



## ■ CHILDREN'S ORTHOPAEDICS

# Radiological hip shape and patient-reported outcome measures in healed Perthes' disease

M. S. Ali,  
M. Khattak,  
D. Metcalfe,  
D. C. Perry

From Alder Hey  
Children's Hospital,  
Liverpool, UK

### Aims

This study aimed to evaluate the relationship between hip shape and mid-term function in Perthes' disease. It also explored whether the modified three-group Stulberg classification can offer similar prognostic information to the five-group system.

### Methods

A total of 136 individuals aged 12 years or older who had Perthes' disease in childhood completed the Patient-Reported Outcomes Measurement Information System (PROMIS) Mobility score (function), Nonarthritic Hip Score (NAHS) (function), EuroQol five-dimension five-level questionnaire (EQ-5D-5L) score (quality of life), and the numeric rating scale for pain (NRS). The Stulberg class of the participants' hip radiographs were evaluated by three fellowship-trained paediatric orthopaedic surgeons. Hip shape and Stulberg class were compared to PROM scores.

### Results

A spherical hip was associated with the highest function and quality of life, and lowest pain. Conversely, aspherical hips exhibited the lowest functional scores and highest pain. The association between worsening Stulberg class (i.e. greater deviation from sphericity) and worse outcome persisted after adjustment for age and sex in relation to PROMIS (predicted mean difference -1.77 (95% confidence interval (CI) -2.70 to -0.83)), NAHS (-5.68 (95% CI -8.45 to -2.90)), and NRS (0.61 (95% CI 0.14 to 1.08)), but not EQ-5D-5L (-0.03 (95% CI -0.72 to 0.11)).

### Conclusion

Patient-reported outcomes identify lower function, quality of life, and higher pain in aspherical hips. The magnitude of symptoms deteriorated with time. Hip sphericity (i.e. the modified three-group classification of spherical, oval, and aspherical) appeared to offer similar levels of detail to the five-group Stulberg classification.

Cite this article: *Bone Joint J* 2023;105-B(6):711–716.

### Introduction

The outcomes of Perthes' disease are usually described in terms of radiological appearance. However, the orthopaedic literature increasingly recommends the use of 'core outcome sets' as a minimum set of outcome domains that should be reported in high-quality studies.<sup>1</sup> The development of core outcomes involves multiple stakeholders, including patients and families, to define which outcomes are important. Radiological appearance, however, is almost invariably ranked as less important than functional outcomes.<sup>1-6</sup> The core outcome set for Perthes' disease consists of 14 measures encompassing various areas which impact the quality of life, as well as the radiological outcome.<sup>7</sup>

Although it is plausible that radiological appearance would be related to patient-reported outcome measures in Perthes' disease, this has not yet been demonstrated. Hailer and Penno<sup>8</sup> investigated 61 patients with 28 years' follow-up and concluded that long-term patient-reported outcome measures (PROMs) exhibited moderate-to-strong correlations with radiological measures of sphericity, femoral head enlargement, and femoral neck growth inhibition. Femoral head sphericity also showed a moderate correlation with the Harris Hip Score,<sup>9</sup> in which hip function was shown to be worse in aspherical hips.

In this study, we have compared the radiological hip shape with mid-term functional outcomes

Correspondence should be sent to D. C. Perry; email: D.C.Perry@liverpool.ac.uk

© 2023 Author(s) et al.  
doi:10.1302/0301-620X.105B6.  
BJJ-2022-1421.R1 \$2.00

*Bone Joint J*  
2023;105-B(6):711–716.

**Table I.** Disease-specific inclusion and exclusion criteria.

| Criteria   |
|--|
| <b>Inclusion criteria</b>  |
| Diagnosis of Perthes' disease made while skeletally immature                   |
| <b>Any of the following radiological features within the femoral epiphysis</b> |
| Flattening   |
| Sclerosis  |
| Fragmentation  |
| Collapse   |
| Reossification   |
| Features may be evident on plain radiographs, or MRI                           |
| Resident within England, Scotland, or Wales                                    |
| Able to understand the study documentation                                     |
| <b>Exclusion criteria (any of the following prior to first diagnosis)</b>      |
| Treatment for developmental hip dysplasia (not including double nappies)       |
| Chemotherapy for malignancy  |
| Sickle cell anaemia  |
| MED or SED   |
| Coagulopathy   |
| Gaucher's disease  |
| Same-sided hip fracture  |
| Hypothyroidism   |
| MED, multiple epiphyseal disease; SED, spondyloepiphyseal dysplasia.           |

in adolescents and adults who have been affected by Perthes' disease in childhood. We hypothesized that the poorest functional outcome scores would be in those with aspherical hips (Stulberg class IV and V),<sup>10,11</sup> while the best outcome scores would be in individuals with spherical hips (Stulberg class I and II). We also hypothesized that older participants in the study would report poorer function than younger ones.

## Methods

**Patient recruitment.** Eligible participants were recruited from a hospital Perthes' disease register. Recruitment was part of the Outcomes Research in Children's Hip Disease (ORCHiD) study. Participants were aged 12 years or older at the time of recruitment, with a prior diagnosis of Perthes' disease that had entered the healed stage. Inclusion and exclusion criteria are outlined in detail in Table I.

Eligible individuals were sent an information pack that included an invitation, information sheet, consent form, and prepaid envelope. Participants and/or their parents were able to contact the research team via email, telephone, or post to discuss participation. Upon receipt of the signed consent forms, questionnaires were sent to the study participants and completed electronically, by paper, or telephone. Patient representatives from the Perthes Association and STEPS Charity contributed to the development and design of the study. Patients were recruited between November 2017 to February 2021.

**Patient-reported outcomes.** The following patient-reported outcomes were collected through the study. Patients were asked to report their symptoms at the time of follow-up.

**PROMIS Mobility v2.0 CAT.** PROMIS Mobility is a set of questions intended to capture physical function related to the lower limbs. The raw scores from PROMIS translate to standardized

**Table II.** Radiological features of participants' hip radiographs.

| Variable              | Participants, n |
|-----------------------|-----------------|
| <b>Stulberg class</b> |                 |
| I                     | 19              |
| II                    | 27              |
| III                   | 37              |
| IV                    | 33              |
| V                     | 20              |
| <b>Hip shape*</b>     |                 |
| Spherical             | 46              |
| Ovoid                 | 37              |
| Aspherical            | 53              |

\*Per the three-group Stulberg classification.

T scores, with normative data suggesting a score of 50 is the population mean with a standard deviation (SD) of 10.<sup>12</sup> The minimal clinically important difference (MCID) of PROMIS Mobility is reported to be 4.2.<sup>13,14</sup>

**Nonarthritic Hip Score.** The Nonarthritic Hip Score (NAHS) is a measure of hip function intended for use in younger patients without arthritic problems or degenerative joint disease,<sup>15</sup> and is widely used in the assessment of Perthes' disease. The score is adjusted and between 0 and 100 where 0 represents a hip without meaningful function, and 100 represents a perfectly functioning hip. The NAHS MCID has been reported to be 8.7.<sup>16</sup>

**EQ-5D-5L.** The EQ-5D-5L is a generic health-related quality of life measure. The questionnaire has five domains relating to activities of daily living with five levels of answer within each domain.<sup>17</sup> Respondents are scored from no problems (score = 1.0) to extreme impairment on all five dimensions (value = -0.594). The MCID for EQ-5D-5L is 0.32.<sup>18</sup>

**Numeric Rating Scale for pain.** This is a unidimensional measure of pain intensity. Among the various versions, the most commonly used is the selection of a whole number between integers of 0 to 10. A score of 0 represents 'no pain' while a score of 10 represents 'worst imaginable pain'. On average, a reduction of one point or reduction of 15% represents a MCID for the NRS.<sup>19</sup>

**Radiological outcome.** The Stulberg classification is the gold-standard measure of hip shape. We used an established classification tree to ensure greater consistency in the descriptions.<sup>10</sup> Radiographs of the hip in the residual/healed stage in late childhood/adolescence were assessed by three fellowship-trained consultant paediatric orthopaedic surgeons (see Acknowledgements). The surgeons independently assessed each image, with the decision of the third surgeon used to resolve discrepancies between the other two by majority vote. We used both the traditional five-group Stulberg classification<sup>10</sup> and the modified three-group classification.<sup>11</sup>

**Age groups of participants.** At the point of PROMs completion, participants were divided into three age categories: 12 to 16 years (adolescents), 17 to 25 years (young adults), and  $\geq$  26 years (adults).<sup>20</sup>

**Statistical analysis.** Kruskal-Wallis one-way analysis of variance was used to compare non-normally distributed continuous variables. As the outcome data were right-skewed, log-linked generalized linear models were used to determine whether

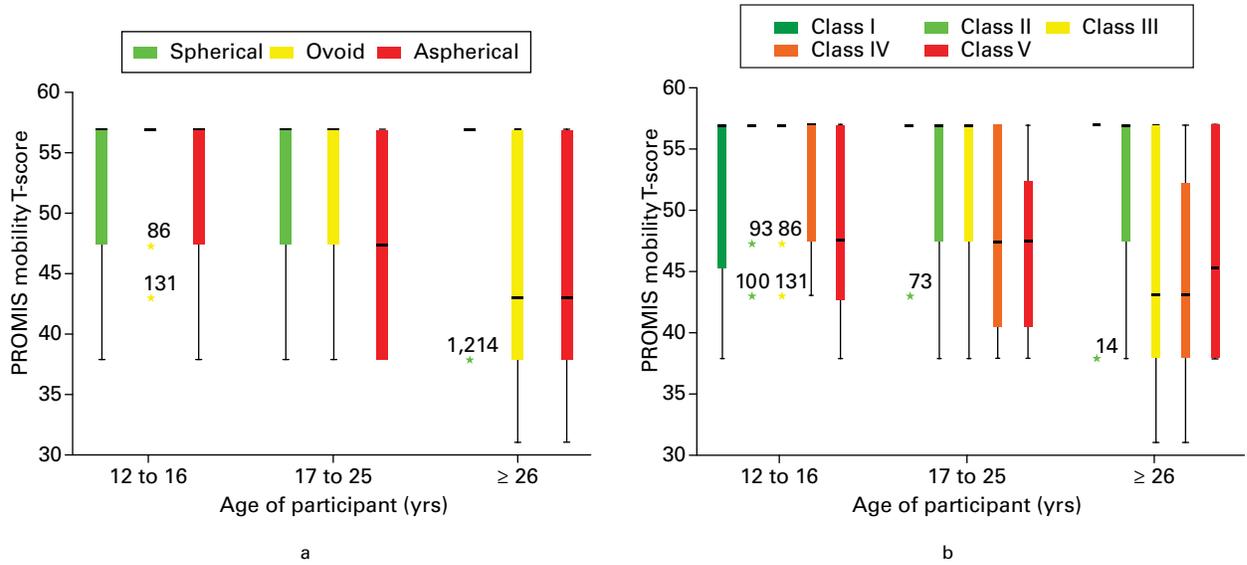


Fig. 1

Boxplot of the effect of a) hip shape and age and b) Stulberg class and age on function as reported by the Patient-Reported Outcomes Measurement Information System (PROMIS) Mobility score. Stars represent 'far out' values according to SPSS (IBM, USA), labelled with their corresponding value from the series of results from the database.

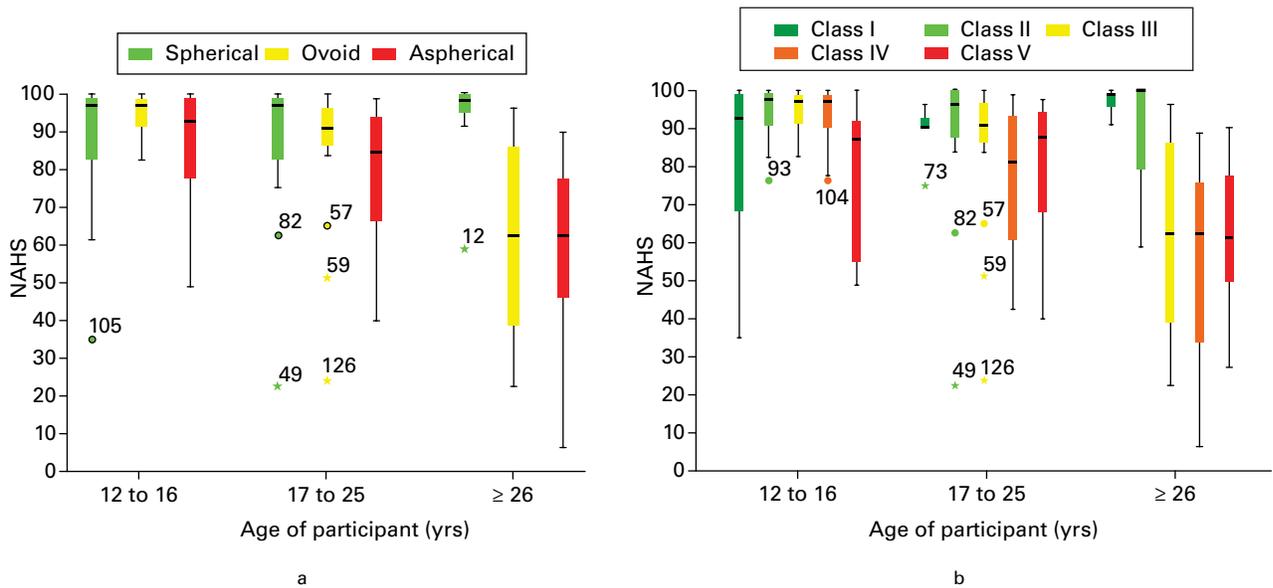


Fig. 2

Boxplot of the effect of a) hip shape and age and b) Stulberg class and age on function as reported by the Nonarthritic Hip Score (NAHS). Circles represent outlier values, and stars represent 'far out' values, according to SPSS (IBM, USA), labelled with their corresponding value from the series of results from the database.

Stulberg class is independently associated with outcome after adjustment for age and sex. Age was included as a continuous variable within these models. Statistical analyses were performed using SPSS v. 26.0 (IBM, USA) and StataIC v. 15 (StataCorp, USA), with  $p < 0.05$  used as the threshold for statistical significance.

**Results**

Of the 856 patients invited to participate, 300 returned questionnaires and of these, 291 returned complete responses. A total of 25 were excluded because they had a history of hip arthroplasty and 130 because hip radiographs were unavailable, which left a study population of 136 participants. The age range was 12 to

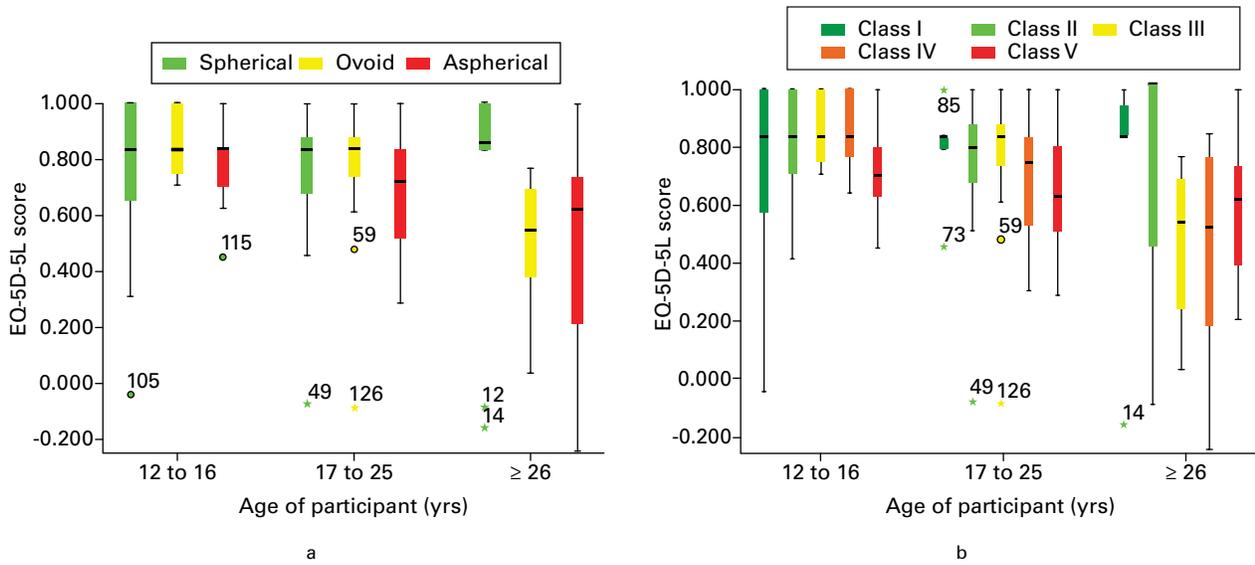


Fig. 3

Boxplot of the effect of a) hip shape and age and b) Stulberg class and age on quality of life as reported by the EuroQol five-dimension five-level questionnaire (EQ-5D-5L) score. Circles represent outlier values, and stars represent 'far out' values, according to SPSS (IBM, USA), labelled with their corresponding value from the series of results from the database.

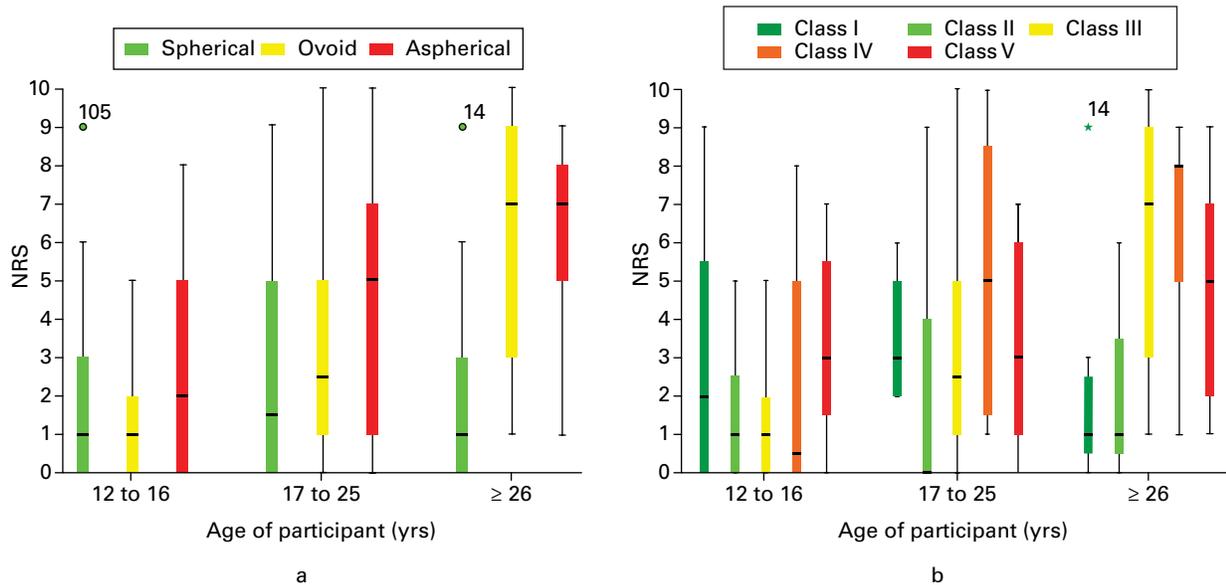


Fig. 4

Boxplot of the effect of a) hip shape and age and b) Stulberg class and age on pain as reported by the numeric rating scale for pain (NRS). Circles represent outlier values, and stars represent 'far out' values, according to SPSS (IBM, USA), labelled with their corresponding value from the series of results from the database.

55 years (mean age at PROM completion 24 years) and 79% (n = 107) were male.

Using the five-group Stulberg classification, most participants had Stulberg group III hips (n = 37) (Table II). Using the three-group classification, most participants had aspherical hips (n = 53), followed by spherical (n = 46), then ovoid (n = 37) (Table II).

Boxplots were plotted to visually represent the distribution of PROM scores for hip shape and age category, and Stulberg class and age category (Figures 1 to 4).

**The relationship between hip shape, age, and sex.** Advancing age was significantly associated with worse outcomes according to PROMIS (p = 0.040) but not NAHS (p = 0.099), EQ-5D-5L (p = 0.134), or NRS (p = 0.984). Male sex was associated with

better outcomes across all four PROMs: PROMIS (median 56.9 (interquartile range (IQR) 47.4 to 56.9) vs 47.4 (37.9 to 56.9);  $p = 0.005$ ), NAHS (92.5 (76.3 to 97.5) vs 75.0 (50.0 to 90.0);  $p = 0.003$ ), EQ-5D (0.837 (0.642 to 1.000) vs 0.678 (0.479 to 0.837);  $p = 0.019$ ), and NRS (2 (0 to 5) vs 5 (1 to 7);  $p = 0.012$ ). In this study, hip shape was not significantly associated with either age ( $p = 0.608$ ) or sex ( $p = 0.063$ ).

**The effect of hip shape on function.** Aspherical hips had significantly worse function compared to spherical hips, as shown by both PROMIS Mobility and NAHS. For both scores, self-reported hip function was worse in participants aged  $\geq 26$  years compared to those aged 12 to 16 years. Within a generalized linear model, each worsening Stulberg class was associated with significant reductions in the PROMIS Mobility T-score (predicted mean difference -1.77 (95% CI -2.70 to -0.83)) and NAHS (-5.68 (95% CI -8.45 to -2.90)).

Participants aged  $\geq 26$  years appeared to show the largest differences in function scores between spherical and aspherical hips, compared to other age categories. The NAHS (Figure 2) appeared to be better at distinguishing the range of hip function captured from participants than the PROMIS Mobility score, as it showed more of a difference than PROMIS between hip shapes and Stulberg classes (Figure 1).

**The effect of hip shape and age on general quality of life.** Within a generalized linear model, worsened Stulberg class was not associated with changes to the EQ-5D-5L (predicted mean difference -0.03 (95% CI 0.72 to 0.11)). Similarly, with regard to function, quality of life was worse in participants aged  $\geq 26$  years compared to those aged 12 to 16 years.

**The effect of hip shape and age on pain.** Within a generalized linear model, worsening Stulberg class was associated with higher pain scores as captured by the NRS (predicted mean difference 0.61 (95% CI 0.14 to 1.08)). Pain in participants aged older than 26 years was reported to be worse than in participants aged 12 to 16 years.

## Discussion

In our study individuals with a spherical femoral head exhibited better function, quality of life, and lower levels of pain than those with aspherical hips. We found that increasing age resulted in a decrease in function and quality of life and increase in pain for all hip shapes with Perthes' disease. Worsening hip shapes was also associated with poorer patient-reported outcomes. Furthermore, we observed that the Stulberg three-group classification appeared to offer similar information to the five-group classification.

Stulberg et al<sup>21</sup> in 1981 were the first to explore the relationship between secondary osteoarthritis and radiological assessment. While the Stulberg classification has largely remained a standard in the assessment of Perthes' disease, few have compared the classification with functional, quality of life, and pain outcomes in a large group of patients. Wiig et al<sup>11</sup> have shown that the Stulberg three-group classification has increased inter-rater reliability, and that the discriminatory ability appeared similar between the three- and the five-group classification in terms of predicting pain, function, and quality of life. The three-group classification is known to be an effective long-term predictor of radiological outcome in Perthes'

disease,<sup>22</sup> and our study suggests that it is effective in predicting patient-reported outcomes. Given the proven usefulness of the three-group classification, with improved interobserver reliability and little detail lost compared to the five-group classification, this appears a more useful tool to describe the outcomes of Perthes' disease. In addition to the three-group classification, other authors have also identified the possibility of additional prognostic information through MRI.<sup>23</sup>

Joint-specific PROMs such as the HHS and the NAHS have exhibited moderate correlations with radiological measures of hip sphericity deviation.<sup>8</sup> Our findings support this as we show a decrease in function in aspherical hips, with the NAHS appearing the most sensitive measure to change in hip shape among the outcomes used within our study. The association between EQ-5D-5L and hip shape was less clear than other functional outcomes, though it seems likely that a quality of life tool was less sensitive to change than other outcomes, which are more specific to pain and limited function.

One potential source of bias is the sampling of our study population, which was affected by non-responders to the invitation and the unavailability of hip radiographs. These biases are likely to have resulted in over-sampling of individuals with poorer-functioning hips compared to the broader Perthes' population. However, we achieved a similar number of patients among all disease severities, which enables a strong analysis and takes into account the correlation between hip shape and outcomes.

Hips in this study were all classified in the healed stage of Perthes' disease, though this is often many years after the onset of disease. Huhnstock et al<sup>22</sup> employed the three-group classification using radiographs five years after disease onset and found that this was reliable at predicting long-term outcomes at this stage. It would be useful to see if long-term outcomes could be reliably determined even earlier, as the minimum time possible to accurately predict long-term outcomes would be useful to drive the minimum duration of follow-up for randomized controlled trials.

In conclusion, this study has demonstrated that hip shape was associated with long-term patient-reported outcomes. Patient-reported function deteriorated with age and the magnitude of this decline was related to the degree of hip deformity. Furthermore, we found that the three-group Stulberg classification appeared to give similar information to the five-group classification.



### Take home message

- Adolescents and young adults with hip shape with greater deviation from sphericity exhibited poorer function and increased pain at the healed stage of Perthes' disease.
- The three-group Stulberg classification appeared to give similar information to the five-group Stulberg classification.

### Twitter

Follow D. Metcalfe @TraumaDataDoc  
Follow D. C. Perry @MrDanPerry

### References

1. Marson BA, Manning JC, James M, et al. CORE-Kids: a protocol for the development of a CORE outcome set for childhood fractures. *BMJ Open*. 2020;10(2):e036224.

2. **Saran N, Varghese R, Mulpuri K.** Do femoral or salter innominate osteotomies improve femoral head sphericity in Legg-Calvé-Perthes disease? A meta-analysis. *Clin Orthop Relat Res.* 2012;470(9):2383–2393.
3. **Haywood KL, Griffin XL, Achten J, Costa ML.** Developing a core outcome set for hip fracture trials. *Bone Joint J.* 2014;96-B(8):1016–1023.
4. **Ollivere BJ, Marson BA, Haddad FS.** Getting the right answer: core outcome sets in orthopaedics. *Bone Joint J.* 2019;101-B(3):233–235.
5. **Crosby BT, Behbahani A, Olujohungbe O, Cottam B, Perry D.** Developing a core outcome set for paediatric wrist fractures: a systematic review of prior outcomes. *Bone Jt Open.* 2020;1(5):121–130.
6. **Marson BA, Craxford S, Deshmukh SR, Grindlay D, Manning J, Ollivere BJ.** Outcomes reported in trials of childhood fractures: a systematic review. *Bone Jt Open.* 2020;1(5):167–174.
7. **Leo DG, Jones H, Murphy R, et al.** The outcomes of Perthes' disease. *Bone Joint J.* 2020;102-B(5):611–617.
8. **Hailer YD, Penno E.** Agreement of radiographic measurements and patient-reported outcome in 61 patients with Legg-Calvé-Perthes disease at mean follow-up of 28 years. *J Pediatr Orthop B.* 2019;28(2):100–106.
9. **Harris WH.** Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am.* 1969;51-A(4):737–755.
10. **Neyt JG, Weinstein SL, Spratt KF, et al.** Stulberg classification system for evaluation of Legg-Calvé-Perthes disease: intra-rater and inter-rater reliability. *J Bone Joint Surg Am.* 1999;81-A(9):1209–1216.
11. **Wiig O, Terjesen T, Svenningsen S.** Inter-observer reliability of the Stulberg classification in the assessment of Perthes disease. *J Child Orthop.* 2007;1(2):101–105.
12. **Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE.** The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol.* 2014;67(5):516–526.
13. **Thissen D, Liu Y, Magnus B, et al.** Estimating minimally important difference (MID) in PROMIS pediatric measures using the scale-judgment method. *Qual Life Res.* 2016;25(1):13–23.
14. **Luo W, Ali MS, Limb R, Cornforth C, Perry DC.** Use of the PROMIS Mobility score in assessing function in adolescents and adults previously affected by childhood hip disease. *Bone Jt Open.* 2021;2(12):1089–1095.
15. **Christensen CP, Althausen PL, Mittleman MA, Lee J, McCarthy JC.** The nonarthritic hip score: reliable and validated. *Clin Orthop Relat Res.* 2003;406(406):75–83.
16. **Rosinsky PJ, Kyin C, Maldonado DR, et al.** Determining clinically meaningful thresholds for the Nonarthritic Hip Score in patients undergoing arthroscopy for femoroacetabular impingement syndrome. *Arthroscopy.* 2021;37(10):3113–3121.
17. **Herdman M, Gudex C, Lloyd A, et al.** Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727–1736.
18. **Bilbao A, García-Pérez L, Arenaza JC, et al.** Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Qual Life Res.* 2018;27(11):2897–2908.
19. **Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W.** Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain.* 2004;8(4):283–291.
20. **Society for Adolescent Health and Medicine.** Young adult health and well-being: A position statement of the Society for Adolescent Health and Medicine. *J Adolesc Health.* 2017;60(6):758–759.
21. **Stulberg SD, Cooperman DR, Wallensten R.** The natural history of Legg-Calvé-Perthes disease. *J Bone Joint Surg Am.* 1981;63-A(7):1095–1108.
22. **Huhnstock S, Wiig O, Merckoll E, Svenningsen S, Terjesen T.** The modified Stulberg classification is a strong predictor of the radiological outcome 20 years after the diagnosis of Perthes' disease. *Bone Joint J.* 2021;103-B(12):1815–1820.
23. **Castañeda P, Serrano Ardila A, Ruiz C, Mijares J.** Radiographic and MRI findings associated with early degenerative joint disease of the hip in patients with legg-calvé-perthes disease. *Revista Mexicana de Ortopedia Pediátrica.* 2016;18(2):83–88.

**Author information:**

M. S. Ali, MBChB, MRes, House Officer, Trauma and Orthopaedics Department, Aintree University Hospital, Liverpool, UK; Department of Women's and Children's Health, Institute of Translational Medicine, Liverpool Women's Hospital, Liverpool, UK.

M. Khattak, MBChB, MRes, MRCS, Specialist Registrar, Trauma and Orthopaedics Department, Alder Hey Children's Hospital, Liverpool, UK.

D. Metcalfe, PhD, MRCS, MRCP, FRCEM, Clinical Lecturer in Emergency Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.

D. C. Perry, MA (Oxon), PhD, FRCS (Orth), NIHR Research Professor, Department of Women's and Children's Health, Institute of Translational Medicine, Liverpool Women's Hospital, Liverpool, UK; Trauma and Orthopaedics Department, Alder Hey Children's Hospital, Liverpool, UK.

**Author contributions:**

M. S. Ali: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing.

M. Khattak: Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

D. Metcalfe: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing.

D. C. Perry: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing.

**Funding statement:**

The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: Versus Arthritis (grant reference 21356).

**ICMJE COI statement:**

This study has been reviewed and approved by independent members of the Perthes Association and the Steps Charity. Patient representatives from the Perthes Association and STEPS Charity contributed to the development and design of the study. D. C. Perry was funded via a NIHR Clinician Scientist Fellowship (CS-2014-14-012) and a NIHR Research Professorship. D. C. Perry is also an editorial board member for *The Bone & Joint Journal*, and was a committee member of the NIHR commissioning board from 2016 to 2021. M. Khattak was funded via a NIHR Academic Clinical Fellowship, and was the recipient of a research grant from AOUK unrelated to this study. D. Metcalfe is an editorial board member for *The Bone & Joint Journal* and is supported by both an NIHR Advanced Fellowship (NIHR302219) and the NIHR Oxford Biomedical Research Centre.

**Data sharing:**

The data that support the findings for this study are available to other researchers from the corresponding author upon reasonable request.

**Acknowledgements:**

The team are grateful to Mr Christopher Prior, Mr Roger Walton, and Mr James Widnall for their contribution.

**Ethical review statement:**

Liverpool Central Research Ethics Committee and Health Research Authority (HRA) (REC reference 17/ES/0113). Recruitment was part of the ORCHID (Outcomes Research in Children's Hip Disease) study (IRAS ID 14201).

**Open access funding:**

The open access fee for this study was covered by the Versus Arthritis grant mentioned above.

**Open access statement:**

This article is distributed under the terms of the Creative Commons Attributions (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original author and source are credited.

This article was primary edited by S. P. F. Hughes.