Cementless total hip arthroplasty (THA) was developed in the 1980s. Early problems with cementless femoral components included failure of bone ingrowth [1], [2], loosening, and thigh pain [3]. It was determined that initial fit and stability of the implant in the femur were found to be critical for success. A robotic system, ROBODOC (THINK Surgical, Inc., Fremont, CA), was developed to address these concerns. The system uses computed tomographic (CT) scans for preoperative planning combined with an active robot [4], which precisely mills the cavity in the femur during the procedure. In addition, this technique also offers the ability to reconstruct the position of the center of the femoral head in 3 dimensions, thereby restoring the leg-length, offset, and anteversion to within a few millimeters of the native femoral head. The ROBODOC system has been used worldwide (Austria, France, Germany, India, Japan, Korea, Switzerland, and the United States) for over 17,000 THAs and over 15,000 total knee arthroplasties (TKAs) since 1994 [5–14]. It received US Food and Drug Administration (FDA) clearance for THA in 2008, and it is currently actively being used in Korea and Singapore for TKA.

Patients from 2 prospective randomized US FDA clinical trials (1994–1998 and 2001–2006) underwent THA using either an active robotic system or a traditional manual technique. These 2 studies were used to determine if any differences existed in radiographic analysis and patient pain and function using the University of California, Los Angeles; visual analog scale; Health Status Questionnaire (HSQ) pain; HSQ role physical; HSQ physical functioning; Harris pain scores; and the total Western Ontario and McMaster Universities Osteoarthritis Index scores at a mean follow-up of 14 years.

Results: The ROBODOC group had statistically significant higher HSQ pain and Harris pain scores and lower Western Ontario and McMaster Universities Osteoarthritis Index scores. There was no statistically significant difference in probability of a revision for wear between the groups ($\chi^2 = 1.80; P = .179$), and no revisions for loosening in either group.

Conclusion: Prior studies have demonstrated improved implant fit and alignment with the use of this active robot system. This long-term study now shows no failures for stem loosening at a mean follow-up of 14 years and small but potentially important improvements in clinical outcomes in the robot group.

© 2017 Elsevier Inc. All rights reserved.
hope, however, was that the patients who had undergone THA using the ROBODOC system would have better long-term clinical or radiographic outcomes. Follow-up of the patients in these 2 clinical trials offer the opportunity to examine these questions. Hence, the aims of this study are to determine whether active robotic THA improves clinical outcomes and results in fewer revisions over a long-term follow-up.

Materials and Methods

Study Design and Setting

This study presents the long-term outcomes of 2 prospective randomized FDA clinical trials. The first randomized multicenter clinical trial of 136 hips in 119 patients took place from 1994-1998 to determine the safety and efficacy of using ROBODOC for cementless primary hip arthroplasty [16]. A second randomized controlled multicenter FDA clinical trial of 97 hips in 97 patients was performed from 2001-2006.

Participants/Study Subjects

The participants considered for this study came from patients enrolled in the 2 clinical trials at a single site and included 125 THAs performed by a single surgeon on 118 patients. The patients in the first study were randomized 1:1 (robot-to-control) and included 21 THAs in the robot group and 20 THAs in the control group, whereas the patients in the second study were randomized 3:1 (robot-to-control) and included 65 THAs in the robot group and 19 THAs in the control group.

Description of Experiment, Treatment, or Surgery

The ROBODOC system consists of ORTHODOC, a preoperative planning software, and ROBODOC, a robotic arm with a high-speed burr as an end-effector. The planning software allows viewing of a CT scan of the patient’s hip and knee in 3 orthogonal views. Once the scan is complete, the CT data is used as input into ORTHODOC, which combines the individual slices to produce a series of 2D images for templating purposes. A 3D surface model of the operative bone is also created in ORTHODOC for intraoperative registration purposes. ORTHODOC contains a library of FDA 510(k) cleared hip replacement implants, depending on the application installed. The surgeon selects an implant model from this library and manipulates the 3D representation of the implant in relation to the 3D bone model to optimally place the implant. Once the surgeon is satisfied with the implant location, the data are written to a transfer media file for use with the ROBODOC during surgery. For the first FDA trial, the ROBODOC system required 3 fiduciary markers to be placed in the femur at a separate outpatient surgery using local anesthetic: 2 at the knee just anterior to the epicondyles and 1 in the greater trochanter. For the hip arthroplasty, the patient was positioned in the lateral decubitus position and the hip exposed using a standard posterior approach. The hip was dislocated and the lower leg placed in a holder with the hip flexed approximately 45°, internally rotated approximately 45° and slightly adducted. The femur was then rigidly fixed to the robot using a clamp that was placed at the level of the lesser trochanter. A bone motion monitor was attached to the bone to detect any motion of the bone during registration and milling and stop the robot if any motion was detected. Registration was performed by hand-guiding the robot to the fiduciary markers. The robot would then approach the marker from several orientations to determine center of the markers. Milling of the femoral canal was then performed by the active robot under direct supervision of the surgeon. The surgeon could pause the milling at any time as desired, but the robot performed the milling according to programmed cut paths. The surgeon could not alter the plan intraoperatively. If the surgeon did not feel that the milling was correct, the only option was to abort the procedure and complete the procedure manually. This did not occur in this study. Once milling was complete, the robot was detached from the femur and wheeled away. The femoral neck osteotomy was performed manually with a saw at the level of a notch that the robot had created. A thrombin-soaked lint free sponge was then placed down the femoral canal to control bleeding. The acetabulum was exposed and prepared in a routine manner using conventional hand-held reamers. The acetabular component was then impacted into place and a trial liner inserted. The sponge was then removed from femur and the preoperatively selected stem was impacted into place. A trial head was placed on the stem and a trial reduction performed in the usual manner. After selecting the appropriate acetabular liner and femoral head, these were implanted and the hip closed in a routine manner. A posterior capsular repair was performed in every case.

Between the first and second FDA clinical trials, the ROBODOC software was improved such that the separate surgery with implantation of fiduciary markers was eliminated and replaced with a surface matching registration technique performed at the time of the hip surgery. Furthermore, the robot cut times were shortened and a more robust error recovery system was used. The purpose of the second study was to establish substantial equivalance to the results of the first study. The second study had only a 6-month follow-up and an uneven randomization of 3:1 (robot-to-control). The remainder of the technique was identical to that described above. The control groups for both studies used conventional preoperative planning with acetate templates and plain film radiographs with magnification markers. The goals of the preoperative plan were the same in each group and the same surgeon performed the planning in both groups. The surgery was performed using conventional manual techniques for THA. With the exception of the planning and execution of femoral bone preparation by the robot, the 2 groups were treated in an identical manner. Patients were implanted with the DePuy AML or Howmedica Osteolock stems in the first clinical trial, and with Zimmer VerSys FMT or Howmedica Osteolock stems in the second clinical trial.

Variables, Outcome Measures, Data Sources, and Bias

Outcomes measured included survivorship and radiographic analysis along with patient pain and function measured using the University of California, Los Angeles (UCLA) Physical Activity Scale; visual analog scale pain; Health Status Questionnaire (HSQ) pain; HSQ role physical; HSQ physical functioning; Harris pain and total scores; and the total Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). All outcomes were measured at a mean follow-up of 13.8 ± 3.2 years for the robot group and 14.2 ± 4.7 years for the control group. Baseline outcome measures were not available for this study.

Statistical Analysis

The 2 studies combined included 125 THAs performed on 118 patients by a single surgeon at a single site. Of these 118 patients, 23 have been confirmed to have died, 18 have been lost to follow-up (15%), 7 declined to participate due to infirmity, and 9 subjects failed to return questionnaires after being consented. Data were available for 15 patients from the first clinical trial (39% lost to follow-up) and 46 patients from the second clinical trial (20% lost to follow-up). Of the 61 patients, 40 (45 hips) belonged to the robot
group and 21 (22 hips) belonged to the control group. The demographics were similar between the 2 groups (Table 1), with the exception that there were more men than women in the robot group. The Charlson comorbidity index was computed for each subject. Differences between the percent of cases in the robot vs control groups on categorical variables were compared using the $\chi^2$ test of independence. Differences in means between the groups were compared using analysis of variance or the Kruskal-Wallis test of independence. Differences in means between the groups were compared using the Kruskal-Wallis analysis of ranks depending on whether the data met the assumptions of parametric statistics. The Kaplan-Meier survival analysis and the log-rank test were used to compare the differences in time to revision between the 2 groups.

Results

The robot group had statistically significant higher HSQ pain ($P = .019$) and Harris pain ($P = .025$) scores and lower WOMAC ($P = .034$) scores than the control group (Table 2; Fig. 1). Although no statistically significant difference was found in visual analog scale pain, HSQ physical and total, total Harris, and UCLA activity scores, superior trends were seen in all these scores for the robot group.

No statistical difference was found in survivorship between the groups with revision for any reason as an end point. There was one revision of the femoral component in each group for postoperative periprosthetic fracture at 2 years for the robot group (not related to robotic procedure) and 3 years for the control group, respectively. There were 3 reoperations in the robot group and 5 in the control group, all for head and liner change due to polyethylene wear. No dislocations occurred after the index surgery in either group. One patient in the robot group suffered a dislocation after reoperation for head and liner change. There were no infections in either group.

The radiographs found that no components in either group showed signs of radiographic loosening. Both groups had similar results for osteolysis (4% robot and 11% control). Femoral osteolysis was limited to the proximal Gruen zones in both groups (4% robot and 6% control). Stress remodeling was seen in both groups receiving the Osteolock implant.

Discussion

The first FDA study (15) established the accuracy and precision of the robot with superior radiographic results (fit, fill, and position) compared with the manual broaching technique. However, the clinical results (HHS, complications, and revision rate) showed no significant difference. There were no robot-related complications, but the surgical time and blood loss were significantly greater in the robot group. The improvements in software, registration, cut times, and error recovery substantially reduced the time and blood loss. The second FDA study was required to establish substantial equivalence. The system received FDA clearance in 2008 and has been used extensively outside the US since the early 2000s.

Although there are several published clinical studies [5–9] using ROBODOC with short-term to midterm follow-up, this is the first long-term follow-up study of patients in a randomized clinical trial. In this study using more modern outcome measures, 3 of the clinical outcome measures (HHS pain, HSQ pain, and WOMAC) showed statistically superior results for the robot group at a mean follow-up of 14 years. These differences are less than the published minimal clinically important differences (MCIDs) for WOMAC [17,18] and HHS [19], but this is to be expected. Published studies of MCIDs are based on preoperative to postoperative changes after THA. Before THA, patient’s quality of life is generally poor and a successful THA can provide dramatic improvement. This is reflected in the large reported MCIDs on these measures before vs after the surgery [17–19]. Preoperative outcome scores were not available for this study and the scores were compared between the 2 groups 14 years after surgery. Differences in quality of life between patients who had a THA are not expected to be as dramatically different as they would be in patients before vs after a THA, so MCIDs as currently calculated cannot be properly applied to this study [20]. Ceiling effects on WOMAC score [21], Harris hip pain and total scores [22], may have also hidden small but possibly important differences when comparing postoperative groups. Therefore, we are reporting statistically significant improved scores in the robot group although these differences are small.

There were no differences in revision rate between the 2 groups. However, there was a trend ($P = .073$) toward a higher revision rate for wear in the conventional group, but the sample sizes were small enough that this was not statistically significant.

The reasons for the improved outcomes found in the robot group are likely multifactorial. It is the opinion of the authors that the differences are due to improved implant sizing and positioning as well as better restoration of femoral offset, anteversion, and leg length, which have been shown in prior studies [23–35]. In addition, the milling process is less traumatic and more accurate than manual broaching, thereby achieving more intimate bone contact with the implant [36].

### Table 1

Demographics of Robot and Control Groups.

<table>
<thead>
<tr>
<th></th>
<th>Control (n = 22)</th>
<th>Robot (n = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>n (%) Mean (SD)</td>
<td>n (%) Mean (SD)</td>
</tr>
<tr>
<td>Male</td>
<td>12 (54.5%) 29.4 (6.5)</td>
<td>35 (77.8%) 32.0 (5.2)</td>
</tr>
<tr>
<td>Female</td>
<td>10 (45.5%) 35.2 (8.4)</td>
<td>10 (22.2%) 32.0 (5.2)</td>
</tr>
<tr>
<td>Index age, y</td>
<td>59.8 (9.4) 59.1 (8.2)</td>
<td>59.8 (9.4) 59.1 (8.2)</td>
</tr>
<tr>
<td>Age at follow-up, y</td>
<td>74.2 (8.2) 72.8 (7.6)</td>
<td>74.2 (8.2) 72.8 (7.6)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>29.4 (6.5) 28.3 (4.7)</td>
<td>29.4 (6.5) 28.3 (4.7)</td>
</tr>
<tr>
<td>Follow-up, y</td>
<td>14.2 (4.7) 13.9 (2.7)</td>
<td>14.2 (4.7) 13.9 (2.7)</td>
</tr>
</tbody>
</table>

The demographics of patients included in the study. There were no statistically significant differences in index age, comorbidities, or BMI, but there were more men than women in the robot group ($\chi^2 = 7.00; P = .008$).

### Table 2

Means (SDs) of Robot and Control Groups.

<table>
<thead>
<tr>
<th>VAS pain score</th>
<th>Control (n = 22); Mean (SD)</th>
<th>Robot (n = 45); Mean (SD)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 best, 100 worst)</td>
<td>6.42 (10.89)</td>
<td>4.69 (10.15)</td>
<td>.112</td>
</tr>
<tr>
<td>(0 best, 100 worst)</td>
<td>72.65 (16.31)</td>
<td>83.75 (20.40)</td>
<td>.019b</td>
</tr>
<tr>
<td>(0 worst, 100 best)</td>
<td>70.88 (35.23)</td>
<td>81.39 (28.23)</td>
<td>.317</td>
</tr>
<tr>
<td>(0 worst, 100 best)</td>
<td>75.49 (26.43)</td>
<td>84.26 (26.71)</td>
<td>.102</td>
</tr>
<tr>
<td>(0 worst, 100 best)</td>
<td>39.09 (7.37)</td>
<td>41.81 (5.05)</td>
<td>.025b</td>
</tr>
<tr>
<td>Total Harris pain score</td>
<td>89.50 (12.03)</td>
<td>93.49 (8.77)</td>
<td>.089</td>
</tr>
<tr>
<td>(0 worst, 100 best)</td>
<td>63.71 (104.53)</td>
<td>683.52 (113.09)</td>
<td>.087</td>
</tr>
<tr>
<td>(0 worst, 100 best)</td>
<td>5.71 (1.45)</td>
<td>6.09 (1.86)</td>
<td>.417</td>
</tr>
<tr>
<td>Total WOMAC</td>
<td>11.32 (11.92)</td>
<td>8.44 (11.48)</td>
<td>.034</td>
</tr>
</tbody>
</table>

Clinical outcomes for the robot and control groups.

Higher scores on the Harris, HSQ, and UCLA and lower scores on the WOMAC and VAS indicate positive functioning and less pain.

HSS, Health Status Questionnaire; SD, standard deviation; UCLA, University of California, Los Angeles; VAS, visual analog scale; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

#### Notes

b Analysis of variance. The remaining pain and function scores were analyzed using the Kruskal-Wallis analysis of ranks.

b Statistically significant difference between the 2 groups ($P < .05$).
There are some limitations to the present study. First, like most long-term studies, there have been many patients who died or lost to follow-up. This resulted in relatively small numbers in each group. Second, this is a single-surgeon study at a single center and the results may not be generalizable. The use of the robot, however, reduces the variable of surgical experience and technique. Third, even if the use of randomized control study design minimizes the expected differences between the groups, small differences can still occur as illustrated by the difference found in the numbers of men and women in the 2 groups. We believe that these potential small differences did not impact the results.

Conclusion

In the current healthcare economic environment, it is important to show value for new technologies that are introduced. The ROBODOC technology is not “new,” at least chronologically, but it has been slow to come into clinical practice in part because of the long FDA process. The present study takes advantage of this long history by showing that robotic THA can improve clinical outcomes in patients undergoing THA at 14 years follow-up compared with manual techniques.

Acknowledgments

This work was supported by funding provided by the THINK Surgical, Inc, which participated in study design, interpretation of the data, and preparation of the manuscript.

References


