Indometacin as prophylaxis for heterotopic ossification after the operative treatment of fractures of the acetabulum

Our study was designed to compare the effect of indometacin with that of a placebo in reducing the incidence of heterotopic ossification in a prospective, randomised trial. A total of 121 patients with displaced fractures of the acetabulum treated by operation through a Kocher-Langenbeck approach was randomised to receive either indometacin (75 mg) sustained release, or a placebo once daily for six weeks. The extent of heterotopic ossification was evaluated on plain radiographs three months after operation. Significant ossification of Brooker grade III to IV occurred in nine of 59 patients (15.2%) in the indometacin group and 12 of 62 (19.4%) receiving the placebo.

We were unable to demonstrate a statistically significant reduction in the incidence of severe heterotopic ossification with the use of indometacin when compared with a placebo (p = 0.722). Based on these results we cannot recommend the routine use of indometacin for prophylaxis against heterotopic ossification after isolated fractures of the acetabulum.

Heterotopic ossification (HO) is commonly described as a complication after the operative treatment of fractures of the acetabulum.\textsuperscript{1-12} The incidence ranges from 24% to 90% with greater frequency and severity with posterior or extensile approaches to the acetabulum.\textsuperscript{7-13} Large amounts of ectopic bone around the hip have been shown to result in severe limitation of the range of movement of the hip with a decreased functional outcome.\textsuperscript{2-5}

Indometacin has been used as prophylaxis for HO with several retrospective studies showing a reduction in bone formation of clinical significance.\textsuperscript{3-5} The results of recent studies are mixed, with some authors showing a dramatic decrease in the incidence of HO and others finding little or no effect.\textsuperscript{5,8-11} The null hypothesis of our study was that there was no difference in the development of severe HO between patients receiving indometacin and those having a placebo.

Patients and Methods

We performed a prospective double-blind controlled clinical trial with the approval of the Institutional Review Boards at two level-I trauma centres to evaluate the efficacy of indometacin compared with a placebo for prophylaxis against HO after the operative treatment of fractures of the acetabulum. Between January 1999 and June 2003, 232 patients with such fractures were treated by operation through a posterior approach. The criteria for exclusion were: age under 18 years, injury to the spinal cord, ankylosing spondylitis, burns, gastro-intestinal bleeding, Glasgow coma scale < 15, cerebrovascular accident and the use of other non-steroidal anti-inflammatory drugs.\textsuperscript{14}

There were 157 patients who were eligible, and 127 were enrolled in the trial, and underwent operative stabilisation of their acetabular fractures through a posterior Kocher-Langenbeck approach.\textsuperscript{9} After fixation and before wound closure, any devitalised muscle was excised.\textsuperscript{12}

The patients were randomised to receive either indometacin (Merck Inc., Whitehouse Station, New Jersey) or a placebo using a computer-generated list. Indometacin (75 mg) sustained release and the placebo were administered to the patients in a blinded fashion by the investigational drug pharmacy. The sustained-release formulation was selected for our study to improve patient compliance since it is given in a single daily dose, and to reduce the costs associated with formulating the placebo. The manufacturer (Merck Inc.) states that “when measured over a 24-hour period, the cumulative amount and time-course of indometacin absorption from a single capsule of Indocin SR are comparable to those of three doses of 25 mg capsules of Indocin”.

©2006 British Editorial Society of Bone and Joint Surgery
doi:10.1302/0301-620X.88B12.18151
There were 63 patients in the indometacin group and 64 in the placebo group. The medication was taken once daily for six weeks. Patient compliance was measured by interview and by obtaining the serum level of indometacin at the first post-operative visit after two weeks. The serum levels were obtained from patients in both treatment groups to maintain the blinded design of the study. The specimens from the indometacin group were unblinded and processed by a research co-ordinator independent of the treating physicians.

The age, gender, height, weight, body mass index (BMI) and injury severity score (ISS) were recorded for each patient at the time of admission.\textsuperscript{15,16} All the fractures were classified using the Orthopaedic Trauma Association classification and that of Letournel and Judet.\textsuperscript{9,17} Standard radiological evaluation including anteroposterior, iliac oblique and obturator oblique views was performed at two, six and 12 weeks after operation.

The primary outcome measurement of the study was the presence of HO according to the classification of Brooker et al\textsuperscript{18} as seen on the radiographs taken at 12 weeks. This time was selected since radiological evidence of HO is usually seen as early as three to six weeks with maximum bone formation by 12 weeks.\textsuperscript{4,5,11,12,15,19,20} The radiographs were scored for the presence of HO using the classification as modified by Moed and Smith.\textsuperscript{21} Grade 0 signifies no HO, grade I describes islands of bone within the soft tissue around the hip, grade II shows bone spurs from the proximal femur or pelvis with more than 1 cm between the opposing spurs, grade III indicates bone spurs from the proximal femur or pelvis with less than 1 cm between the opposing spurs and grade IV represents complete bony ankylosis of the hip. The radiographs taken at 12 weeks were assessed in a blinded fashion by MAK, MJB, JAG, SHS, JFK; and a consensus score was obtained and recorded.

**Statistical analysis.** The data were analysed using SPSS statistical software (SPSS Inc., Chicago, Illinois) and the ‘intention-to-treat’ principle, in which all the enrolled subjects with Brooker scores after 12 weeks were included.\textsuperscript{22-25} The basic premise of an intention-to-treat analysis is the inclusion of all randomised patients in the treatment groups to which they were assigned, regardless of their adherence to the entry criteria, of the treatment which they actually received and of subsequent withdrawal or deviation from the protocol.\textsuperscript{22} Descriptive statistics were used to describe sample clinical details and to identify significant differences in measurements by centre, age and gender. For the descriptive analysis, categorical variables such as the Brooker score, hip dislocation and type of fracture were compared using Pearson chi-squared tests. Continuous variables such as age, ISS and surgical delay were compared using independent sample t-tests, or one-way analysis of variance (ANOVA) when appropriate. For testing the association between two categorical variables, Fisher’s exact test was used whenever the observed number of subjects in any category fell below ten. A non-parametric Mann-Whitney U test was used to compare the median of the distribution of the Brooker grade across the two treatment groups. A p-value < 0.05 was considered to be significant. For understanding the association between the severity of HO and various risk factors, the Brooker grades were classified into three categories as follows: none (grade 0), mild (grade I and grade II), and severe (grade III and grade IV). A multivariate ordinal logistic regression of severity status was performed while controlling for treatment group, gender, type of fracture, age, ISS and surgical delay.

**Results**

There were no significant differences between the two centres with respect to age, BMI, ISS, gender, type of fracture or complications, or between the two treatment groups with respect to age, gender, BMI, ISS, type of fracture, mean time to surgery and overall incidence of complications.

\begin{table}
\centering
\begin{tabular}{|l|l|l|l|}
\hline
 & Indometacin (n = 63) & Placebo (n = 64) & p-value \\
\hline
Mean age in yrs (range) & 37 (18 to 71) & 39 (18 to 76) & 0.35 \\
Number of men (%) & 49 (78) & 51 (80) & 0.79 \\
Mean ISS\textsuperscript{*} (range) & 14 (5 to 55) & 12 (4 to 30) & 0.13 \\
Number with hip dislocation (%) & 37 (59) & 51 (80) & 0.01 \\
Mean time to operation in days (range) & 5 (0 to 17) & 4 (0 to 21) & 0.40 \\
\hline
\end{tabular}
\caption{Clinical details of the indometacin and placebo groups}
\end{table}
There were significantly more dislocations of the hip in the placebo group \( (p = 0.01; \text{Table I}) \).

**Incidence of heterotopic ossification.** There were 121 patients with sufficient follow-up to be included in the analysis. Six were lost to follow-up, four in the indometacin group and two in the placebo group, leaving 59 in the former and 62 in the latter group. Figure 1 presents the relative distribution of Brooker grades in the two groups. The most frequent grade in both groups was 0 and the observed median grade in both was I. Noting that the Brooker grade represented an ordinal rating of severity of HO, we carried out the non-parametric Mann-Whitney U test for the medians in the two groups. The resulting \( p \)-value was 0.11, indicating no significant difference between the groups.

The overall proportion of HO in the study group as a whole was 53\% (64) and of severe HO, 17.3\% (21). There were 27 patients (45.8\%) in the indometacin group with evidence of HO and 37 (59.7\%) patients in the placebo group. Nine patients (15.2\%) in the indometacin group and 12 (19.4\%) in the placebo group had severe HO. The 95\% confidence interval (CI) of the difference in the proportions of severe HO between the indometacin and the placebo groups was -9.4\% to +17.6\%, indicating no statistically significant difference between the two groups (chi-squared test, \( p = 0.722 \)). Fisher's exact test revealed no significant association between the Brooker categories (none, mild, severe) and the treatment groups \( (p = 0.334) \).

**Multivariate analysis of heterotopic ossification.** A multivariate ordinal logistic regression analysis was carried out with severity as outcome and the treatment group as the primary covariate of interest. Age, ISS, surgical delay, gender, dislocation and type of fracture were included in the model as additional covariates. Interactions between group and dislocation, group and surgical delay and dislocation and surgical delay were also controlled for in the regression model. Group, type of fracture, dislocation, age, and the interaction between group and dislocation were significant at the 5\% level. The main effect of group being significant seemed to be in apparent contrast to the findings in the univariate analyses. However, this had to be interpreted with caution given the significance of the interaction between the group and dislocation, once this was removed from the model the main effect due to the group lost statistical significance. Thus, it appeared that the main effect of the group was substantially confounded by its interaction with dislocation. We therefore decided to report results from separate logistic regression fitted within subgroups determined by dislocation.

Table II gives the odds ratios and associated 95\% CIs based on the multivariate logistic regression analyses stratified by dislocation. For subjects without dislocation of the hip, the placebo group were approximately ten times more likely to be in the severe HO category than the indometacin group \( (p = 0.008) \). Also, for this subgroup of patients, those with an associated type of fracture were approximately 12 times more likely to be in the severe group. For the subgroup with dislocation of the hip, however, the only significant association was found with surgical delay, with higher delay being associated with a severe HO category. Neither the group variable nor type of fracture was significant at the 5\% level.

**Compliance.** Of the original 127 patients 97 (76\%) stated that they had medication for six weeks. Six patients were lost to follow-up and 24 did not complete the entire six weeks because of side effects in 14 and loss of medication in ten (Table III). The difference in the proportion of patients who completed medication for six weeks between the two treatment groups was significant (Fisher's exact test, \( p = 0.003 \)).

**Levels of indometacin.** The serum levels of indometacin were determined in 37 (59\%) of the 63 patients who received this drug. Of the 37 tested, 21 (57\%) had measurable levels and 16 (43\%) did not. There were 26 patients (41\%) in whom the serum indometacin level was not assessed because of refusal to have the test or because of an improperly processed sample. Table IV gives the grade of HO as judged against the level of indometacin. Of the 21 patients with measurable levels, 12 developed HO, which was severe in three, and nine did not. Of the 16 in whom the levels were not detected, seven developed HO, which was severe in four, and nine did not. Seven of the 13 patients who withdrew from the indometacin group because of side-effects had levels of serum indometacin assessed. Four of the seven had detectable levels, and one of these had a gastro-intestinal bleed. A patient with a perforated ulcer was treated with indometacin, but did not have the levels tested. One patient who received indometacin developed nonunion of a fracture of the tibia, but did not have detectable serum levels when assessed.
The best indication for treatment was when the difference in the proportion of severe HO between the placebo and treatment groups could be considered to be clinically important. The 95% CI for the difference in the proportions of severe HO between the placebo and treatment groups which we reported was -9.4% to +17.6%. This can be interpreted as follows: The treatment effect of indometacin was somewhere between 9.4% worse or 17.6% better than the placebo. Whether this was significant or not could not be determined from the data. Whether this was significant or not depended on the clinical application of these data. Alternatively, these data may be examined using the concept of the number needed to treat. The best indication for treatment with indometacin was demonstrated by the upper limit of the CI. If this was used to calculate the number needed to treat, the data suggest that six patients would have to be treated with indometacin for one to benefit. However, if this was based on the 4% difference demonstrated between the indometacin and placebo groups in our study, 25 patients would have required treatment before one would receive benefit. The decision as to the clinical value of our findings is at the discretion of the individual clinician, but given the potential for serious complications as shown by the two gastro-intestinal complications in our study, poor patient tolerance and poor compliance, the large number of patients needed to be treated before obtaining any benefit does not justify the routine use of indometacin as prophylaxis for HO.

Patient compliance can affect the success of any drug treatment regimen. To assess this in our series serum levels of indometacin were obtained in 59% of patients who received the drug, but levels were detected in only 57% of these. Similar numbers of patients lost their medication in each group, but significantly more withdrew from the indometacin group because of side effects, indicating that the drug was poorly tolerated by many patients and that compliance with this regime was difficult. Complications with indometacin have been reported to include gastric ulcers, decreased platelet function, renal toxicity and impaired healing of a fracture. A limitation of our study was the use of an intention-to-treat model for the analysis of the data. This required the inclusion of all patients in the groups to which they had been randomly assigned, regardless of their adherence with the entry criteria, the treatment which they actually received or their subsequent withdrawal or deviation from the protocol. We used this method since it was unclear from the literature exactly how long treatment with indometacin was required for to provide an effect, with between one and six weeks suggested. Given this uncertainty, we decided to analyse the patients in the groups to which they had been originally assigned since this represented the clinical situation most accurately.

A second criticism related to the power of this study. The inability to detect a difference between the indometacin and placebo groups with respect to the severity of HO could be attributed to the size of the groups. However, if we assume that the effect sizes obtained represented the true effect, we would require approximately 278 subjects in each group to be able to conclude that our observed differences were significant at a 5% level with 80% power based on a two-group three-category chi-squared test. This indicated that the effect size in our study was quite small and perhaps clinically insignificant. Given the complexity of sampling, the complications and the difficulties with compliance, we believe that continuing the study to obtain a sample size four times larger was unrealistic and unlikely to add information pertinent to the clinician.

This study was unable to demonstrate a statistically significant reduction in the incidence of severe HO with the use of indometacin compared with a placebo. More patients withdrew from the indometacin group because of side-effects compared with the placebo group. Patient compli-
pliance with indomethacin was poor, and serious gastrointestinal complications occurred in two patients treated with this drug. Therefore, we cannot recommend the routine use of indomethacin for prophylaxis against HO in isolated fractures of the acetabulum exposed through a Kocher-Langenbeck approach.

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


