Mental Health and Outcomes in Primary Total Joint Arthroplasty

Carlos J. Lavernia, MD,* Jose C. Alcerro, MD,† Larry G. Brooks, PhD,‡ and Mark D. Rossi, PhD§

Abstract: A consecutive series of 640 total joint arthroplasty patients was interviewed before surgery and at a minimum of 2 years following surgery. Statistical analyses were conducted to examine the effect of psychological distress and other patient characteristics on outcomes (Western Ontario and McMaster Universities Osteoarthritis Index, Short Form 36, and Quality of Well-Being index). Before and after surgery, distressed subjects had significantly lower scores than nondistressed subjects for most dependent measures ($P$ range, $0.05 \leq 0.001$). All mean outcomes improved by follow-up in both groups ($P \leq 0.001$) except mental health scores of nondistressed subjects. Stepwise regression analysis found that low baseline mental health score, non-Hispanic ethnicity, and fewer years since procedure were the strongest predictors of worse Western Ontario and McMaster Universities Osteoarthritis Index scores at follow-up. Although the magnitude of improvement is similar to nondistressed subjects, distressed patients do not achieve comparable functional and psychosocial outcomes. Keywords: total knee arthroplasty, total hip arthroplasty, mental health, outcomes.

Osteoarthritis (OA) of the hip or knee is a major cause of disability in elderly Americans [1]. When conservative treatments such as rehabilitation and medications have failed, total joint arthroplasty becomes a more attractive option. Most patients have positive functional outcomes following knee or hip arthroplasty surgery in terms of substantial pain relief, improvement in functional abilities, and improvement in quality of life [2-6]. The greatest improvement in pain and physical function after hip or knee arthroplasty occurs within the first 3 to 6 months after surgery [7,8]. Improvement in health-related quality of life after surgery is also evident, including domains such as social function, mental health, and vitality [2-4,6].

However, not all patients undergoing knee or hip joint arthroplasty report improvement in function. Approximately 15% to 30% of patients who have had knee arthroplasty, and a smaller number who have had hip arthroplasty, report little to no improvement compared with the general population or are not satisfied with the results shortly after surgery [3,9,10]. Lack of improvement is usually related to ongoing pain, restricted range of motion, and unsatisfactory function. In some cases, the lack of positive outcomes and dissatisfaction with surgery can be attributed to surgical issues or complications [11,12]; however, most of the time, no technical or medical reasons can be identified that influence these negative outcomes [13,14].

Several recent articles have examined the relationship between preoperative mental health status and outcomes from total hip arthroplasty (THA) or total knee arthroplasty (TKA) [15-24]. Most of the articles have concluded that preoperative psychological distress predicts outcome measures such as pain, function, satisfaction, or quality of life. However, most of the articles have relatively small samples [19] or have follow-ups of 2 years or less [15,17,18,20-24]. Research has demonstrated that poor outcomes at 1 year frequently improve by the fifth year [19]. This suggests that studies with findings of the effect of psychological distress on outcomes in the short term may not generalize to outcomes in the long term. In the present study, we defined psychological distress as having a Short Form 36 (SF-36) mental health score of less than 52, based on data from the SF-36 manual [25]. In our study, we assess the effects of preoperative mental health status on
postoperative pain, mental health, quality of life, and functional outcomes and change scores in THA and TKA with a large sample followed over an average of more than 5 years.

**Materials and Methods**

**Subjects**
A total of 2037 primary (890 THA, 1147 TKA) total joint arthroplasty procedures were studied retrospectively. All interventions were performed by a single surgeon. Inclusion criteria were unilateral knee or hip arthroplasties with diagnosis of OA, avascular necrosis, or rheumatoid arthritis, plus a minimum 2-year follow-up assessment. Cases were excluded because of history of bilateral arthroplasty (954), incomplete follow-up (345), diagnosis (98), death (40), lost to follow-up (27), and missing SF-36 data (10). This resulted in a final sample of 563 patients. No compensation was provided for study participation. Subjects volunteered to participate in this Institutional Review Board–approved study. Sociodemographic information was obtained from questionnaires administered to patients at either their first office visit or their preoperative evaluation. Follow-up assessment was performed yearly thereafter following the intervention. Each patient was interviewed by phone or in person by a trained member of the research team. All patients’ final follow-up was used for statistical analysis.

**Psychosocial and Functional Outcome Measures**
We assessed preoperative and postoperative psychosocial and functional status of the patients on several self-report measures: SF-36 [25], Quality of Well-Being index (QWB) [26,27], and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [28]. The SF-36 is a well-validated measure of physical and emotional status commonly used in arthroplasty outcomes research [17,18,21,22]. The SF-36 has been shown to be a valid and reliable instrument to measure psychometric properties in the elderly [2-4,29], as well as in African-Americans, Hispanics, and non-Hispanic whites [30-32]. Published investigations have also shown that the instrument is both specific to the domains expected to be affected by mental disorder and sensitive to the severity of mood disturbance [33]. We categorized our patients as distressed based on an SF-36 mental health cutoff score of less than 52 because (1) the user’s manual for the SF-36 demonstrates that a score on the mental health subscale of less than 52 has been shown to be highly predictive for the presence of a psychiatric disorder [25] and (2) previous research has shown that the mental health scale is less influenced by physical functioning than is the mental component scale [34]. We also included the pain and function subscales of the SF-36 as outcome variables, as they have been shown to be the most sensitive to change in OA patients following surgery [32].

The QWB Index was developed by Kaplan et al [35] to assess general health-related quality of life that combines values for symptoms and functioning. This index has been validated for use in a variety of populations, including African Americans and Hispanics [35]. The QWB scores range from 0 (for death) to 1 (for symptom-free optimal functioning).

The WOMAC has been shown to be a valid and reliable instrument to document perceived abilities in individuals with knee and hip joint arthritis [28,36,37]. Consequently, it has become a standard for evaluating results of total hip and knee arthroplasty surgery [38]. The WOMAC is a disease-specific instrument with 3 dimensions that measure the patient’s perception of pain, joint stiffness, and physical function. It consists of 24 items (5 for pain, 2 for stiffness, and 17 for function). Point values from 0 to 4 are assigned to each response, and scores are totaled for each category. The WOMAC total score, physical function, pain, and stiffness dimensions were used as dependent measures.

Last, we analyzed range of motion data preoperatively and postoperatively for all patients to determine if there were differences in objective functional measurements between the distressed and nondistressed groups. For TKA patients, specific measurements included passive and active flexion and extension. For THA patients, measurements included flexion, abduction, adduction, internal rotation, and external rotation.

**Analyses**
$t$ Tests and $\chi^2$ analyses were run to determine if there were any differences between the distressed and nondistressed groups on relevant demographic and clinical variables. The sample was then stratified into those who had THA (n = 241) and those individuals who had TKA (n = 322). Analysis of covariance (ANCOVA) was used to compare the scores of distressed and nondistressed subjects on multiple psychosocial and functional measures at baseline and follow-up controlling for significant differences in demographic variables identified between the distressed and nondistressed groups. In addition, ANCOVA was used to determine if there were any differences between distressed and nondistressed subjects in change scores (calculated by subtracting baseline from follow-up scores) on the same set of psychosocial and functional measures. Paired-sample $t$ tests were used to assess for significant differences in scores on outcome measures before and after surgery. A further set of analyses was conducted to compare subjects by type of procedure (TKA or THA), regardless of mental health status, on functional and psychosocial measures at baseline and follow-up. Because of skewness and kurtosis, all WOMAC-related follow-up scores were transformed using a natural logarithmic transformation; however, data in tables are displayed in the original scale of measurement to...
facilitate ease of interpretation. The SF-36 social functioning and bodily pain scores at baseline were not normally distributed. As a result, analyses for those variables were handled using the Kruskal-Wallis test. Missing data were handled by casewise deletion.

Stepwise regression analysis was conducted to determine whether demographic characteristics, clinical characteristics, or baseline SF-36 subscale scores predicted WOMAC pain, stiffness, and functional outcome following either TKA or THA. Demographic variables entered into the model included age at follow-up, gender, and ethnicity. Clinical characteristics selected for the analyses were procedure type (THA or TKA) and years since procedure. Short Form 36 subscales included in the analysis were the baseline mental health score, physical functioning score, and bodily pain score. Demographic variables, clinical characteristics, and baseline SF-36 predictors were entered into separate blocks. Functional outcomes were measured using the natural logarithm of the WOMAC subscale scores at follow-up. Criterion set for entry into the model was $P < .05$.

## Results

### Demographics and Clinical Variables

Average time from surgery to the subject’s last follow-up was 5.5 years ($SD = 2.9$). Of the 563 subjects with SF-36 mental health scores, 145 (25.8%) were classified as distressed using the mental health subscale of the SF-36. Table 1 shows the demographic characteristics of patients by mental health status. Body mass index was only available for 346 participants (81 distressed; 265 nondistressed). Significant differences were found between the distressed and nondistressed groups on several demographic variables: age at procedure ($P < .05$), ethnicity ($P < .01$), and gender ($P < .001$). No significant differences were found between distressed and nondistressed patients in terms of marital status ($P < .07$), body mass index ($P = .31$), race ($P = .37$), and type of arthroplasty procedure ($P = .55$). Preoperatively, there were no differences between the distressed and nondistressed groups on any of the knee or hip range of motion data ($P range = .10-.96$). Postoperatively, distressed patients displayed less active flexion ($P < .05$); however, none of the other range of motion variables were significantly different between the distressed and nondistressed groups ($P range = .16-.99$). Correlational analyses between range of motion variables and mental health scores preoperatively and postoperatively revealed few significant correlations, with none higher than 0.16. Thus, range of motion variables were minimally related to mental health scores.

### Psychosocial and Functional Outcomes—TKA

The TKA subjects who were classified as psychologically distressed presented with significantly worse scores preoperatively for nearly all psychosocial and functional measures with the exception of the WOMAC pain scale. Postoperatively, fewer differences remained between the distressed and nondistressed groups. Psychosocial and functional measures that remained significantly different were the SF-36 social function scale ($P < .001$), SF-36 mental health scale ($P < .001$), SF-36 bodily pain scale ($P < .05$), and the QWB ($P < .05$). Analyses of change scores showed significantly greater improvement postsurgery by the distressed group on multiple measures, including the SF-36 social function scale ($P < .05$), SF-36 mental health scale ($P < .001$), WOMAC function ($P < .05$), WOMAC stiffness ($P < .05$), and WOMAC total score ($P < .05$). Within-subjects analyses revealed improvement on every outcome measured at follow-up in both the distressed and nondistressed group ($P < .001$), with the exception that the nondistressed group did not show improvements in mental health. All ANCOVA analyses controlled for demographic variables of age, ethnicity, and gender.

### Psychosocial and Functional Outcomes—THA

For THA subjects classified as distressed, preoperative scores on most psychosocial and functional measures indicated significantly worse function compared with nondistressed patients. The exceptions were the SF-36 bodily pain scale and the WOMAC stiffness scale. Postoperatively, distressed subjects continued to display lower scores on a range of measures, including the SF-

<table>
<thead>
<tr>
<th>Table 1. Demographic and Baseline Characteristics of Patients who Were Categorized as Distressed and Nondistressed Preoperatively</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Gender (%)</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Race (%)</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
</tr>
<tr>
<td>Non-Hispanic</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Marital status (%)</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Divorced</td>
</tr>
<tr>
<td>Single</td>
</tr>
<tr>
<td>Widow</td>
</tr>
<tr>
<td>Separated</td>
</tr>
<tr>
<td>Procedure (%)</td>
</tr>
<tr>
<td>THA</td>
</tr>
<tr>
<td>TKA</td>
</tr>
</tbody>
</table>

* $P < .05$.
† $P < .01$.
‡ $P < .001$.  

1278 The Journal of Arthroplasty Vol. 27 No. 7 August 2012
Outcomes by Type of Procedure (TKA vs THA)

The THA patients displayed significantly worse scores at baseline than did the TKA patients on the SF-36 physical function scale ($P < .01$), SF-36 social function scale ($P < .05$), QWB ($P < .05$), WOMAC function ($P < .01$), WOMAC pain ($P < .01$), and total WOMAC ($P < .01$). At follow-up, no differences were found between TKA and THA patients on any of the outcome measures. Analysis of change scores showed that THA patients scored significantly greater improvements on multiple measures, including the SF-36 physical function scale ($P < .01$), SF-36 social function scale ($P < .01$), WOMAC function ($P < .01$), WOMAC pain ($P < .01$), and total WOMAC ($P < .01$). All ANCOVA analyses controlled for demographic variables of age, ethnicity, and gender.

Regression Analysis

Stepwise selection of the predictor variables resulted in models that accounted for 13.1% of the variance in follow-up WOMAC function score, 10.4% of the variance in follow-up WOMAC stiffness score, and 9.3% of the variance in follow-up WOMAC pain score (all follow-up WOMAC variables natural log-transformed). Of all the variables entered into the regression models, only the type of procedure (THA or TKA) did not significantly predict any of the outcome variables. The following variables were retained in at least one of the models: gender, ethnicity, age at follow-up, years since procedure, baseline SF-36 mental health score, SF-36 bodily pain score, and baseline SF-36 physical functioning score. Some variables that were not significant in the final model were retained because they were significant predictors in an earlier block. Years since procedure, ethnicity, and baseline mental health score were the only predictors that significantly predicted all 3 dependent variables (follow-up WOMAC function, stiffness, and pain). Fewer numbers of years since procedure ($P < .001$) predicted higher WOMAC scores, as did lower baseline SF-36 mental health score (Table 2). Hispanic ethnicity significantly predicted better outcomes at follow-up.

Discussion

Total joint arthroplasty of the knee and hip has been shown to lead to dramatic improvements in OA patients’ functioning, pain levels, and quality of life. However, previous research has shown that not all patients respond equally well to surgery [39-45], and this situation holds as well for joint arthroplasty surgery. A nontrivial fraction of knee and hip arthroplasty patients, some 15% to 30%, reports not being happy following surgery [3,10]. This lack of improvement often is not due to “technical failures” or complications in surgery, and the cause of the lack of improvement remains largely unknown. Multiple studies that have examined outcomes following various types of surgery have found that psychological distress and depression are associated with poor outcomes [15,17-24,46-48]. However, most of the studies have had a relatively brief follow-up ($\leq 2$ years).

We found that 26% of the patients in our study were categorized as mentally distressed at baseline (as defined by an SF-36 mental health subscale score of less than 52), although they did not present with greater limitations on objective measures of disease (range of motion). Regardless of time, these patients had significantly worse scores on measures of well-being, pain, stiffness, function, and total WOMAC. However, when we examined the change in measures from one period to the next, there was only one significant difference (SF-36 mental health score) in the relative improvements in these measures between those who were distressed and nondistressed. Thus, our findings indicate that although distressed patients have worse baseline measures, their improvement following surgery is just as substantial as for nondistressed patients. Because of the initially low

### Table 2. Multivariate Regression Results Predicting WOMAC Subscales at Follow-Up

<table>
<thead>
<tr>
<th>Predictor</th>
<th>WOMAC Function</th>
<th>WOMAC Pain</th>
<th>WOMAC Stiffness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>$-0.034^{*}$</td>
<td>$-0.088^{*}$</td>
<td>$-0.147^{*}$</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>$0.116^{†}$</td>
<td>$0.145^{†}$</td>
<td></td>
</tr>
<tr>
<td>Age (follow-up)</td>
<td>$0.164^{‡}$</td>
<td>$0.198^{‡}$</td>
<td></td>
</tr>
<tr>
<td>Years since procedure</td>
<td>$-0.121^{‡}$</td>
<td>$-0.198^{‡}$</td>
<td></td>
</tr>
<tr>
<td>SF-36 physical functioning</td>
<td></td>
<td>$-0.088^{*}$</td>
<td></td>
</tr>
<tr>
<td>SF-36 bodily pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 mental health score</td>
<td>$-0.155^{‡}$</td>
<td>$-0.168^{‡}$</td>
<td></td>
</tr>
</tbody>
</table>

* $P < .05$, † $P < .01$, ‡ $P < .001$. 

36 mental health scale ($P < .05$), SF-36 physical function scale ($P < .01$), QWB ($P < .01$), WOMAC function ($P < .01$), WOMAC pain ($P < .01$), WOMAC stiffness ($P < .01$), and total WOMAC ($P < .01$). However, analysis of change scores showed few differences in degree of change between distressed and nondistressed subjects. Distressed patients showed significantly greater improvements in SF-36 mental health score ($P < .001$) and SF-36 social functioning score ($P < .05$). Within-subjects analyses revealed improvement on every outcome measured at follow-up in both the distressed and nondistressed group ($P < .001$), with the exception that the nondistressed group did not show improvements in mental health. All ANCOVA analyses controlled for demographic variables of age, ethnicity, and gender.
baseline scores, they still have worse outcome measures following surgery. The lone exception was the SF-36 mental health score, which showed a dramatic improvement in the distressed group while remaining essentially stable in the nondistressed group. The overall percentage of distressed patients declined from 26% at baseline to 9% at follow-up. This finding suggests that arthroplasty reduced the prevalence of psychological distress in this population, likely because of a reduction in physical dysfunction associated with joint OA. Further supporting that hypothesis, a post hoc correlational analysis found a weak, inverse correlation between the change in mental health score and the change in total WOMAC ($r = -0.26$), indicating that improvements in WOMAC score were associated with improvements in mental health score.

Regression analyses were used to predict functional outcomes from arthroplasty. As in previous studies, demographic variables (gender and age) predicted outcomes. Ethnicity also predicted outcome, but in a direction different than anticipated. Hispanic patients had lower total WOMAC scores at follow-up than did non-Hispanic patients. Time since procedure also predicted outcome, with shorter time being associated with higher WOMAC scores. Lastly, lower baseline SF-36 physical functioning and mental health scores predicted worse functional outcome at follow-up. We have previously reported that phenomenon [49]. Thus, the relationship between mental health and outcome appears to be complex, with surgery affecting mental health outcomes and preoperative mental health status predicting postoperative functional outcomes.

These results may be viewed as encouraging. In comparison with the dramatic improvement in all measures for all individuals following surgery, the final difference in outcomes between distressed and nondistressed patients is relatively small for most measures. Nonetheless, it might be important to determine whether actively treating emotional health in distressed presurgical patients will allow their outcomes to equal those of nondistressed patients.

One of the strengths of our study is its demographic composition. Our sample has a much larger percentage of Hispanic Americans (76.5%) than other published studies. This potentially enables the extension to a new demographic group of findings that mental health status predicts arthroplasty outcomes. On the other hand, the ethnic differences in functional outcome, with better outcomes for Hispanic subjects, may have biased the results and merit further research.

Future research is also needed to further explore the causal forces. If it can be shown that mental health treatment will increase the improvement in distressed patients' surgical outcomes such that they are equivalent to nondistressed patients, a brief mental health screener could be widely and easily applied to identify distressed patients before surgery. The patients identified as distressed could then receive a psychiatric consultation, and possibly treatment, both before and following surgery to attempt to improve emotional health and the risk of poorer outcomes. If preoperatively distressed patients can achieve outcomes equivalent to nondistressed patients, overall patient satisfaction might be improved and subsequent health care utilization that results from “poor” outcomes might be reduced.

Although not the main goal of this article, our results show an interesting relationship between outcomes following joint arthroplasty surgery and site of procedure. We found that the site of procedure was a significant determinant of some of our outcome measures. Previous work has found that patients undergoing THA have greater improvements in outcomes, and better ultimate outcomes, than patient receiving TKA [7]. In our study, hip arthroplasty patients entered the study having significantly worse scores on several measures of well-being and disability. However, at their last follow-up, hip patients scored no differently than knee patients on any of the measures. This equalization was due to the fact that the hip arthroplasty patients had significantly greater improvements following surgery than did the knee arthroplasty patients. This result, as well as our results for distressed patients, highlights the importance of taking into account both the absolute level of outcome measures at any point in time and also the magnitude of the change in measures between periods.

Our study has several limitations. First, our measure of psychological distress is based on the SF-36 mental health subscale. Although the SF-36 is a well-validated, reliable scale, it is not designed to diagnose specific mental health disorders. For example, the fact that the patients in this study all faced significant pain and disability with their OA may have biased upward our finding of psychological distress. For this group of patients, a category cutoff of 52 for the SF-36 mental health scale might be too high. However, several other studies have used a similar cutoff on the SF-36 mental health or mental component scales to identify psychological distress or emotional health issues in total joint arthroplasty patients [16,17,46]. A second limitation was the lack of inclusion of a variable assessing patient expectations of the procedure. Expectations might have assisted in explaining additional variance in outcomes and is an area deserving further research. A third limitation is one inherent to all studies that assess subjective measures such as well-being, social function, pain, and stiffness. Despite having used only instruments that have all been validated in related populations, subjective measurements such as these may be biased through patient report.
In summary, our study found that those individuals who were categorized as distressed had worse self-perceived, but not objective, OA disability before surgery than nondistressed individuals. Following surgery, distressed patients had rates of improvement on most outcome measures that matched nondistressed patients. However, because of the worse starting point, there were still significant differences between distressed and nondistressed patients following surgery on most self-perceived outcome measures. This result extends findings from previous studies to a time point on average 5.5 years later. The results of this study indicate that the relationship between mental health and outcome from surgery is at least bidirectional, with surgery leading to significant improvements in mental health status, whereas worse preoperative mental health status is a predictor of relatively worse outcomes from surgery. We believe that the results of this study, and the demonstrated feasibility of using a simple mental health assessment tool in a clinical setting, could lead to future research that might help improve recognition of mental health concerns in surgical patients and ultimately improve outcomes for distressed patients.

Acknowledgments

We would like to thank Mercy Hospital, Mercy Foundation, and the Arthritis Surgery Research Foundation Inc for their financial support to perform this investigation.

References

25. Ware J. SF-36 physical and mental health summary scales: user's manual; 1994.


