Management of unstable slipped upper femoral epiphysis: a meta-analysis

Stefanie Lowndes†, Anil Khanna‡, David Emery†, Julius Sim‡, and Nicola Maffulli§*

†Trauma and Orthopaedic Surgery, University Hospital of North Staffordshire, Stoke on Trent, UK; ‡School of Health and Rehabilitation, Keele University, Stoke on Trent, UK, and §Barts and The London School of Medicine and Dentistry, London, UK

Introduction: The management of unstable slipped upper femoral epiphysis (SUFE) is controversial, with a high risk of developing avascular necrosis (AVN). We meta-analysed two areas of concern: reduction of the slip and the timing of treatment.

Methods: A search of Medline, CINAHL and Embase identified only retrospectively relevant studies: four regarding the role of reduction and five regarding the timing of treatment. The incidence of AVN was compared between reduced and unreduced SUFEs, and between those treated within 24 h of symptom onset and those treated thereafter.

Areas of agreement: Analysis of the pooled data gave an odds ratio of 2.20 ($P = 0.290$) in favour of the unreduced group, who had a lower risk of developing AVN. The odds ratio was 0.50 in favour of the group treated within 24 h from symptom onset ($P = 0.441$). However, though clinically important, these effects were not statistically significant.

Areas of controversy: The timing of treatment is somewhat inconsistent: two studies favour management more than 24 h after the onset of symptoms, while for three unstable SUFEs are best managed within 24 h.

Growing points: Despite the non-significant results from the meta-analysis, it can be suggested that, if reduction is to be performed, it should be undertaken cautiously, as it may be associated with increased AVN. The ideal time for management of unstable slip is probably within 24 h of symptom onset.

Areas timely for developing research: There is a strong need for multicentre, randomized, controlled trials in this area.

Keywords: slipped upper femoral epiphysis/reduction/avascular necrosis/meta-analysis
Introduction

Slipped upper femoral epiphysis (SUFE) is a serious juvenile hip disorder whose appropriate clinical management has been a topic of debate for many years.\(^1\) The classification criteria developed by Loder et al.\(^2\) have revolutionized the concept of the SUFE. These criteria are based on whether the epiphysis is stable or unstable. The previous practice of classifying SUFEs based on the duration of symptoms—which divides slips into acute, chronic and acute-on-chronic subgroups—is prognostically unhelpful.\(^3\) According to Loder et al.,\(^2\) 95% of SUFEs are stable and these have a minimal complication rate. In contrast, 5% of SUFEs that are unstable have a much more substantial complication rate, which has been reported to be as high as 84%.\(^2,4\) The primary and most serious complication of SUFE is avascular necrosis (AVN) of the head of the femur.

The management of the stable slip is relatively straightforward and uncomplicated. The management of the unstable SUFE, however, is a controversial topic. Owing in part to the rarity of the condition, high-quality literature regarding the optimum treatment pathway is not available, yet the high complication rate in unstable SUFEs makes this a very important clinical issue.

There is no universally accepted way of managing these slips. The literature is divided as to which interventions contribute to, or alternatively serve to minimize, the risk of AVN. This review assesses two main areas of debate relating to the management of unstable slips and their effect on rates of AVN in particular; namely, the role of reduction and the timing of treatment.

Most papers that assess the management of unstable SUFE provide information of limited value due to extremely small patient samples. This study therefore aims to enhance the value of these individual studies by pooling the available evidence through a meta-analysis, allowing a larger patient sample to be considered than is possible from any single study.

Materials and methods

Inclusion criteria

The objective of this study was to review the published data regarding different management options and their impact on the prognosis of unstable SUFE. The outcome of the different management pathways was assessed by examining the resultant rates of AVN. Only studies using the Loder et al.\(^2\) criteria of instability were included in the
analysis, owing to their prognostic superiority to criteria based on chronicity of symptoms, and to ensure that the patient groups of different studies were comparable. According to these criteria, the hip is considered to be unstable when the magnitude of the pain experienced prevents the patient from weight bearing, even with the assistance of crutches.\textsuperscript{2}

Although several methods have been described historically to treat SUFEs, each tending to report a slightly different risk relating to the development of AVN, only results relating to hips that had been ultimately managed by internal fixation were considered, as the literature suggests this to be the most modern and least risky management option.\textsuperscript{5}

The following questions concerning the management pathway of unstable SUFEs were considered, in relation to the development of AVN.

- Is it possible to reduce an unstable SUFE and not increase the risk of AVN?
- Is there an optimum time for treatment that has the lowest rate of AVN?

As most individual studies do not address both of these issues, a subgroup of studies relating to each question was formed and each subgroup was assessed independently.

**Search strategy**

For electronic searches (June 2008), the following search terms were utilized to identify relevant papers:

1. MeSH term: ‘slipped epiphysis’,
2. ‘Slipped upper femoral epiphysis’ or ‘slipped capital femoral epiphysis’ or ‘SUFE’ or ‘SCFE’,
3. (1) or (2),
4. ‘Reduction’ or ‘reduce’,
5. MeSH term: ‘femur head necrosis’,
6. ‘Avascular necrosis’,
7. (5) or (6),
8. ‘Acute’ or ‘unstable’,
9. (3) and (4) and (7) and (8).

Limiters were applied to the search. Only studies that were specific to humans were searched for. Non-English language articles were also excluded from this search. The search was performed on Medline, EMBASE and the Cochrane Controlled Trials Register. Following these initial database searches, the bibliographies of any identified articles
were searched and similarly the bibliographies of any papers identified in this manner. This was intended to ensure that any papers missed in the database searches were ultimately identified. Finally, paper copies of relevant journals from the previous 6 months were manually searched to collect any articles not yet electronically categorized. The search results were screened to remove any duplicate studies.

**Selection of studies**

For inclusion in the meta-analysis, each study needed to provide data on one or both of the following comparisons: (i) patients who had had their SUFEs reduced and those who had not and (ii) patients treated within 24 h and those treated later than 24 h. The electronic search strategy described above identified 32 potentially relevant articles from Medline and an additional two from EMBASE (after screening out duplicate papers). These were articles relating either specifically to the management of unstable SUFEs, or to SUFEs in general, with specific consideration of the unstable subgroup. Additional studies were found through the manual search method described previously. No Cochrane reviews were identified.

The reports identified were then evaluated for their relevance and their appropriateness for inclusion in a meta-analysis. The papers were examined to identify any studies that used the same patient sample to consider different aspects of the management of the unstable SUFEs; in such cases the sample was included from one report only.

Following application of the aforementioned inclusion and exclusion criteria, the final number of studies deemed appropriate for inclusion in this review was eight, published between 1993 and 2007. All were case–control studies in the form of retrospective reviews. No randomized controlled trials were identified. From these eight studies, information was extracted on patients with unstable SUFEs and the overall incidence of AVN. Information relating to reduction and timing of treatment was then collected, along with the incidence of AVN for these subgroups. Only one study had appropriate information regarding both reduction and treatment timing; the remainder had information relating to one question only.

**Statistical analysis**

The relative incidence of AVN was determined (i) between reduced versus unreduced epiphyseal slips and (ii) between treatment within 24 h versus beyond 24 h, and was expressed as odds ratios, with 95% confidence intervals (CIs). Where a cell in the contingency table had a frequency of zero (whereby the resulting odds ratio will have a value of
either zero or infinity), the odds ratio was calculated after first adding 0.5 to each cell frequency.\textsuperscript{6} An odds ratio of unity indicates no difference between the comparison groups in the odds of developing AVN.

Heterogeneity of the individual odds ratios was assessed by calculating the $I^2$ statistic. This expresses the percentage of the total variation in estimated effects across studies that is attributable to true heterogeneity in these estimates rather than to chance fluctuation; values of 25, 50 and 75\% are considered to represent low, moderate and high heterogeneity, respectively.\textsuperscript{7} The odds ratios from the individual studies within each subgroup were then pooled using a fixed effects model in the absence of heterogeneity or a random effects model where heterogeneity was detected. A fixed effects model assumes that studies are homogeneous, and that each study is estimating a common ‘true’ effect in the population, whereas a random effects model assumes that individual studies are estimating a distribution of differing population effects, which in turn may represent different subcategories of study.\textsuperscript{8} The contribution of individual odds ratios to the pooled odds ratio was weighted by the inverse variance method (fixed effects)\textsuperscript{9} or the DerSimonian–Laird method (random effects),\textsuperscript{10} such that the more precise study estimates had greater weight in determining the pooled estimate. Additionally, study reports were examined for evidence of clinical heterogeneity in relation to the underlying study populations, the diagnostic criteria used and the interventions performed.

\section*{Results}

Of the eight suitable studies included in the meta-analysis, four provided information regarding reduction and five provided information regarding timing of management. One report (Kennedy \textit{et al.})\textsuperscript{11} contained data relating to both comparisons and was thus appropriate for inclusion in the analysis for both questions. Table \textsuperscript{1}\textsuperscript{12}–\textsuperscript{17} shows descriptive information available from the eight studies included in the review.

\textbf{Is reduction associated with an increased risk of AVN?}

There were four studies that provided the necessary information to allow comparison of outcome in unstable SUFEs that had been either reduced or left unreduced (Table 2). From these studies, there were a total of 85 patients with unstable SUFEs; 49 underwent reduction, 36 did not. Not all of this information was explicitly provided in each of these studies; in some cases it had to be inferred from other
information. The data extracted from each of the studies are shown in Table 2, together with odds ratios for the development of AVN. There was no evidence of heterogeneity in the odds ratios from individual studies ($I^2 = 0\%$) and the overall odds ratio was therefore derived through a fixed effects model. The odds ratios are displayed graphically in Figure 1.

The odds ratio for the pooled data was 2.20 (95% CI 0.51–9.54; $P = 0.290$), which indicates that the odds of developing AVN were over twice as large following reduction as when no reduction was undertaken. However, as the 95% CI contains unity, this difference is not statistically significant at $P \leq 0.05$; we cannot conclude that there is a true difference in the development of AVN between reduction and non-reduction.

**Does the time of treatment affect the overall outcome following an unstable slip?**

There were five studies that provided the necessary information to allow comparison between patients with unstable SUFEs who were treated within 24 h of onset of symptoms and those treated more than 24 h after symptom onset. From these studies, there was a total sample
of 130 unstable SUFEs; 56 were treated within 24 h and 74 were treated after 24 h of symptom onset. Again, in some studies, this information had to be inferred from other details provided in the study.

The data extracted from each of the studies are shown in Table 3, together with odds ratios for the development of AVN. The odds ratios are also displayed graphically in Figure 2. Owing to moderate heterogeneity of the odds ratios from individual studies ($I^2 = 60\%$), the overall odds ratio was derived through a random effects model.

The overall odds ratio was 0.50 (95% CI 0.09–2.92; $P = 0.441$), indicating that the odds for developing AVN if treatment occurs within 24 h were half than those of developing AVN if treatment occurs beyond this point. Again, however, this difference is non-significant, as the CI includes unity.

**Discussion**

This meta-analysis compiled and analysed the available published evidence regarding the management pathway for the unstable SUFE. The meta-analysis looked at two controversial issues regarding the management of this condition: (i) reduction and (ii) timing of *in situ* fixation.
Four studies were identified as providing appropriate information relating to AVN rates in unstable SUFEs that had been reduced compared to those that had not and five studies were deemed to provide such information regarding AVN rates in unstable SUFEs managed within 24 h of symptom onset versus those managed after 24 h of symptom onset. It must be acknowledged that the retrospective studies on which this review is based are susceptible to bias and are therefore classified as rather low in the hierarchy of levels of evidence for treatment.

Table 3  Comparative rate of AVN in reduced and unreduced unstable SUFEs for treatment \( \leq 24 \) h and treatment \( >24 \) h from symptom onset.

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Whole sample</th>
<th>Treatment ( \leq 24 ) h</th>
<th>Treatment ( &gt;24 ) h</th>
<th>Odds ratio (95% confidence interval)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peterson et al.</td>
<td>41</td>
<td>37 4</td>
<td>25 1</td>
<td>12 3</td>
<td>0.16 (0.02, 1.70)</td>
</tr>
<tr>
<td>Gordon et al.</td>
<td>16</td>
<td>14 2</td>
<td>10 0</td>
<td>4 2</td>
<td>0.09 (0.00, 2.17)</td>
</tr>
<tr>
<td>Kennedy et al.</td>
<td>27</td>
<td>23 4</td>
<td>7 2</td>
<td>16 2</td>
<td>2.29 (0.27, 19.66)</td>
</tr>
<tr>
<td>Loder et al.</td>
<td>30</td>
<td>16 14</td>
<td>1 4</td>
<td>15 10</td>
<td>6.00 (0.58, 61.84)</td>
</tr>
<tr>
<td>Kalogrianitis et al.</td>
<td>16</td>
<td>8 8</td>
<td>5 1</td>
<td>3 7</td>
<td>0.09 (0.01, 1.08)</td>
</tr>
<tr>
<td>Total</td>
<td>130</td>
<td>98 32</td>
<td>48 8</td>
<td>50 24</td>
<td>0.50 (0.09, 2.92)</td>
</tr>
</tbody>
</table>

*Reference category for odds ratio = treatment \( >24 \) h.

Fig. 2  Forest plot of odds ratios, with 95% CIs, for development of AVN in relation to treatment \( \leq 24 \) h versus treatment \( >24 \) h. The overall effect is indicated by the diamond marker and those for individual studies by square markers, whose size is proportional to the weighting of the study concerned in determining the overall effect. The vertical dashed line indicates an odds ratio of unity (no difference).
effectiveness. However, there is currently no higher level of information available regarding unstable SUFEs, and moreover the feasibility of randomized controlled trials to examine the questions addressed in this review would be constrained by the comparative rarity of AVN as an outcome.

In summary, while the results from the two analyses showed a difference between the two comparison groups, in neither case was the effect statistically significant. That is, while a difference was found in the total sample of patients included within the meta-analysis, this cannot be generalized to the population of such patients. Nonetheless, the magnitude and direction of the observed effects warrant some comment.

Reduction and avascular necrosis

Three of the four studies included in the analysis of reduction when compared with no reduction showed an odds ratio favouring no reduction (the odds ratio from the remaining study showed no difference). The overall estimate showed the odds of AVN among patients who had been reduced to be more than double the odds among those whose SUFE was unreduced, though this difference was not statistically significant. Moreover, it should be interpreted with caution on other grounds. Rattey et al. suggested that their findings relating to AVN and reduction in acute SUFEs may have been affected by the fact that severe slips are more likely to be reduced; the severe slip is more mobile and lends itself better to a successful reduction. If slips that are reduced tend to be more severe, this would inflate the incidence of AVN among reduced SUFEs, as these severe slips have a greater incidence of AVN. Further studies on the effect of reduction on the development of AVN should therefore seek to control for slip severity.

The mechanism whereby increased slip severity may result in an increased risk of AVN is related to the debate regarding the cause of impaired blood supply to the femoral head. Damage to the blood supply of the femoral head has long been considered to be an iatrogenic phenomenon. Maeda et al. published results of a study in 2001 in which they performed angiography of the blood supply to the femoral head before and after reduction or treatment. This showed that in three of the five unstable SUFEs, there was impaired filling of the superior retinacular arteries before any reduction or treatment was initiated. In these cases, therefore, it was demonstrated that the vascular damage was a direct result of the injury, rather than being an iatrogenic phenomenon.
It is important to stress in any discussion relating to reduction, that forceful manipulative reduction is never appropriate in SUFEs.\textsuperscript{20} With increased force of reduction comes diminished control and damage to the vascular supply is almost inevitable.\textsuperscript{14,19} Most of the studies in this analysis admit to having reduced their slips inadvertently (i.e. the reduction occurred simply due to the general anaesthetic and positioning for surgery) or through closed reduction. In one of the studies, a minority of the slips required open reduction.\textsuperscript{11} Even though it is commonly acknowledged that forceful manipulative reduction contributes to an increased incidence of AVN, this tendency may extend to other methods of reduction. It was not possible to examine the rates of AVN associated with different methods of reduction within this review. This is a topic that needs further study.

Despite the recent controversies regarding the role of reduction, the majority of authors still favour reduction, as is reflected in the data reported in Table 1, where more unstable SUFEs were reduced than were not. The reason that reduction is favoured is that it restores, to some extent, the normal anatomy of the hip. The anatomy of the hip must be grossly distorted to produce the extent of pain that is associated with unstable SUFEs, and that makes walking impossible. If a more normal position of the hip is achieved by reduction prior to internal fixation, there is less chance of severe degeneration occurring later, and normal function will likely be maintained for longer, assuming AVN is avoided.\textsuperscript{3} Importantly, it also significantly decreases the likelihood of needing reconstructive surgery following the slip, which in itself carries a further risk of AVN.\textsuperscript{21}

Although this analysis produced a non-significant overall effect, the magnitude of the pooled odds ratio is relatively large, suggesting a clear need for further research on this question, involving large studies.

**Timing of treatment: association with AVN**

The five studies within the analysis based on the timing of screw fixation were less consistent in their findings, but the results of the meta-analysis showed that treatment within 24 h was associated with a 50% reduction in the odds of AVN relative to treatment later than 24 h. Although this effect was not statistically significant, clinically it is of a potentially important magnitude.

In their series of unstable SUFEs, Loder \textit{et al.}\textsuperscript{2} hypothesize that the more severe slips presented earlier, therefore allowing treatment to be commenced within 24 h of onset of symptoms. More severe slips, as already noted, have a greater risk of AVN. Unfortunately, data are as yet unavailable to allow an analysis of the combined effect of slip
severity and timing of treatment on the rate of AVN. However, as the incidence of AVN in the studies reviewed was lower with early management than with late management, despite the tendency for more severe slips to present earlier, it is reasonable to surmise that early management helps to counteract the development of AVN.

There remains, however, the question of what to do if a patient first presents beyond 24 h of symptom onset. Kalogrianitis et al.\textsuperscript{17} found that all but one of their patients who developed AVN did so following treatment that occurred between 24 and 72 h of symptom onset, whereas none of those treated either side of this period did so. They infer from this that treatment within 24 h is ideal, but that if this is not possible, treatment should be delayed for at least a week. These authors postulate that the reason for this ‘unsafe window’ of time is related to the inflammatory process occurring within the hip. Effusion and inflammation contribute to the compromise of the vascular supply via a tamponade effect and it takes at least 1 week for this to settle. Hence, they suggest that surgical intervention, which risks disrupting the blood supply further, should occur either before the inflammatory response has become established or after it has subsided.\textsuperscript{17} Clearly, further research, which would lead to more definitive guidance on the timing of treatment, is required.

\textit{Limitations of the present study}

Although this meta-analysis can be expected to have produced more robust findings than those of any of the individual studies already performed, the number of studies that were deemed appropriate for analysis, and hence the number of patients on which the meta-analysis was based, are small. The power of the meta-analysis to detect a significant association was thereby limited. Any research regarding unstable SUFEs is hampered by their infrequent occurrence, and a single definitive study on one or other of the questions examined in this review might well be prohibitive. For example, a randomized study seeking to detect an odds ratio of 2.0 for the development of AVN between the two methods of management would require approximately 660 patients (based on the incidence of AVN reported in this review and assuming 80\% power and a two-tailed 5\% significance level). In the absence of a multicentre randomized study with a sample of this size, pooling estimates of effect from smaller individual studies, may therefore represent the most feasible way of examining different methods of managing SUFE.

Although the timing of treatment was taken from the onset of symptoms, one included study (Peterson \textit{et al.})\textsuperscript{12} measured the time from
presentation. It was decided to include this study in the current analysis. However, the lack of complete information regarding the time from symptom onset at which treatment occurred raises the possibility of ascertainment bias regarding the timing of treatment.

Owing to limited information provided in some of the primary studies, this meta-analysis was unable to consider the interrelationship between certain key factors. Thus, any conclusions regarding the effect of reduction versus non-reduction, or of intervention within versus beyond 24 h, may be confounded by between-group differences in the severity of the slips. Equally, the effect of treatment timing may be distorted by a between-group difference in the proportion of unstable slips that were reduced prior to fixation.

**Recommendations**

Given the lack of statistically significant results from this meta-analysis, no definitive clinical recommendations can be formulated. Nonetheless, given the magnitude of the odds ratio for AVN in relation to reduction, and in the absence of more definitive data, it is prudent to advise that if unstable SUFEs are to be reduced, such reduction should be performed with caution, owing to the possible increased risk of AVN.

There is no clear evidence as to the best time to manage unstable SUFEs, but what evidence there is suggests that if reduction is to occur, it is best performed sooner rather than later. This is also in line with what we know of the pathology of AVN. There is, however, a suggestion that treatment that cannot be instituted immediately should be delayed for some days; this issue requires clarification through future studies.

This meta-analysis has demonstrated that the data available regarding unstable SUFEs are inadequate to allow any extensive recommendations regarding optimal management to be made. The evidence that is available is low level; a large multicentre study, with adequate control for confounding factors such as slip severity, would be hugely beneficial in this particularly contentious area of orthopaedics.

**Conclusions**

A meta-analysis of four studies describing the effects of reduction on rates of AVN in unstable SUFEs found there to be no statistically significant difference between the reduced and unreduced groups. The magnitude of the observed odds ratio does, however, raise the possibility that reduction may have a clinically important effect on the
development of AVN. Reduction may be necessary to prevent loss of function and premature degeneration of the femoral head, which would require management by surgery, which itself carries an inherent risk of AVN. Reduction should, however, be undertaken with extreme caution.

A meta-analysis of five studies considering the optimal time of management of unstable SUFEs found treatment within 24 h from the onset of instability to be associated with a lower risk of AVN than treatment beyond this time. Again, the odds ratio was statistically non-significant, but of a clinically important magnitude.

In relation to both of these questions, there is a need for further research, so that more definitive recommendations can be made as to the appropriate management pathway for SUFE.

**References**