The cerebral palsy hip classification is reliable
AN INTER- AND INTRA-OBSERVER RELIABILITY STUDY

We have tested the reliability of a recently reported classification system of hip morphology in adolescents with cerebral palsy in whom the triradiate cartilage was closed. The classification is a six-grade ordinal scale, based on the measurement of the migration percentage and an assessment of Shenton’s arch, deformity of the femoral head, acetabular deformity and pelvic obliquity.

Four paediatric orthopaedic surgeons and four physiotherapists received training in the use of the classification which they applied to the assessment of 42 hip radiographs, read on two separate occasions. The inter- and intra-observer reliability was assessed using the intraclass correlation coefficient and found to be excellent, with it ranging from 0.88 to 0.94. The classification in our study was shown to be valid (based on migration percentage), and reliable. As a result we believe that it can now be used in studies describing the natural history of hip displacement in cerebral palsy, in outcome studies and in communication between clinicians.

Population-based studies have shown that displacement of the hip affects approximately one-third of children with cerebral palsy1-3 and has been directly related to gross motor function as determined by the Gross Motor Function Classification System (GMFCS).4 A wide variety of treatments are used in the management of this type of hip displacement ranging from bracing and botulinum injection to surgical reconstruction and salvage surgery. Understanding the natural history and clinical outcome is hampered by the lack of a valid and reliable classification system. Children in developmental dysplasia of the hip (DDH), the Severin5 classification is widely used to describe the outcome of treatment. However, the inter- and intra-observer reliability of this classification system has been questioned.6

In a recent study, a new system describing hip morphology in adolescents with cerebral palsy was proposed.7 It uses a six-grade ordinal scale based on the measurement of hip migration percentage8 and assessment of the integrity of Shenton’s arch, deformity of the femoral head, acetabular deformity and pelvic obliquity. Our aim in this study was to test the inter- and intra-observer reliability of this system.

Patients and Methods
Four orthopaedic surgeons (JGR, BJS, PS, AK) and four physiotherapists (AF, TH-I, JR, PT) were instructed in the grading of hip radiographs using the classification system.7 Two of the surgeons were fellowship trained in paediatric orthopaedic surgery and had been in practice for a mean of six years (4 to 10). The other two surgeons had completed their orthopaedic residencies and were in the final months of their paediatric orthopaedic fellowship. The four physiotherapists had a mean clinical experience of 20 years (16 to 26) and were actively involved in research and clinical work with patients with cerebral palsy. All assessors were introduced to the classification system by an oral presentation and given with written instructions and guidelines. They were given printed radiographs as well as a pen or pencil, ruler and calculator.

The hip classification. The following six grades as originally described by Robin et al7 are described below and illustrated in Figure 1.

Grade I: normal hip . A morphologically normal hip at skeletal maturity has a migration percentage < 10%. Shenton’s arch is intact and the femoral head is round and is covered by a well-developed acetabulum. This includes an everted, well-defined lateral acetabular margin, and a normal teardrop and sourcil. Pelvic obliquity is < 10°.

Grade II: nearly normal hip. A nearly normal hip has a migration percentage ≥ 10% and ≤ 15%. Shenton’s arch is intact and the femoral head is round or almost round. The
Grade I: Normal Hip—
Migration Percentage <10%
1. Shenton’s arch intact
2. Femoral head round (within 2mm using Mose circles)
3. Acetabulum – normal acetabular development with a normal horizontal sourcil, an everted lateral margin and normal tear drop development
4. Pelvic obliquity less than 10 degrees

Grade II: Near Normal Hip—
Migration Percentage ≥10% ≤15%
1. Shenton’s arch intact
2. Femoral head round or almost round
3. Acetabulum – normal or near normal development
4. Pelvic obliquity less than 10 degrees

Grade III: Dysplastic Hip—
Migration Percentage >15% ≤30%
1. Shenton’s arch intact or broken by less than or equal to 5mm
2. Femoral head round or mildly flattened
3. Acetabulum normal or mildly dysplastic including blunting of the acetabular margin and a widened tear drop
4. Pelvic obliquity less than 10 degrees

Grade IV: Subluxated Hip—
Migration Percentage >30% <100%
1. Shenton’s arch broken by more than 5mm
2. Femoral head variable deformity – Appendix 1
3. Acetabulum variable deformity – Appendix 2
4. Pelvic obliquity variable – Appendix 3

Grade V: Dislocated Hip—
Migration Percentage ≥100%
1. Shenton’s arch completely disrupted
2. Femoral head variable deformity – Appendix 1
3. Acetabulum variable deformity – Appendix 2
4. Pelvic obliquity variable – Appendix 3

Grade VI: Salvage Surgery
1. Valgus osteotomy
2. Arthrodesis
3. Excision arthroplasty (Castle) ± valgus osteotomy (McHale)
4. Replacement arthroplasty

Radiographs and diagrams showing the hip classification system in cerebral palsy. Reproduced with permission from Robin et al.7

acetabulum may be normal or only slightly deformed with the lateral acetabular margin slightly blunted. The teardrop is slightly widened, but the sourcil should be well formed. Pelvic obliquity is < 10°.

Grade III: dysplastic hip. A dysplastic hip has a migration percentage > 15% and ≤ 30%. Shenton’s arch is intact or broken by ≤ 5 mm. The femoral head may be round or slightly flattened. The acetabulum is normal or mildly dysplastic with blunt lateral acetabular margin or a gothic arch. There is a widened teardrop and a poorly developed sourcil. Pelvic obliquity is < 10°.

Grade IV: subluxed hip. A subluxed hip has a migration percentage > 30% and < 100%. There is always some contact of the femoral head with the true acetabulum. Shenton’s arch is
broken by more than 5 mm. The femoral head has variable deformity from none to severe with more than half being affected. The acetabulum likewise has variable deformity from being normally shaped to having a large gothic arch. Grade V: dislocated hip. A dislocated hip has a migration percentage ≥ 100%. There is no remaining contact between the femoral head and the true acetabulum. Shenton’s arch is completely disrupted. As in grade IV, there is variable deformity of the femoral head and acetabulum. Shenton’s arch is affected. The acetabulum likewise has variable deformity from being normally shaped to having a large gothic arch. Grade VI: salvage surgery - loss of hip. These hips have undergone salvage surgery because of painful dislocation or subluxation. Salvage surgery may involve valgus osteotomy alone, excision arthroplasty plus or minus valgus osteotomy, arthrodesis or replacement arthroplasty.

Radiographs of 42 hips of adolescents with cerebral palsy aged between 14 and 19 years and in whom the tri-radiate cartilage was closed were drawn from a large, population-based series of children with cerebral palsy born between January 1990 and December 1992 from the cerebral palsy register of the State of Victoria, Australia.9 The radiographs were selected by the first author (MLM). He had previously classified the radiographs based on the published classification system.7 Drawing from approximately 80 possible radiographs, 42 were selected in a non-randomised way to provide a balanced representation of all six grades (grade I, 3 hips; grade II, 10 hips; grade III, 6 hips; grade IV, 8 hips; grade V, 9 hips; grade VI, 6 hips). The radiographs were taken using a standardised positioning protocol employed in the Radiology Department, which has been shown to improve reliability.10 They were read on two separate occasions by all eight readers. Each reader was asked to classify each of the hips into one of six grades as defined by the classification system. All identifying information was removed from the radiographs which were presented in two sessions in a different order and separated by four weeks.

Statistical analysis. This was undertaken using the intraclass correlation coefficient. This is equivalent to the weighted kappa value in which less weight is assigned to agreement since categories are further apart.11 The intraclass correlation coefficient was interpreted using established conventions for kappa in which < 0 is poor agreement, 0 to 0.2 slight, 0.2 to 0.4 fair, 0.4 to 0.6 moderate, 0.6 to 0.8 substantial, and > 0.8 almost perfect agreement.12 In addition to inter- and intra-observer reliability, comparison was also made with an established standard of consensus grade agreed upon by the first and senior authors (MLM, HKG) who were blinded to the results of the eight readers. An asymptotic symmetry test was used to test for bias, in order to determine if one rating was consistently higher or lower than the other. All statistical analysis was performed using Stata 10.0 (StataCorp, College Station, Texas).13

Results

Intra-observer reliability. The results are summarised in Table I. The intraclass correlation coefficients ranged from 0.88 to 0.94 demonstrating excellent agreement. A symmetry test indicated that there was no bias when comparing the first and second readings of any of the raters.

Inter-observer reliability. The results are summarised in Table II. Having established the intra-observer reliability on their first and second reading, the data were analysed for inter-observer reliability. The reliability was excellent for both readings, with the mean intraclass correlation coefficient increasing slightly between the two readings (0.84 to 0.88 and 0.89 to 0.92, respectively). When taken individually, the physiotherapists showed excellent reliability, with noted improvement (although this did not reach statistical significance in the intraclass correlation coefficient between the two readings (0.84 and 0.92)). The orthopaedic surgeons were similarly excellent in terms of reliability (0.88 and 0.89).

Agreement with the established standards. The data were also analysed for agreement with the established standard (Table III). Agreement with the established standard was excellent with a mean intraclass correlation coefficient of 0.91 (0.74 to 0.96). Of 16 readings, 15 scored an intraclass correlation coefficient greater than 0.85, with one reader recording an intraclass correlation coefficient of 0.74 which is the only reading not in the ‘almost perfect’ category when compared with the established standard.

Patterns of disagreement. The data were further analysed to determine if there were individual subjects (radiographs)
which consistently demonstrated poorer agreement. Table IV highlights those which led to a difference of more than two points on the hip scale which was assigned. Radiographs 14 and 16 showed repeated disagreement between readers, three and four times, respectively. A similar analysis was carried out on subjects which had been consistently rated differently by readers and the established standard. Table V summarises these results. In all but one case, the assigned grade by the readers was higher than the established standard.

Discussion
The migration percentage as described by Reimers is the most widely used measurement of hip displacement in children with cerebral palsy. In his original report he suggested that measurement of the migration percentage was more reliable than the use of the centre-edge angle in spastic subluxation of the hip. The centre-edge angle was found to be more variable depending on the position of the child when the radiograph was taken and to be non-linear. He estimated an SEM of 10% for the migration percentage based on an estimate of the error involved in measuring a line to the nearest millimetre. In addition, two groups have studied the reliability of the measurement of the migration percentage. Using carefully standardised methodology, Parrott et al14 found that an experienced reader would be expected to measure the migration percentage on a single radiograph to within 5.8% of the true value and hence a change in migration percentage between two radiographs taken at different times to within 8.3% of the true value. More recently, Faraj, Atherton and Stott15 found that a change of more than 13% in the migration percentage, in repeated measurements by one assessor, was required in order to be 95% confident of a ‘true’ change. However, both of these studies examined only a single radiograph. In a clinical setting, most management decisions are based on the examination of a series of radiographs taken at different times.

Advantages of the use of the migration percentage are that it is easily understood, has good face validity, acceptable reliability and there is an extensive body of literature which has related specific values of this important clinical outcome.1-3,8,10,14-17 For example, a threshold value of 40% has been shown to be associated with instability and a rapid increase in the migration percentage in most hips studied.16 Most definitions of subluxation and dislocation as well as thresholds for intervention and management decisions, are commonly based on the migration percentage. A further advantage is that it can be measured from a good quality anteroposterior radiograph of the hip once the ossific nucleus is present.

Measurement of the migration percentage from serial hip radiographs has also been proposed as the most practical method of carrying out hip surveillance programmes in children with cerebral palsy.17 Given the frequency of silent hip displacement in children with cerebral palsy, a number of authors have proposed systematic hip surveillance for the early detection of hip displacement.10,17 It is therefore widely used for the detection of early hip displacement, the longitudinal follow-up of children who have been identified with hip displacement and for monitoring the outcome of various interventions.18-20 There are inherent limitations to the measurement of the migration percentage from a two-dimensional anteroposterior radiograph of the hip. Dislocation of the hip is usually posterolateral in children with cerebral palsy, but is occasionally anterior.21 Both anterior and posterior dislocations do not always show a marked increase in the migration percentage although there is almost always some degree of laterisation before the femoral head escapes from the acetabulum. Axial imaging using CT com-

Table III. Details of reader agreement with the established standards

<table>
<thead>
<tr>
<th>Rater</th>
<th>ICC (95% CI)</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading A</td>
<td>Reading B</td>
</tr>
<tr>
<td>1</td>
<td>0.94 (0.90 to 0.97)</td>
<td>0.92 (0.87 to 0.97)</td>
</tr>
<tr>
<td>2</td>
<td>0.92 (0.87 to 0.97)</td>
<td>0.93 (0.88 to 0.97)</td>
</tr>
<tr>
<td>3</td>
<td>0.94 (0.90 to 0.98)</td>
<td>0.93 (0.89 to 0.97)</td>
</tr>
<tr>
<td>4</td>
<td>0.86 (0.78 to 0.94)</td>
<td>0.87 (0.82 to 0.95)</td>
</tr>
<tr>
<td>5</td>
<td>0.96 (0.94 to 0.98)</td>
<td>0.95 (0.92 to 0.98)</td>
</tr>
<tr>
<td>6</td>
<td>0.91 (0.86 to 0.96)</td>
<td>0.96 (0.93 to 0.98)</td>
</tr>
<tr>
<td>7</td>
<td>0.90 (0.85 to 0.96)</td>
<td>0.93 (0.89 to 0.97)</td>
</tr>
<tr>
<td>8</td>
<td>0.74 (0.60 to 0.88)</td>
<td>0.89 (0.83 to 0.96)</td>
</tr>
</tbody>
</table>

* ICC, intraclass correlation coefficient; CI confidence interval

Table IV. Details of patterns of disagreement (individual subjects (radiographs) which had a difference in grade greater than 2 between their ratings). Gradings according to the classification of Robin et al

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gradings</th>
<th>Number of readers</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>II, V</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>II, IV</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>IV, VI</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>II, IV</td>
<td>1</td>
</tr>
<tr>
<td>25</td>
<td>IV, VI</td>
<td>1</td>
</tr>
<tr>
<td>38</td>
<td>II, IV</td>
<td>1</td>
</tr>
</tbody>
</table>

Table V. Details of patterns of disagreement (individual subjects (radiographs) which had a difference in grade > 2 between their rating (either A or B) and the established standard. Gradings according to the classification of Robin et al

<table>
<thead>
<tr>
<th>Subject</th>
<th>Grade</th>
<th>Established standard</th>
<th>Number of times</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>III</td>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>V</td>
<td>II</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>IV or VI</td>
<td>II</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>VI</td>
<td>IV</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>IV or V</td>
<td>II</td>
<td>3</td>
</tr>
<tr>
<td>25</td>
<td>IV or VI</td>
<td>II</td>
<td>3</td>
</tr>
<tr>
<td>38</td>
<td>I</td>
<td>III</td>
<td>1</td>
</tr>
</tbody>
</table>
implemented by three-dimensional reconstruction, and MRI offer more detailed information. However, although CT may be useful when planning reconstruction of the hip in an individual patient, neither CT nor MRI is appropriate for the classification of hips in population-based studies or in repeated longitudinal examinations in a single subject. CT examination of the hips exposes children to a significant dose of ionising radiation and MRI usually requires general anaesthesia. Given the widespread use of the migration percentage in studies of hip displacement in children with cerebral palsy, we considered that this was the key continuous variable to be used in the classification of the hip in such patients. However, with the development of severe displacement, various combinations of deformity of the femoral head, acetabular deformity (gothic arch) and pelvic obliquity make the measurement of the migration percentage increasingly difficult. The hip classification system was therefore designed to describe the morphology of hips after skeletal maturity in cerebral palsy using a combination of both quantitative (migration percentage) and qualitative features. The classification, based on the use of a continuous variable migration percentage was consciously based on the GMFCS for children with cerebral palsy and its relationship to the gross motor function measure.

The study by Ward et al of the Severin system for DDH found unacceptably low intra- and inter-observer reliability. This re-inforces the importance of establishing the validity and reliability of a new classification system before its widespread acceptance and application. Previous studies have discussed the influence of the expertise of the readers on interobserver agreement. In our study, both physiotherapists and orthopaedic surgeons served as readers. There was a spectrum of clinical experience in both groups. There was a subtle increase in the interobserver reliability between the first (A) and second (B) reading which could be attributed to improved application of the system secondary to learning from the experience of the first reading. In order to minimise the influence of recall bias, the radiographs were presented four weeks apart in addition to being blinded and presented in a different order.

In our study the classification has been shown to be valid and reliable when used by experienced physiotherapists and paediatric orthopaedic surgeons. As a result it can now be used to describe the natural history and outcome of hip displacement in cerebral palsy. Basing the classification on the migration percentage allows comparison with studies reporting a change in the migration percentage in younger children. However, these rarely show the morphological features which we found to be part of hip disease in adolescents with cerebral palsy. We believe that the changes in hip morphology accelerate during the adolescent growth spurt. Although the rate of change may slow with the onset of skeletal maturity, it is known that scoliosis, pelvic obliquity and changes in hip status may occur in the third and fourth decades.

The classification may be useful as a means of communication regarding the clinical status of hip morphology in adolescents with cerebral palsy as well as a tool for evaluating the
Anteroposterior radiograph of the hips of a 16-year-old girl with spastic quadriplegia, gross motor function classification system level V. The right hip (grade VI) has been excised for intractable pain. However, there has been proximal migration with contact between the femur and pelvis causing pain. The left hip is grade II with a mildly flat femoral head which is completely covered (migration percentage 0%) and there is moderate pelvic obliquity of 30°.

outcome of treatment and natural history studies. The combined use of the GMFCS and the classification may also be useful. For example, adolescents at GMFCS level I and II, who are long-term independent community walkers may require a normal or near morphologically normal hip (Fig. 2). Those with lower levels of mobility or who are non-walkers may be able to function with less normal hip development (Fig. 3). Grade-VI hips have been included in the classification to allow the documentation of hips which have been painful and required excision as opposed to those which are dislocated (grade V), but have not yet required salvage (Fig. 4). The relationship between hip morphology, pain and sitting and standing function deserves further exploration.

Our results show that the hip scale in cerebral palsy is a reliable tool for the evaluation of radiological findings of hip morphology in adolescents with cerebral palsy in whom the triradiate cartilage is closed. The interobserver reliability was excellent, showing consistent agreement between readings for individual raters.

Supplementary material

Appendices relating to the classification of deformity of the femoral head, acetabular deformity and pelvic obliquity are available with the electronic version of this article on our website at www.jbjs.org.uk

We wish to thank A. Fosang, T. Hastings-Ison, J. Rodda and P. Thomason for their help in grading the radiographs. We also acknowledge the help of M. Sheedy in the preparation of this manuscript.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References