A 2-year follow-up study was conducted on 46 patients to assess Harris Hip Scores (HHS) and early distal migration of the femoral stem after total hip arthroplasty using a fully cemented femoral stem. The mean age was 73.4 years. The Fullfix stem (Mathys A.G., Bettlach, Switzerland) is characterised by a mat polished surface and a flange aimed to compress the cement during the insertion process. Early distal migration was determined in 36 patients using the computer assisted EBRA-FCA method (Einzel-Bild-Roentgen-Analyse femoral component analysis). At 2 years, subsidence reached a mean value of 0.44 mm (95% CI: 0.19, 0.70), whereafter the stem appeared stable, i.e. distal migration reached a plateau. Females showed significantly higher migration than males (p < 0.01). EBRA-FCA proved to be an accurate method to measure early migration, with a standard deviation below 1 mm (0.56 mm in the interval 1-2 years). Harris Hip Scores (HHS) at 2 years follow-up averaged 89.6 (95% CI: 79.5, 99.8). As expected considering the small extent of migration, there was no correlation between subsidence and HHS (p = 0.5).

Key words: hip ; arthroplasty ; femoral stem ; cemented ; migration.

INTRODUCTION

Total hip arthroplasty is currently an established treatment for osteoarthritis of the hip as well as for hip fractures in elderly patients (31).
It appears that too rough a surface favours cement abrasion which may lead to early loosening of the shaft (10, 34, 35, 39, 41). Whether a highly polished or rough stem surface affects subsidence and generation of cement particles remains controversial. Many studies consider stem migration in the distal direction, i.e. subsidence, as a predictive factor for early loosening (7, 13, 22, 23, 27, 43), the most prominent migration generally being observed during the first year (1, 25, 26). Migration of more than 2 mm in the first 2 years appears to be associated with early loosening (27, 29, 30, 43). For this reason, more and more authors propose that the dynamics of migration of new or altered designs of prostheses and implant shafts should be examined first on a limited number of patients prior to implantation on a large scale (7, 13, 23, 27, 43). EBRA-FCA, which uses repeated images of a single radiography, offers a non-invasive method to analyse migration of implanted hip cups and femoral stems (3, 29, 30).

The Fullfix® femoral prosthesis was developed by Mathys AG, Bettlach, Switzerland in 1995/1996 and was first used in 1997. The design characteristic of this stainless steel stem is a so-called flange, expected not only to provide auto-centering of the stem but also to compress the cement during the implantation process in such a way that it acts against subsidence of the implant (1, 11, 12, 20, 24, 32, 42, 44). The close contact to bone in the medial proximal region aims at optimal transmission of forces, precise alignment, close fitting of the implant and hence strong compression of bone cement (24, 32, 38). The specific cross-section of the lateral part is meant to provide high rotational stability, whereby shear forces are expected to be confined to unloaded zones. Furthermore, the mat polished surface should favour adhesion of the cement, hence reduce the risk of cement abrasion. A retrospective study was carried out to investigate the effect of these design features on the migration behaviour of the stem.

The aim of this study was to quantify the migration of the implant. In addition, the functional outcome of the hip joint was assessed using the Harris Hip Score. Finally, the relationship between the amount of migration and the Harris Hip Score was explored.

**PATIENTS AND METHODS**

The present study is a non-randomised one-center retrospective follow-up study based on patients recruited in a medium-size hospital in Belgium. All patients were treated by the same senior surgeon with a long experience in total hip arthroplasty. A fully cemented Fullfix® stem (Mathys A.G., Bettlach, Switzerland) was implanted in all patients (fig 1). A cementless titanium-coated RM polyethylene cup was usually implanted in younger patients, and a cemented polyethylene cup in the older category.

Follow-up clinical and radiological examinations at periods 1, 3, 6, 12 and 24 months after surgery were reviewed retrospectively, as well as the base-line evaluation immediately prior to surgery. The median follow-up for the radiological assessment of migration was 25 months; the inter-quartile range (25% and 75% percentiles) was 14-37 months.
Inclusion / exclusion criteria and lost to follow-up

Patients were included if they had osteoarthritis, a traumatic fracture or inflammatory arthritis. There were no cases with rheumatoid arthritis, although this was not an exclusion criterion.

 Determination of subsidence

The EBRA-FCA method based on repeated images of a single radiography (Einbild Röntgenanalyse) was used to assess migration. One of the main features of this technique is to reduce the inherent measurement error by estimating the mean profile based on subsequent radiological measurements using well-defined reference lines. Eventual stem migration over time usually follows a curvilinear shape, i.e. after initial subsidence it will reach a persisting plateau at 2 or 3 years after operation. The more measurements points are available over time, the more precise will be the estimation of subsidence.

The measurement accuracy is expected to be 1.5 mm, i.e. the true value is known with an implicit lack of precision of 1.5 mm, i.e. twice the standard deviation or the 95% interval. As compared to roentgen stereophotogrammetric analysis (RSA), generally accepted as the gold standard, the 95% percentiles of differences between EBRA and RSA were ± 1.6 mm. This is better than the very similar “Dusseldorf Migration Analysis - Femoral Component Analysis” (DMA-FCA), with which the 95% percentiles of differences with RSA amount to ± 2.5 mm. In terms of reliability an inter-observer coefficient of 0.84 was noted slightly lower than the coefficient of 0.89 reported with DMA-FCA.

All images were processed by the same technician trained to work with digitised radiographic images.

There were several observations with apparent upward migration, i.e. showing positive migration values. As the factors responsible for these measurement errors were expected to act symmetrically, i.e. random in both directions, these values were not altered.

To assess the reliability of determining the reference points by the user, a sample of n = 20 images were evaluated by two persons independently. The intra class correlation (ICC) coefficient according to Bland and Altman’s method was calculated for this purpose. The ICC between the two ratings was 0.99. Hence there is no observer bias when marking the reference lines.

Assessment of Harris Hip Score

The Harris Hip Score is a questionnaire based assessment of the functional outcome after total hip arthroplasty. It basically comprises a quantitative measure for pain, range of movement and a basic assessment of the daily activities such as climbing stairs. The total score yields a fairly good functional indication for the clinical outcome.

Statistical methods

For a description of the Harris-Hip Score over time, Box plots giving the 25%, 50% and 75%-percentiles, respectively, were used. Likewise, the distribution of subsidence was analysed. To allow for a tabular analysis the time-points at examination were classified into the following non-overlapping categories: 0, 1, 2-3, 4-6, 7-12, 13-24, 25-36 and > 37 months.

To determine the monthly migration rate a linear random effects regression model of subsidence against time since operation was examined. To account for the curvilinear trend, both a linear and a quadratic term of time were specified. The SAS procedure ‘Mixed’ was used for this purpose with patients treated as block variable, i.e. a compound symmetry covariance structure was specified to handle the correlation among patients. Quantile-quantile-plots (i.e. QQ-plots) and graphs plotting residuals against predicted values (also known as Tukey-Anscombe-plot) were used to judge the assumptions of the model. Finally, to explore the relationship between the Harris Hip score and migration, the records at the longest interval since operation were used as outcome and correlated among themselves.

All statistical analyses were carried out using SAS version 9.1 (SAS Institute, Cary NC, USA).

RESULTS

Basic patient parameters and follow-up characteristics

Follow-up examinations

Fifty-six patients were operated during the time period between 30.07.1997 and 8.11.2000. The median date of operation was 9.7.1998. Seven patients were excluded as they did not provide at least one follow-up examination after the operation: 6 males (1 with bilateral surgery) and 1 female, the mean age was 72 years.
From the remaining 47 hips, 46 had a unilateral arthroplasty whereas one female patient had bilateral surgery, the operations being 2 months apart.

One patient died (F21) 7 months after operation from causes unrelated to the arthroplasty. This patient had been examined at 4 follow-up visits without any complications due to the surgery. Mainly due to the restrictions imposed by the EBRA-system (at least 4 sequential good x-ray images need to be at disposal), only 34 cases (≈ 59%), i.e. 34 patients out of the 56 operated cases were available for an assessment of the migration.

Osteoarthritis (35 out of 46 patients, i.e. 76%) was the main indication for surgery, followed by fractures (15%). There was one patient with inflammatory arthritis and two patients with avascular necrosis.

The median duration of the follow-up for the assessment of the Harris Hip Score was 10 months (25% and 75% Percentiles : 6-23 months) and 9 months after exclusion of 3 cases (1 due to dislocation and 2 due to wrong shaft sizing). The maximal follow-up time for the HHS data was 3 years (37 months).

The median duration of the follow-up for the determination of subsidence was substantially longer : the median duration was 25 months (25% and 75% Percentiles : 14-37 months). The maximal follow-up time for the migration data was 49 months.

The mean age of the patients (n = 46) was 73.4 years (Table I), the fraction of females and males was 29 (63%) and 17 (37%). Table I gives a detailed description of the follow-up examinations. It can be seen that a few patients only provided a complete follow-up history.

All patients sample and analysis set

During the analysis it turned out that two female patients (age : 77 and 80 yrs, respectively, no. 2 and 3 in the case series) had received undersized implants ; both were in a severe varus position. For these reasons, a supplementary analysis was carried out excluding these patients from the full data set. One hip dislocated 14 days following surgery in another patient (F24, male of 64 yrs.). Since this could possibly mislead the subsequent migration assessment, this patient was excluded as well. In the regression analysis shown later, these observations were clearly conspicuous as lying off the expected line in the qq-plot. Since two of these patients were at the very beginning (No. 2 and 3) of the operation series and had apparently both an undersized prosthesis, there are objective reasons as to repeat the analysis excluding these patients as the operation technique was not yet fully standardised. Luxation in turn (Patient F24) is quite a
known complication but occurred repeatedly in this patient and can therefore produce very aberrant results as revealed by the residual analysis (F24 showed a migration of 2.3 mm at 1 month and 2.6 mm at 9 months, respectively). In summary, excluding these patients would probably induce much more bias than preserving them.

**Migration**

Table II a shows a summary of the migration assessment. Likewise, table IIb gives an account of the results excluding the mentioned three patients.

Highest median migration across all follow-up examinations was 0.61 mm at 13-24 months and 0.47 mm, respectively, excluding the three patients. Upon excluding the three patients, maximum average subsidence was 0.44 mm at the period of 13-24 months (cf. table IIb). With regard to the range of the 95% confidence intervals it appears that the values are within 1 mm across all time-points. The regression analysis did reveal a statistically significant increase of subsidence across time (i.e. decrease as related to the cranial-distal y-direction). Both the linear term and the quadratic term were significant: $p = 0.0004$ and $p < 0.011$, respectively. The curvilinear term indicates, however, that the increase appears to be stable after approximately 2 years. A graphical summary of the regression analysis showing estimated least squares means including their 95% confidence limits is given in figure 3 (unlike the linearly parameterised model, the model used for the graph was based on nominally parameterised time-intervals where points were fitted by a polynomial line).

Regarding patient-specific features, the following effects were examined: gender, BMI and age. Apparently, females had higher migration values than males: delta = 0.32 mm ($p < 0.01$). BMI showed a minor tendency only, towards increasing subsidence, i.e. an inverse relationship: $p < 0.125$. 

---

**Table IIa.** Summary of data of migration assessment (Full patient’s information). (Median, Minima, Maxima, Mean, 95% Confidence Interval and Standard deviation). Values < 0 indicate apparent upward migration

<table>
<thead>
<tr>
<th>Follow-up examination</th>
<th>No hips</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Lower 95% CL for Mean</th>
<th>Upper 95% CL for Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>31</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>–</td>
<td>–</td>
<td>0.00</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>0.07</td>
<td>-0.81</td>
<td>2.25</td>
<td>0.18</td>
<td>-0.10</td>
<td>0.46</td>
<td>0.58</td>
</tr>
<tr>
<td>2-3</td>
<td>25</td>
<td>0.03</td>
<td>-0.94</td>
<td>2.26</td>
<td>0.14</td>
<td>-0.10</td>
<td>0.37</td>
<td>0.58</td>
</tr>
<tr>
<td>4-6</td>
<td>20</td>
<td>0.29</td>
<td>-1.18</td>
<td>1.73</td>
<td>0.21</td>
<td>-0.07</td>
<td>0.48</td>
<td>0.59</td>
</tr>
<tr>
<td>7-12</td>
<td>27</td>
<td>0.42</td>
<td>-0.70</td>
<td>3.43</td>
<td>0.52</td>
<td>0.18</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>13-24</td>
<td>25</td>
<td>0.61</td>
<td>-0.77</td>
<td>7.40</td>
<td>0.97</td>
<td>0.31</td>
<td>1.60</td>
<td>1.61</td>
</tr>
<tr>
<td>25-36</td>
<td>17</td>
<td>0.45</td>
<td>-1.35</td>
<td>4.93</td>
<td>0.52</td>
<td>-1.20</td>
<td>0.17</td>
<td>1.35</td>
</tr>
<tr>
<td>+37</td>
<td>15</td>
<td>0.61</td>
<td>-1.19</td>
<td>12.67</td>
<td>3.00</td>
<td>-0.22</td>
<td>0.57</td>
<td>4.99</td>
</tr>
</tbody>
</table>

**Table IIb.** Complementary analysis set to Table II a, excluding patients : F02, F03 and F24

<table>
<thead>
<tr>
<th>Follow-up examination</th>
<th>No hips</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Lower 95% CL for Mean</th>
<th>Upper 95% CL for Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>0.05</td>
<td>-0.81</td>
<td>0.51</td>
<td>0.06</td>
<td>-0.09</td>
<td>0.21</td>
<td>0.31</td>
</tr>
<tr>
<td>2-3</td>
<td>24</td>
<td>0.03</td>
<td>-0.94</td>
<td>0.75</td>
<td>0.05</td>
<td>-1.1</td>
<td>0.21</td>
<td>0.38</td>
</tr>
<tr>
<td>4-6</td>
<td>20</td>
<td>0.29</td>
<td>-1.18</td>
<td>1.73</td>
<td>0.21</td>
<td>-0.07</td>
<td>0.48</td>
<td>0.59</td>
</tr>
<tr>
<td>7-12</td>
<td>23</td>
<td>0.34</td>
<td>-0.70</td>
<td>0.94</td>
<td>0.25</td>
<td>0.06</td>
<td>0.43</td>
<td>0.42</td>
</tr>
<tr>
<td>13-24</td>
<td>21</td>
<td>0.47</td>
<td>-0.77</td>
<td>1.52</td>
<td>0.44</td>
<td>0.19</td>
<td>1.70</td>
<td>0.56</td>
</tr>
<tr>
<td>25-36</td>
<td>16</td>
<td>0.39</td>
<td>-1.35</td>
<td>1.53</td>
<td>0.25</td>
<td>-0.15</td>
<td>0.64</td>
<td>0.75</td>
</tr>
<tr>
<td>+37</td>
<td>11</td>
<td>0.26</td>
<td>-1.19</td>
<td>1.32</td>
<td>0.14</td>
<td>-0.45</td>
<td>0.73</td>
<td>0.87</td>
</tr>
</tbody>
</table>
All these effects were observed based on the restricted data set omitting the three mentioned cases. Whereas the two patients with undersized stems showed migration greater than 10 mm, the patient with luxation had subsidence amounting to 2.6 mm. Using the full data set no significant effects at all were observed. Moreover, the qq-plot displayed a skewed non-normal pattern including these observations.

**Functional assessment using Harris Hip Score**

Figure 4 and table III show the distribution of the Harris Hip Score at baseline and at the various follow-up examinations.

From an initial mean value of 30.5 the Harris hip scores (HHS) steadily rise within the very first month. Thereafter, the recovery is more smoothed and seems to reach its plateau at 24 months post-operatively. Mean value and its 95% confidence limits at 24 months are: 89.6 (range: from 79.5 to 99.8).

HHS correlated inversely with age at operation, Pearson’s $r = -0.3$, $p = 0.031$; no gender specific difference was apparent, Anova, $F_{1,52} = 0.05, p < F = 0.8$. The extreme values shown in figure 4 might be partially due to age, as 3 out of the 5 persons showing HHS less than 60 points were of age 75 years or higher.

**Correlation with subsidence**

In order to test the predictive potential of migration we checked whether high subsidence values were related to lower Harris Hip scores. Due to missing values this analysis could be done for 26 patients (after exclusion).

This analysis revealed no substantial correlation, with $p = 0.5$.

**DISCUSSION**

The present data gives evidence of fairly low migration for the fully cemented stem studied,
ASSESSMENT OF EARLY MIGRATION AND CLINICAL EVALUATION OF A CEMENTED FEMORAL STEM

...during the first three years following surgery. The time profile of subsidence is consistent with pattern D shown in an earlier EBRA study by Krismer et al (29): the stem appears to be stable after 3 years, after an initial subsidence of fairly small extent. At 3 years follow-up or later, the migration reached a median value of 0.61 mm (Inter-quartile range: from -0.44 to 8.90). However, when excluding two cases with inadequate implant size and one which presented an early dislocation, the median value for stem migration at 3 years or greater was estimated as 0.26 mm (Inter-quartile range: from 0.67 to 1.02). The mean value at 2 years was 0.25 mm (95% CI: -0.15, 0.64). Compared with other studies, these findings are rather good. Krismer et al reported a median value of 0.3 mm for migration at years 3-5 (Inter-quartile range: from 0.6 to 0.0) (29). Other studies report much more pronounced migration such as a mean total of 2.18 mm at > 5 years for the CPT stem (44). In a combined group using Charnley and Stanmore stems, a total migration of 1.93 mm was observed at the 2-year follow-up period (43).

Regarding the accuracy of the EBRA measurements, the results prove reliable, since a curvilinear time-profile, typically observed in most studies during the first two years following surgery, was apparent. The standard deviation for the subsidence values was 0.75 mm at 2-3 years and 0.87 mm at +3 years, respectively. The 5% and 95% percentiles at 2-3 years were at -1.53 and 1.35, respectively, consistent with an estimated accuracy of +/-1.5 mm reported by Biedermann et al (3). Given our results, a subsidence greater than 0.5 mm can well be detected with a sample of n = 20 patients with a power of 90% (SD = 0.65, a = 0.05, 2-sided).

Since only anteroposterior radiographs were studied, migration in rotation could not be assessed. However, as hypothesised earlier, strong rotational stability may help to decrease subsidence (25).

Regarding the Harris Hip score the final values reached after 2 years (i.e. average score of 89.6) are in-line with other short-term follow-up studies. Several studies report an average HHS in the range of 85.9-92 at follow-up times of 5.3 up to 10 years (4, 8, 9, 37).

There was no learning curve regarding the calendar time of operation. That is, patients treated early did not perform worse as compared to those from later periods of the study. However, it is noteworthy that there appears to be considerable variability in the pre-operative examination, reflecting the heterogeneity of the patients.

Although the follow-up time is short, it is nevertheless of interest to note that no clinical or radiological failure has been observed.

Finally, no correlation was found between the extent of migration and the corresponding Harris Hip Score. This might be due to both low migration and the combined effect of a relatively small sample size and a rather short follow-up period, restricting the statistical power for detecting such an association.

There remains the issue whether complete absence of subsidence is a necessary condition to reduce the risks of aseptic loosening on the long term. Among the group of fully cemented shafts...
basically two distinct directions can be discerned. On the one side the ‘force closed stems’ are highly polished stems allowing for subsidence to some extent. Several studies demonstrate that such subsidence occurs over time but does not necessarily end in a failure. For instance, stems of this design such as the Exeter stem have rather good long-term results as shown in the Swedish registry. On the other side the ‘shape closed stems’ are designed to counteract subsidence by means of a collar or a rough polish surface. The Lubinus SP II shaft with these characteristics also shows quite good long-term results in the Swedish registry.

Up to now superiority of one approach against the other could not be established. Hence, it would be unjustified to judge a fully cemented shaft based only upon measuring migration, or making a statement about the predictive value of subsidence with respect to aseptic loosening without taking into account the specific design features. The Fullfix stem clearly represents the paradigm of achieving no subsidence at all. Based on this study it could be shown that the design of this stem achieves this claim of no subsidence. However, stems that achieve initially good results will not necessarily show good long-term results as well. For instance, the pre-coated Harris shaft provided with a collar and a rough surface showed a marked failure rate on the long term. It is noteworthy that the surface roughness of this shaft is about five times higher as compared to the Fullfix, i.e. 3.42 µm vs. 0.7 µm (highly polished stems like the Exeter have values around 0.05 µm). This is further consistent with the excellent performance of the cemented Müller straight stem (rugosity 0.7-0.8 µm) as reported in the Scandinavian registry. The low roughness hence might prevent stress on
the cement interface through lower friction forces. At least the present 2-year results show that a ‘shape closed stem’ with a mat polish surface can prevent subsidence, probably due to the markedly strong collar and the highly pressurised cement mantle.

Whether the position of the Fullfix stem remains stable over an extended period of time needs to be demonstrated in long-term studies. In addition, the association between the rate of migration and the risk of aseptic loosening needs to be examined on a longer time scale.

Acknowledgments

The statistical analysis was financed by the manufacturer Mathys Bettlach AG.

REFERENCES


