the remaining twenty-eight patients were lost to follow-up. An additional fifteen patients from the original study were recruited from an adjacent state for follow-up at our center, and the total nonoperatively treated cohort in the present study was therefore fifty-eight hips in fifty-six patients.

Range-of-motion treatment consisted of active exercises, traction, hip adductor tendon releases, and abduction plaster casting as necessary to maintain ≥30° of hip abduction. Weight-bearing was allowed as tolerated. Hip motion was actively monitored by the treating physician at regular clinic visits. Brace treatment involved prescribed full-time wear of the Atlanta Scottish Rite orthosis to position the legs in ≥30° of hip abduction. If abduction to <30° was noted at a clinic visit, plaster abduction casting was initiated. Six of the thirty-two hips in the bracing group and ten of the twenty-six hips in the range-of-motion treatment group lost range of motion and were treated in a weight-bearing Petri cast for one month. Radiographs were obtained in the braces and Petri casts to verify adequate abduction of the hip and appropriate positioning of the femoral head. The lateral pillar classification was assessed at the fragmentation stage, and the Stulberg classification was assessed at skeletal maturity[2,13].

The patients who had been treated nonoperatively in our region were invited to return for clinical examination, radiographs, and determination of outcome measures including the Iowa Hip Score (IHS), the Short Form-36 (SF-36), and the Nonarthritic Hip Score (NAHS). The latter is a validated instrument for young adults with hip disease and may detect hip dysfunction in younger patients[14].

Clinical examination at the time of the latest follow-up included hip range of motion, femoroacetabular impingement tests, and the presence of the Trendelenburg sign. A positive impingement sign on physical examination has been shown to be specific, but not sensitive, for femoroacetabular impingement[15]. The anterior impingement test was considered positive if the patient had pain with hip flexion to 90° (or a smaller angle as tolerated) in internal rotation and if the pain was relieved by external rotation of the hip. The lateral impingement test was considered positive if the patient had pain with lateral abduction of the hip in a supine position with the pelvis stabilized. The posterior impingement test was considered positive if the patient had pain with extension and external rotation of the hip.

Radiographic parameters evaluated include the Tönnis grade, femoral head-neck offset, articular-trochanteric distance, head size ratio, and sourcil type (normal or dysplastic)[16]. We also attempted to measure joint space narrowing and the alpha angle but found these measurements to have poor reliability because of the lack of joint congruency and the asphericity of the femoral head. The head-neck offset ratio could be measured reliably but was not necessarily reflective of impingement, since the impingement in Legg-Calvé-Perthes disease results from contact of the large femoral head with the acetabulum rather than from the presence of a lesion at the femoral head-neck junction[16]. The head size ratio was calculated as the size of the unaffected femoral head divided by that of the affected femoral head, with a ratio of <0.9 representing coxa magna[1]. The articular-trochanteric distance was defined as the distance between two lines, one at the top of the femoral head and one at the top of the trochanter, that were perpendicular to the Perkins line[8]. This distance was considered negative if the trochanter extended more proximally than the femoral head did.

TABLE I Correlates of Poor Outcome Scores*

	Nonarthritic Hip Score			Iowa Hip Score		
	Poor or Fair†	P Value‡	95% CI of Difference§	Poor or Fair†	P Value‡	95% CI of Difference§
Physical examination						
Impingement		0.0008	24%, 76%		0.0040	18%, 67%
Yes	67% (22/33)			76% (25/33)		
No	19% (4/21)			33% (7/21)		
Trendelenburg sign		0.0165	12%, 64%		0.0070	17%, 63%
Yes	75% (12/16)			88% (14/16)		
No	38% (14/38)			47% (18/38)		
Radiographic measures						
Tönnis grade		0.58	–37%, 15%		1.00	–24%, 28%
0 or 1	43% (13/30)			60% (18/30)		
2 or 3	54% (13/24)			58% (14/24)		
Coxa magna#		0.37	–14%, 44%		0.033	4%, 61%
Yes	49% (17/35)			66% (23/35)		
No	33% (5/15)			33% (5/15)		

*Excluding hips that required reconstructive surgery (pelvic osteotomy or arthroplasty). †A score of <80. ‡Calculated with the Fisher test. §CI = confidence interval. #A head size ratio of <0.9. Four hips in which the ratio could not be measured accurately because of extremely distorted hip anatomy were also excluded.

Informed consent was obtained from all families and patients. Initial enrollment in the study was performed prior to the availability of the institutional review board process. However, institutional review board approval for the return visits and radiographs was obtained in 2007.

Statistical Analysis

Categorical data were analyzed with use of standard chi-square methods and two-way contingency tables. Continuous variables were analyzed with use of the Student t test. A p value of <0.05 was considered significant. No Bonferroni correction was utilized, although this may increase the risk of reporting a false positive result as significant[17]. A lack of significance should be interpreted as meaning that the sample size was too small to detect a possible difference between the groups rather than that there was no difference. Multivariate logistic regression was utilized to test which combination of factors affected the probability of a particular outcome as measured with the IHS and the NAHS. The factors that were considered were chronologic age at presentation, physiologic age at presentation (based on osseous development), bilateral Legg-Calvé-Perthes disease, treatment type, sex, lateral pillar classification, Stulberg classification, and duration of follow-up. A backward elimination process was used to remove any nonsignificant factors. Where appropriate, the difference and its 95% confidence interval were reported.

Source of Funding

No external funding source was used in this study.

Results

Demographics

Fifty-six patients (fifty-eight hips) from the original cohort of patients treated nonoperatively in our region returned for follow-up evaluation and radiographs (see Appendix). No differences in functional outcomes were detected between the patients treated with bracing and those treated with range-of-motion exercises, although the duration of follow-up was significantly longer in the range-of-motion group (see Appendix). Compared with the nonoperatively treated patients from the original cohort who did not return for follow-up, our study cohort had a similar distribution of age at diagnosis, sex, bilateral disease, lateral pillar classification, and Stulberg classification (see Appendix). This indicates that the fifty-six patients who participated in our study were similar to the 145 patients who did not return.

Additional Surgery

Four patients required surgical reconstruction (see Appendix) and were excluded from the analysis of clinical and radiographic outcomes. Kaplan-Meier survivorship analysis indicated that this represented a 7% rate of reconstructive surgery (either total hip arthroplasty or periacetabular osteotomy) at twenty years following nonoperative treatment. One additional patient underwent hip arthroscopy and was included in the outcomes analysis. We did not find a significant association between reconstructive surgery (total hip arthroplasty or periacetabular osteotomy) and lateral pillar classification or age at symptom onset.

Clinical Findings and Outcome Measures

Thirty-one patients had a positive anterior impingement test on physical examination, eighteen had a positive lateral impingement test, and fourteen had a positive posterior impingement test. Sixteen patients had a Trendelenburg sign, and four patients had a leg-length discrepancy of >2 cm.

Forty-one patients (76%) reported at least occasional hip pain, although only four patients had severe or extreme pain. Twenty-one patients (39%) had pain daily or several times weekly. Pain was most commonly reported in the groin (twenty-six patients). Pain generators included prolonged sitting (twenty-three patients), car rides (twenty-one), squatting (nineteen), and standing (seventeen). Five patients took pain medications (typically nonsteroidal anti-inflammatory drugs) daily or weekly, and an additional eighteen took pain medications occasionally.

Femoroacetabular impingement was associated with groin pain (p = 0.028), use of pain medication (p = 0.005), pain with car rides (p = 0.02), pain with sitting (p = 0.047), and pain with squatting (p = 0.018). Any impingement (anterior, lateral, or posterior) was associated with the presence of pain, a lower IHS, and a lower NAHS (p = 0.0004, 0.0014, and 0.0007, respectively). Further results are shown in Table I.

The mean for the NAHS (which has a maximum score of 100) was 79 (range, 35 to 100). Only twenty-seven hips (50%) achieved a good or excellent score (≥80). The mean for the IHS (which has a maximum score of 100) was 74 (range, 43 to 100). Only twenty-two hips (41%) achieved a good or excellent score (≥80). Poor or fair IHS results (a score of <80) were found in 48% of Stulberg type I or II hips, 61% of type III hips, and 77% of type IV hips. Poor or fair NAHS results (a score of <80) were found in 22% of Stulberg type I or II hips, 72% of type III hips, and 69% of type IV hips. The majority of points lost on the IHS and the NAHS were lost because of poor pain scores. The mean SF-36 physical component summary score was 46 (range, 30 to 56) and the mean mental component summary score was 42 (range, 27 to 51); the mean value for these subscores in the adult population is 50. Limited abduction on physical examination was statistically associated with a fair or poor IHS and NAHS (p = 0.0013 and 0.0024, respectively). Limited hip flexion was associated with a fair or poor NAHS (p = 0.0087). A body mass index of >30 kg/m^2 (obesity) was associated with a lower IHS and a lower NAHS (p = 0.02 and 0.0002, respectively).

Radiographic Findings

Excluding the four patients who had required reconstructive surgery, follow-up radiographs were obtained for fifty-four affected hips. The presence of coxa magna could only be assessed in fifty hips, and thirty-five (70%) of these hips had coxa magna. Twenty-seven hips (50%) had a decreased femoral head-neck offset ratio (<0.17), although it is unclear whether such a head-neck offset ratio is indicative of a cam impingement lesion in the setting of coxa magna and coxa breva associated with Legg-Calvé-Perthes disease. Interestingly, thirty-eight (76%) of the fifty unaffected contralateral hips also had a decreased femoral head-neck offset ratio. There was no significant association between the head-neck offset ratio and impingement on physical examination, although the altered anatomy of hips with sequelae of Legg-Calvé-Perthes disease makes radiographic measurements of the head-neck offset

TABLE II Demographics and Clinical Outcomes According to Lateral Pillar Type

	Lateral Pillar Type					
	A (N = 1)	B (N = 39)	B/C (N = 7)	C (N = 11)	P Value (Test*)	Difference
At enrollment						
Chronologic age of onset‡ *(yr)*	7.4 (7.4, 7.4)	8.0 (6.1, 11.6)	7.5 (6.2, 10.3)	7.0 (6.0, 9.2)	0.147 (ANOVA)	
Physiologic age of onset‡ *(yr)*	Unknown	6.3 (3.5, 11.5)	4.3 (3.5, 5.5)	6.4 (3.5, 12.0)	0.323 (ANOVA)	
Stulberg type§					0.0031 (Fisher)	B better than B/C or C
I or II	1 (100%)	21 (54%)	2 (29%)	0 (0%)		
III	0 (0%)	11 (28%)	2 (29%)	6 (55%)		
IV or V	0 (0%)	7 (18%)	3 (42%)	5 (45%)		
At final follow-up						
Reconstructive surgery§					1.00 (Fisher)	
Yes	0 (0%)	4 (10%)	0 (0%)	0 (0%)		
No	1 (100%)	35 (90%)	7 (100%)	11 (100%)		
Nonarthritic Hip Score†‡	100 (100, 100)	80 (35, 100)	85 (61.3, 100)	68 (50, 95)	0.181 (ANOVA)	
Iowa Hip Score†‡	100 (100, 100)	80 (43, 99.7)	73 (51, 91)	72 (53, 95.3)	0.921 (ANOVA)	
Impingement on exam.†§					0.046 (Fisher)	C has higher rate
Yes	0 (0%)	20 (57%)	3 (43%)	10 (91%)		
No	1 (100%)	15 (43%)	4 (57%)	1 (9%)		
Pain in affected hip†§					1.00 (Fisher)	
Yes	0 (0%)	26 (74%)	4 (57%)	11 (100%)		
No	1 (100%)	9 (26%)	3 (43%)	0 (0%)		

*ANOVA = analysis of variance. †Excluding hips that required reconstructive surgery (pelvic osteotomy or arthroplasty). ‡Values for the lateral pillar types are given as the mean (based on the number of patients for age and on the number of hips for all other parameters), with the range in parentheses. §Values for the lateral pillar types are given as the number of hips, with the percentage in parentheses.

TABLE III Demographics and Clinical Outcomes According to Stulberg Classification

	Stulberg Type				
	I or II (N = 24)	III (N = 19)	IV or V (N = 15)	P Value (Test*)	Difference†
At enrollment§					
Chronologic age at onset *(yr)*	7.4 (6.1, 9.9)	7.8 (6.3, 9.8)	8.1 (6.0, 11.6)	0.22 (ANOVA)	
Physiologic age at onset *(yr)*	5.0 (3.5, 6.8)	5.8 (3.5, 7.0)	7.6 (4.0, 12.0)	0.03 (ANOVA)	95% CI: (0.45, 4.89) for (IV + V) – (I + II)
At final follow-up					
Reconstructive surgery#				0.54 (Pearson)	
Yes	1 (4%)	1 (5%)	2 (13%)		
No	23 (96%)	18 (95%)	13 (87%)		
Nonarthritic Hip Score‡§	88.2 (35.0, 100.0)	72.1 (35.0, 100.0)	68.5 (36.3, 95.0)	0.016 (ANOVA)	95% CI: (–34.7, –4.73) for (IV + V) – (I + II), and (2.52, 29.7) for (I + II) – III
Iowa Hip Score‡§	75.5 (53.0, 99.7)	74.1 (43.0, 99.3)	70.5 (51.0, 95.3)	0.609 (ANOVA)	
Any impingement on exam.‡#				0.006 (Pearson)	(IV + V) > (I + II)
Yes	9 (39%)	12 (67%)	12 (92%)		
No	14 (61%)	6 (33%)	1 (8%)		
Pain in affected hip‡#				0.04 (Fisher 2-way)	
Yes	14 (61%)	16 (89%)	11 (85%)		
No	9 (39%)	2 (11%)	2 (15%)		

*ANOVA = analysis of variance. †CI = confidence interval. ‡Excluding hips that required reconstructive surgery (pelvic osteotomy or arthroplasty). §Values for the Stulberg types are given as the mean (based on the number of patients for age and on the number of hips for all other parameters), with the range in parentheses. #Values for the Stulberg types are given as the number of hips, with the percentage in parentheses.

ratio difficult to interpret. A dysplastic sourcil was present in nineteen affected hips (35%) compared with eight unaffected hips (15). The mean articular-trochanteric distance was 8 mm on the affected side (range, −10 to 33 mm) compared with 22 mm on the unaffected side (range, −5 to 36 mm).

Degenerative changes were common. Fourteen hips (26%) had no osteoarthritic changes, sixteen hips (30%) had mild changes (Tönnis grade 1), twelve (22%) had moderate changes (grade 2), and twelve (22%) had severe osteoarthritic changes (grade 3). Two unaffected hips had mild osteoarthritis consistent with Tönnis grade 1 (4%), three (6%) had severe osteoarthritic changes consistent with grade 3, and the remaining unaffected hips had no osteoarthritic changes.

Association of Classifications with Adult Outcomes

We next sought to determine whether the lateral pillar classification measured at the fragmentation stage predicted clinical outcomes in adulthood (Table II). The lateral pillar classification was closely associated with the Stulberg classification (p = 0.0063). Impingement on examination was associated with a high lateral pillar classification (p = 0.041). We did not find other factors at the latest follow-up that were associated with the lateral pillar classification.

We also evaluated whether the Stulberg classification as determined at skeletal maturity was associated with impingement, hip pain, or osteoarthritis in adulthood (Table III). A more severe Stulberg classification was associated with impingement on physical examination (p = 0.0495). The Stulberg classification was associated with the NAHS but not with the IHS. The Stulberg classification was also associated with the Tönnis grade (p = 0.012), with a worse Stulberg classification being associated with greater osteoarthritic changes (Fig. 1).

We further examined specific radiographic findings to see whether the lateral pillar classification or the Stulberg classification was associated with radiographic abnormalities in adulthood (Tables IV and V). The lateral pillar and Stulberg classifications were associated with the articular-trochanteric distance, with more trochanteric overgrowth in severely affected hips (p = 0.0068 and 0.001, respectively). Greater lateral pillar involvement was also associated with a dysplastic sourcil (p = 0.0362).

A multivariate logistic regression analysis was performed to evaluate associations among patient demographics, patient outcomes, and the lateral pillar and Stulberg classifications. Only the Stulberg classification was a significant contributor to the NAHS (p = 0.032). The Stulberg type I and II hips had

TABLE IV Radiographic Outcomes According to Lateral Pillar Type

	Lateral Pillar Type					
	A (N = 1)	B (N = 39)	B/C (N = 7)	C (N = 11)	P Value (Test*)	Difference†
Reconstructive surgery‡					1.000 (Fisher)	
Yes	0 (0%)	5 (13%)	0 (0%)	1 (9%)		
No	1 (100%)	34 (87%)	7 (100%)	10 (91%)		
Dysplastic sourcil‡§					0.0362 (Fisher)	B > B/C, and B > C
No	0 (0%)	27 (77%)	3 (43%)	5 (45%)		
Yes	1 (100%)	8 (23%)	4 (57%)	6 (55%)		
Coxa magna‡§#					0.305 (Fisher)	
Yes	0 (0%)	20 (65%)	6 (86%)	9 (82%)		
No	1 (100%)	11 (35%)	1 (14%)	2 (18%)		
Head-neck offset ratio‡§					1.000 (Fisher)	
<0.17, cam deformity	1 (100%)	18 (51%)	3 (43%)	5 (45%)		
≥0.17	0 (0%)	17 (49%)	4 (57%)	6 (55%)		
Articular-trochanteric distance§** *(mm)*	26.4 (26.4, 26.4)	9.9 (–10.2, 32.8)	8.8 (0, 19.6)	0.1 (–8.6, 14.9)	0.0068 (ANOVA)	95% CI: (1.36, 51.3) for (A – C), and (1.30, 18.4) for (B – C)
Tönnis grade‡§					0.804 (Fisher)	
0 or 1	1 (100%)	19 (54%)	3 (43%)	7 (64%)		
2 or 3	0 (0%)	16 (46%)	4 (57%)	4 (36%)		

*ANOVA = analysis of variance. †CI = confidence interval. ‡Values for the lateral pillar types are given as the number of hips, with the percentage in parentheses. §Excluding hips that required reconstructive surgery (pelvic osteotomy or arthroplasty). #A head size ratio of <0.9. Four hips in which the ratio could not be measured accurately because of extremely distorted hip anatomy were also excluded. **Values for the lateral pillar types are given as the mean (based on the number of hips), with the range in parentheses.

better NAHS results, with an odds ratio of 0.101 for a fair or poor NAHS in type I or II hips compared with type IV or V hips. No variables were predictive of the IHS; the lateral pillar classification was weakly predictive in the initial full model, but was subsumed by the Stulberg classification in the final model. Chronologic age at presentation, physiologic age at presentation, bilateral disease, treatment type (brace or range of motion), sex, and duration of follow-up were not significant

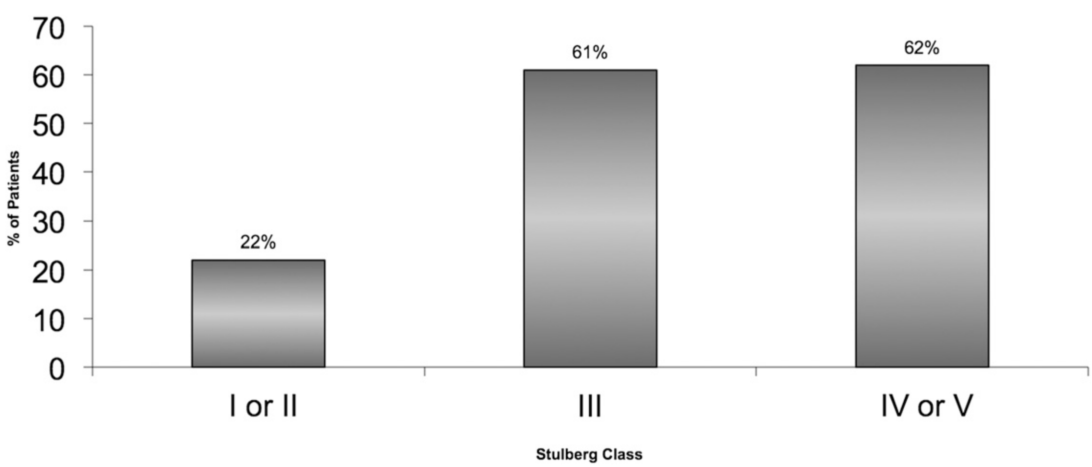

Fig. 1

Percentage of hips with Tönnis grade 2 or 3 changes according to the Stulberg classification. At a mean of twenty years of follow-up, the rate of arthritis in the Stulberg type III hips was similar to that in the type IV and V hips.

TABLE V Radiographic Outcomes According to Stulberg Classification

	Stulberg Type				
	I or II (N = 24)	III (N = 19)	IV or V (N = 15)	P Value (Test*)	Difference†
Reconstructive surgery‡				0.308 (Fisher)	
Yes	1 (4%)	1 (5%)	2 (13%)		
No	23 (96%)	18 (95%)	13 (87%)		
Dysplastic sourcil‡§				0.0502 (Fisher)	Suggests (I + II) > (IV + V)
Yes	4 (17%)	8 (44%)	7 (54%)		
No	19 (83%)	10 (56%)	6 (46%)		
Coxa magna‡§#				0.047 (Fisher)	(I + II) > (IV + V)
No	10 (48%)	4 (24%)	1 (8%)		
Yes	11 (52%)	13 (76%)	11 (92%)		
Unable to calculate	2	1	1		
Head-neck offset ratio‡§				0.780 (Fisher)	
<0.17, cam deformity	13 (57%)	8 (44%)	6 (46%)		
≥0.17	10 (43%)	10 (56%)	7 (54%)		
Articular-trochanteric distance§** *(mm)*	13.4 (–10.2, 32.8)	5.6 (–6.9, 14.3)	2.4 (–8.6, 14.9)	0.001 (ANOVA)	95% CI: (1.0, 14.6) for (I + II) – III, and (3.7, 18.4) for (I + II) – (IV + V)
Tönnis grade‡§				0.012 (Fisher)	(I + II) > III, and (I + II) > (IV + V)
0 or 1	18 (78%)	7 (39%)	5 (38%)		
2 or 3	5 (22%)	11 (61%)	8 (62%)		

*ANOVA = analysis of variance. †CI = confidence interval. The CI is given if p < 0.05. ‡Values for the Stulberg types are given as the number of hips, with the percentage in parentheses. §Excluding hips that required reconstructive surgery (pelvic osteotomy or arthroplasty). #A head size ratio of <0.9. The ratio could not be measured accurately in four hips because of extremely distorted hip anatomy. **Values for the Stulberg types are given as the mean (based on the number of hips), with the range in parentheses.

contributors to the model. The residual chi-square had a p value of 0.3870, indicating a satisfactory fit.

Discussion

This study found that a cohort of patients with Legg-Calvé-Perthes disease evaluated twenty years after nonoperative treatment (involving either range-of-motion exercises or weight-bearing abduction bracing) commonly have hip pain and dysfunction. Four patients had already undergone joint replacement or periacetabular osteotomy, and 76% of the fifty-four hips not requiring further surgery were at least occasionally painful. At least half of the patients had a poor or fair outcome according to the IHS and the NAHS. Hip pain was associated with a positive impingement sign on physical examination. Although it is difficult to differentiate between impingement and other intra-articular sources of hip pain, there were associations among the presence of impingement signs, hip pain, and a poor Stulberg classification.

In comparison to prior reports, this study indicated a high prevalence of degenerative changes in hips with Legg-Calvé-Perthes disease at twenty years of follow-up (see Appendix). Previous retrospective studies have indicated the long-term outcomes following nonoperative treatment of Legg-Calvé-Perthes disease[1-5,18]. Stulberg et al. categorized the radiographic results of nonoperative treatment in 171 hips at skeletal maturity as spherical congruency (types I and II), aspherical congruency (III and IV), and aspherical incongruency (V). Ninety-nine patients had thirty to forty years of follow-up. In these patients, osteoarthritic changes were observed in 16% of Stulberg type I or II hips, 60% of type III hips, and 75% of type IV or V hips; 5% of hips had undergone a fusion or arthroplasty procedure[2]. In contrast, our study indicated similar results but at only twenty years of follow-up; approximately 20% of Stulberg type II hips had osteoarthritis; 60% of Stulberg type III, IV, and V hips had osteoarthritis; and 7% of hips had required reconstructive surgery. Gower and Johnston reviewed thirty-five patients who were initially treated nonoperatively[3]. At a mean follow-up of thirty-six years, five (14%) had undergone reconstructive surgery, with three having required arthroplasty. Eighteen (60%) of the remaining thirty had activity-related pain, and thirteen (43%) had joint space narrowing[3]. McAndrew and Weinstein provided an update of the Gower and Johnston paper with a follow-up of nearly fifty years[1]. By that time, eleven (39%) of twenty-eight patients had required total joint arthroplasty, and six (21%) had an IHS of <80. In our series with twenty years of follow-up, 7% of patients had undergone reconstructive surgery, and approximately 60% had an IHS of <80.

Thus, we found higher rates of reconstructive surgery, patient-reported pain, and arthritis than those in the previous studies. There are several possible explanations for our findings. It may be that weight-bearing nonoperative treatment produces inferior results compared with previous techniques of prolonged bed rest and non-weight-bearing. Alternatively, patients in today's culture may have higher expectations following the treatment of Legg-Calvé-Perthes disease. Stulberg type III, IV, and V hips are known to develop radiographic signs of arthritis by mid- to late adulthood[2]. Through hip arthroscopy and improved magnetic resonance imaging, we now understand that radiographic changes are the end-stage findings of long-standing structural hip abnormalities[12,19]. Furthermore, new techniques are available for treatment of hip pain in the young adult, including periacetabular osteotomy for instability, hip arthroscopy or surgical hip dislocation for labral repair, and head-neck osteochondroplasty for impingement[10,12,20]. Although we focused on impingement rather than instability in this study, one patient did require a periacetabular osteotomy for symptoms of instability. Also, the more severely affected hips commonly had a dysplastic sourcil.

The majority of unaffected contralateral hips had a decreased femoral head-neck offset ratio. It is possible that some of the patients had a subclinical contralateral presentation of Legg-Calvé-Perthes disease. Alternatively, there may be an underlying structural hip abnormality seen bilaterally in children who develop Legg-Calvé-Perthes disease. Other studies have documented structural abnormalities of the unaffected hip and retroversion of the acetabulum in patients with Legg-Calvé-Perthes disease[21-24].

The Stulberg classification at skeletal maturity was associated with impingement on physical examination, pain, and arthritis. A classification of Stulberg type IV or V was associated with lower NAH scores even when age at presentation, time to follow-up, and sex were considered as candidate variables in the full multivariate regression model. The hips requiring reconstructive surgery were all rated as having a lateral pillar type of B and represented a variety of Stulberg types.

There are several limitations to this long-term, prospective multicenter study. First, excluding deceased and incarcerated patients, only 59% of the nonoperatively treated patients from our center and 47% from our region returned for evaluation. At approximately twenty years after their initial treatment, many patients had been lost to follow-up. Second, it should be emphasized that this is not a population-based study, and it is unclear whether the patients enrolled in the original study are a representative cross-section of all patients with Legg-Calvé-Perthes disease. The distribution of lateral pillar and Stulberg types in the study cohort was similar to that in the cohort of nonoperatively treated patients who did not return for follow-up. In the original study design, there was an exclusion provision for patients who had a seemingly benign clinical course with minimal pain, full range of motion, and limited radiographic involvement; these patients received nonoperative treatment regardless of the treatment protocol at their study center and were not included in the study. Finally, there could have been a selection bias because patients who were more symptomatic may have been more likely to return for follow-up in order to seek help; alternatively, patients who were doing well may have been more likely to return and to report better results than the patients in the cohort who did not return.

Furthermore, all patients initially underwent nonoperative treatment, so these results do not reflect the natural history of untreated Legg-Calvé-Perthes disease. In contrast to historical studies[1,2,4,18], all children in this series were ambulatory during disease treatment. Since initiation of the original multicenter study, the efficacy of the Atlanta Scottish Rite orthosis for the treatment of Legg-Calvé-Perthes disease (particularly its role in achieving hip containment) has been called into question[25-27]. The usage of this brace has been discontinued at our center. It is possible that improved results would be seen in a cohort treated with restricted weight-bearing or with alternative bracing strategies to achieve improved containment.

In summary, our patient population had a high prevalence of osteoarthritis and low clinical outcome scores at only twenty years of follow-up. Approximately 20% of Stulberg type II hips and 60% of type III, IV, and V hips had radiographic evidence of osteoarthritis. This prevalence is worse than in previous long-term reports of the nonoperative treatment of Legg-Calvé-Perthes disease. The behavior of the Stulberg type III hips appeared to be more similar to that of the type IV or V hips in previous reports. A severe Stulberg classification was associated with pain, impingement on physical examination, and arthritis. Although previous studies have shown only mild radiographic changes in young adults with Legg-Calvé-Perthes disease, in reality many patients are quite symptomatic. This report on nonoperative treatment provides useful outcomes data that can serve as a benchmark for future studies evaluating the results of treatment of Legg-Calvé-Perthes disease.

Appendix

eA Tables showing patient demographic data, demographics and outcomes according to treatment type, a comparison of patients who did and did not return for follow-up, data on the patients who required reconstructive surgery, and previously reported long-term results of nonoperative treatment of Legg-Calvé-Perthes disease are available with the online version of this article as a data supplement at jbjs.org. ■

A. Noelle Larson, MD
Department of Orthopedic Surgery, Mayo Clinic,
200 1st Street S.W., Rochester, MN 55905

Daniel J. Sucato, MD
John Anthony Herring, MD
Richard Browne, PhD
Adriana DeLaRocha, MS
Texas Scottish Rite Hospital for Children,
2222 Welborn Street, Dallas, TX 75219.
E-mail address for J.A. Herring: Tony.Herring@tsrh.org

Stephen E. Adolfsen, MD
Pediatric Orthopedic Associates,
P.A. 585 Cranbury Road, East Brunswick, NJ 08816

Derek M. Kelly, MD
Campbell Clinic Orthopaedics, 1400 South Germantown Road,
Germantown, TN 38138

Jeffrey E. Martus, MD
Vanderbilt Children's Hospital, 2200 Childrens Way,
4202 DOT, Nashville, TN 37232-9565

John F. Lovejoy, MD
Children's National Medical Center,
111 Michigan Avenue N.W., Washington, DC 20010

References

1. McAndrew MP, Weinstein SL. A long-term follow-up of Legg-Calvé-Perthes disease. J Bone Joint Surg Am. 1984;66:860-9.

2. Stulberg SD, Cooperman DR, Wallensten R. The natural history of Legg-Calvé-Perthes disease. J Bone Joint Surg Am. 1981;63:1095-108.

3. Gower WE, Johnston RC. Legg-Perthes disease. Long-term follow-up of thirty-six patients. J Bone Joint Surg Am. 1971;53:759-68.

4. Perpich M, McBeath A, Kruse D. Long-term follow-up of Perthes disease treated with spica casts. J Pediatr Orthop. 1983;3:160-5.

5. Kelly FB Jr, Canale ST, Jones RR. Legg-Calvé-Perthes disease. Long-term evaluation of non-containment treatment. J Bone Joint Surg Am. 1980;62:400-7.

6. Mose K, Hjorth L, Ulfeldt M, Christensen ER, Jensen A. Legg Calvé Perthes disease. The late occurence of coxarthrosis. Acta Orthop Scand Suppl. 1977;169: 1-39.

7. Herring JA, Kim HT, Browne R. Legg-Calve-Perthes disease. Part II: prospective multicenter study of the effect of treatment on outcome. J Bone Joint Surg Am. 2004;86:2121-34.

8. Wiig O, Terjesen T, Svenningsen S. Prognostic factors and outcome of treatment in Perthes' disease: a prospective study of 368 patients with five-year follow-up. J Bone Joint Surg Br. 2008;90:1364-71.

9. Catterall A. Adolescent hip pain after Perthes' disease. Clin Orthop Relat Res. 1986;209:65-9.

10. Ganz R, Gill TJ, Gautier E, Ganz K, Krügel N, Berlemann U. Surgical dislocation of the adult hip a technique with full access to the femoral head and acetabulum without the risk of avascular necrosis. J Bone Joint Surg Br. 2001;83:1119-24.

11. Rebello G, Spencer S, Millis MB, Kim YJ. Surgical dislocation in the management of pediatric and adolescent hip deformity. Clin Orthop Relat Res. 2009; 467:724-31.

12. Roy DR. Arthroscopic findings of the hip in new onset hip pain in adolescents with previous Legg-Calve-Perthes disease. J Pediatr Orthop B. 2005;14:151-5.

13. Herring JA, Kim HT, Browne R. Legg-Calve-Perthes disease. Part I: Classification of radiographs with use of the modified lateral pillar and Stulberg classifications. J Bone Joint Surg Am. 2004;86:2103-20.

14. Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. Clin Orthop Relat Res. 2003;406: 75-83.

15. Nogier A, Bonin N, May O, Gedouin JE, Bellaiche L, Boyer T, Lequesne M; French Arthroscopy Society. Descriptive epidemiology of mechanical hip pathology in adults under 50 years of age. Prospective series of 292 cases: clinical and radiological aspects and physiopathological review. Orthop Traumatol Surg Res. 2010;96(8 Suppl):S53-8.

16. Clohisy JC, Carlisle JC, Beaulé PE, Kim YJ, Trousdale RT, Sierra RJ, Leunig M, Schoenecker PL, Millis MB. A systematic approach to the plain radiographic evaluation of the young adult hip. J Bone Joint Surg Am. 2008;90 Suppl. 4:47-66.

17. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1:43-6.

18. Danielsson LG, Hernborg J. Late results of Perthes' disease. Acta Orthop Scand. 1965;36:70-81.

19. Zilkens C, Holstein A, Bittersohl B, Jäger M, Haamberg T, Miese F, Kim YJ, Mamisch TC, Krauspe R. Delayed gadolinium-enhanced magnetic resonance imaging of cartilage in the long-term follow-up after Perthes disease. J Pediatr Orthop. 2010;30:147-53.

20. Ganz R, Klaue K, Vinh TS, Mast JW. A new periacetabular osteotomy for the treatment of hip dysplasias. Technique and preliminary results. Clin Orthop Relat Res. 1988;232:26-36.

21. Eijer H, Berg RP, Haverkamp D, Pécasse GA. Hip deformity in symptomatic adult Perthes' disease. Acta Orthop Belg. 2006;72:683-92.

22. Ezoe M, Naito M, Inoue T. The prevalence of acetabular retroversion among various disorders of the hip. J Bone Joint Surg Am. 2006;88:372-9.

23. Larson AN, Stans AA, Sierra RJ. Ischial spine sign reveals acetabular retroversion in Legg-Calvé-Perthes disease. Clin Orthop Relat Res. 2011;469:2012-8.

24. Sankar WN, Flynn JM. The development of acetabular retroversion in children with Legg-Calvé-Perthes disease. J Pediatr Orthop. 2008;28:440-3.

25. Meehan PL, Angel D, Nelson JM. The Scottish Rite abduction orthosis for the treatment of Legg-Perthes disease. A radiographic analysis. J Bone Joint Surg Am. 1992;74:2-12.

26. Martinez AG, Weinstein SL, Dietz FR. The weight-bearing abduction brace for the treatment of Legg-Perthes disease. J Bone Joint Surg Am. 1992;74:12-21.

27. Wang L, Bowen JR, Puniak MA, Guille JT, Glutting J. An evaluation of various methods of treatment for Legg-Calvé-Perthes disease. Clin Orthop Relat Res. 1995;314:225-33.