Results of Cervical Arthroplasty Compared with Anterior Discectomy and Fusion: Four-Year Clinical Outcomes in a Prospective, Randomized Controlled Trial

Rick C. Sasso, MD, Paul A. Anderson, MD, K. Daniel Riew, MD, and John G. Heller, MD

Background: The published two-year results of the pivotal U.S. Food and Drug Administration investigational device exemption trial with the use of the Bryan cervical disc arthroplasty compared with anterior cervical discectomy with fusion for treating single-level degenerative cervical disc disease revealed a significantly superior overall success rate in the arthroplasty group. The purpose of this study was to evaluate the midterm safety and effectiveness of the Bryan disc as an alternative to arthrodesis following anterior cervical discectomy.

Methods: A prospective, multicenter randomized clinical trial was undertaken for the treatment of persistent radiculopathy or myelopathy due to single-level cervical disc herniations or spondylosis. Patients were randomized to treatment with either the Bryan disc (the arthroplasty group; 242 patients) or anterior cervical discectomy and fusion (the fusion group; 221 patients). Patients completed preoperative and postoperative self-assessment forms at specified intervals and had radiographs made preoperatively, at six weeks, and at three, six, twelve, twenty-four, and forty-eight months after surgery. The primary outcome measure was overall success, a composite variable of safety and efficacy measures. Numerous secondary measures were assessed. The follow-up data for up to twenty-four months have been previously published. We report in the present study the forty-eight-month data collected on 181 patients who received the Bryan disc and 138 patients who underwent anterior cervical discectomy and fusion.

Results: The study groups were demographically similar. Substantial reduction in Neck Disability Index scores occurred in both groups compared with preoperative values. The greater improvement in the Neck Disability Index score in the Bryan disc cohort persisted through the four-year follow-up period (p < 0.001). The four-year overall success rates were 85.1% and 72.5% for the arthroplasty and fusion groups, respectively (p = 0.004). The improvement in the arm pain score was substantial for both groups and significantly greater in the Bryan disc cohort (p = 0.028), and the neck pain scores showed persistently greater improvement in the Bryan disc group at forty-eight months of follow-up (p = 0.001). Short Form-36 physical component score improvement remained greater among the Bryan disc cohort (p = 0.007). The mean range of motion for the Bryan disc was 8.08° and 8.48° at twenty-four and forty-eight months, respectively. Total and serious adverse event rates were similar between the groups.

Conclusions: The forty-eight-month follow-up data in the present report showed consistent, sustained significantly superior outcomes for cervical spine arthroplasty compared with cervical spine fusion. The arthroplasty cohort continued to show significantly greater improvements in Neck Disability Index, neck pain score, arm pain score, and Short Form-36 physical component score, as well as the primary outcome measure, overall success, at forty-eight months following surgery.

Level of Evidence: Therapeutic Level I. See Instructions to Authors for a complete description of levels of evidence.

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. In addition, one or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. Also, one or more of the authors has had another relationship, or has engaged in another activity, that could be perceived to influence or have the potential to influence what is written in this work. The complete Disclosures of Potential Conflicts of Interest submitted by authors are always provided with the online version of the article.
Anterior cervical discectomy and fusion is a common treatment for radiculopathy and myelopathy. The procedure is well tolerated and, in correctly selected patients, the results are highly successful. Longer follow-up has revealed that up to 25% of patients may develop recurrent radicular symptoms from adjacent segment disc degeneration. Furthermore, reoperations may be required to treat complications of fusion, such as nonunion and bone graft collapse or expulsion. In select patients, cervical arthroplasty is a potential substitute for cervical fusion after anterior neural decompression.

In the United States, a prospective, randomized controlled clinical trial was performed to evaluate the safety and effectiveness of the Bryan Cervical Disc (Medtronic Spinal and Biologics, Memphis, Tennessee) compared with fusion with allograft and plate stabilization. The indications for surgery were radiculopathy or myelopathy in patients with single-level cervical disc disease. The twenty-four-month results of this trial have been published. The overall success and other patient-reported outcome measures indicated that treatment with the Bryan cervical disc arthroplasty achieved significantly superior results and allowed patients to return to work sooner. The current investigation was prompted by the relatively short follow-up period for the implanted motion-bearing device in the prior study. Further, to assess cost-effectiveness, knowledge of the durability of improvement in self-reported outcomes compared with preoperative states is essential. The purpose of this study was to evaluate clinical functional outcomes four years after either Bryan cervical disc arthroplasty or anterior cervical disectomy and fusion, with an emphasis on assessing any deterioration in outcomes over time.

Materials and Methods

Subjects and Study Design

The study design has been reported elsewhere. Briefly, eligible patients were at least twenty-one years old and had radiculopathy or myelopathy from single-level cervical disc disease secondary to disc herniation or focal osteophytes that had not responded to at least six weeks of nonoperative management. All investigational sites had institutional review board approval, and all patients provided voluntary informed consent to participate in the study. Patients were randomly assigned in a 1:1 ratio to one of two treatment groups: arthroplasty with use of an artificial disc, the Bryan Cervical Disc (Medtronic Spinal and Biologics) (Fig. 1), or fusion with anterior cervical plate stabilization and bone allograft. The surgical procedures were performed at thirty investigational sites by sixty-five investigators and coinvestigator surgeons. The procedure in the fusion group was standardized by using both a commercially available allograft (Cornerstone; Medtronic Spinal and Biologics) and a single anterior cervical plating system (Atlantis; Medtronic Spinal and Biologics). The patients in the arthroplasty group were treated with a two-week postoperative course of a nonsteroidal anti-inflammatory drug of their surgeon’s choice. Recommendations for immobilization with either soft or hard cervical collars, or the imposition of activity restrictions, were left to the discretion of the surgeon for both patient cohorts.

The initial study end point was twenty-four months for the investigational device exemption clinical trial. However, as a condition for approval of this device, the U.S. Food and Drug Administration (FDA) required the manufacturer to extend its follow-up of enrolled subjects to ten years after surgery. This required each investigative center to reapply for institutional review board approval. Furthermore, renewed informed consent was acquired from willing study participants. Patients were evaluated, according to protocol-defined intervals, preoperatively, at the time of surgery and discharge, at six weeks, and at three, six, twelve, twenty-four, and forty-eight months postoperatively.

Outcomes Assessment

Pain and function were assessed with use of the Neck Disability Index (NDI)\(^6\)\(^7\), the Short Form-36 (SF-36)\(^8\), and numeric rating scales for neck and arm pain. Standardized neurological examinations, including motor and sensory function and reflexes, were recorded by the investigator or nursing staff. Neurological success was defined as maintenance or improvement of all three neurological parameters (motor and sensory function and reflexes). Radiographs were made preoperatively, prior to hospital discharge, and at three, six, twelve, twenty-four, and forty-eight months after surgery. All images were stored centrally and read by independent radiologists. All adverse events were recorded prospectively; categorized, evaluated for causality, and graded for severity with use of World Health Organization (WHO) criteria\(^9\). All were then reviewed for accuracy of categorization, causality, and severity by an independent physician.

The primary end point for the study was a composite measure termed overall success, consisting of the primary effectiveness and safety measures. To have the outcome be considered an overall success, patients had to achieve all of the following: an improvement of 15 points in the NDI, neurological improvement, no serious (WHO grade-3 or 4) adverse events related to the implant or surgical implantation procedure, and no subsequent surgery or intervention that would be classified as a treatment failure.

Patients were evaluated for flexion-extension motion of the cervical spine with use of the Cobb measurement technique on dynamic lateral radiographs of the cervical spine\(^7\). For each measurement, the means from two reviewers were calculated and used for analysis.

Statistical Methods

The primary analysis dataset consisted of all patients who received one of the study treatments. Statistical comparisons were primarily based on the observed and recorded follow-up data. A small number of patients required an additional surgical procedure (removal, revision, or supplemental fixation); their outcomes were recorded as a treatment failure for overall success—the primary study end point. For other outcome variables, the last-observation-carried-forward technique was used for all future evaluation periods.

To compare patients’ demographic and preoperative measures, an analysis of variance was used for continuous variables and the Fisher exact test was used for categorical variables. For comparisons of postoperative mean...
scores or mean score improvements measured in continuous scales, such as NDI scores, analysis of covariance was used with the preoperative score as the covariate. To assess the significance of improvement in the outcome measures within each treatment group, a paired t test was used. For comparison of success or event rates, the Fisher exact test was used to assess the superiority hypothesis.

One-sided p values were reported for most clinical outcomes as defined in the protocol except for surgery and return-to-work data, adverse events, and additional surgical procedures, which were two-sided. A p value of ≤0.05 was considered as significant.

Preoperative Comparison
From May 2002 to October 2004, a total of 463 enrolled subjects were randomly assigned to the study groups, with 242 assigned to the arthroplasty group and 221 to the fusion group. Two-year follow-up was achieved for 230 patients in the arthroplasty group and 194 in the fusion group whose data had been previously reported. The current study involves the 181 arthroplasty and 138 fusion patients at the four-year follow-up (Fig. 2). Preoperative characteristics of the patients and preoperative clinical measures were similar in the two groups (see Appendix). There were no differences in demographics, disease severity (NDI and pain scores), and treated levels between the two groups preoperatively. Similarly, there were no differences between the treatment groups with regard to these preoperative variables for patients with complete forty-eight-month follow-up.

Source of Funding
Medtronic funded the investigational device exemption clinical trial and its continuation as a postapproval study. The clinical trials identification number is NCT00437190.

Results
Overall Success
At every time point postoperatively, the primary outcome measure of overall success was significantly superior for the arthroplasty group compared with the fusion cohort (Fig. 3). At the four-year postoperative mark, overall success was achieved in 85.1% of the patients in the arthroplasty group and
72.5% in the fusion group (Table I) ($p = 0.004$). No deterioration over time was observed in either group.

**Neck Disability Index**

At every time point postoperatively, the two groups were significantly improved from their preoperative state and the arthroplasty group was significantly superior compared with the arthrodesis cohort (Fig. 4). At the four-year postoperative mark, the mean NDI was 13.2 (95% confidence interval [CI]: 10.9 to 15.6) in the arthroplasty group and 19.8 (95% CI: 16.5 to 23.2) in the arthrodesis group ($p < 0.001$). Improvement was seen rapidly within six weeks after surgery, and a plateau was reached by three months in the arthroplasty group and by six months in the fusion group.

**Neck Disability Index Success**

At every time point postoperatively, the percentage of the arthroplasty group who had a reduction of ≥15 points in NDI scores, a criterion for overall success, was significantly higher than that of the fusion group (Fig. 5). At the four-year postoperative mark, NDI success was achieved in 90.6% of the arthroplasty group and in 79.0% of the fusion group ($p = 0.003$).

**Neurological Success**

Neurological success rates at forty-eight months were similar to those observed at twenty-four months. The mean rates were 92.8% and 89.9% in the arthroplasty group and fusion group, respectively, and were not significantly different between the groups (Table I and Appendix).

### TABLE I Rates of Overall Success, Neck Disability Index Success, and Neurological Success at Forty-eight Months

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Arthroplasty Group*</th>
<th>Fusion Group*</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall success</td>
<td>154/181 (85.1)</td>
<td>100/138 (72.5)</td>
<td>0.004</td>
</tr>
<tr>
<td>Neck Disability Index success</td>
<td>164/181 (90.6)</td>
<td>109/138 (79.0)</td>
<td>0.003</td>
</tr>
<tr>
<td>Neurological success</td>
<td>167/180 (92.8)</td>
<td>124/138 (89.9)</td>
<td>0.234</td>
</tr>
</tbody>
</table>

*The values are given as the number with successful outcome/total number with available data, with the percentage in parentheses. †One-sided p values were from Fisher exact test.

Fig. 3

Overall clinical success. The difference between the arthroplasty and fusion groups was significant at all time points.

---

**Results of Cervical Arthroplasty Compared with Anterior Discectomy and Fusion**
Arm Pain
The score for arm pain improved rapidly from a mean preoperative score of 71.2 for both groups to 16.6 (95% CI: 13.1 to 20.2) and 22.4 (95% CI: 17.7 to 27.1) at forty-eight months of follow-up for the arthroplasty and fusion groups, respectively (p = 0.028). Small but significant differences in improvement of the arm pain score were detected between the groups at twelve and forty-eight months, favoring the arthroplasty group over the fusion group (see Appendix).

Neck Pain
The mean preoperative score for neck pain was 75.4 and 74.8 for the arthroplasty and fusion groups, respectively, while at forty-eight months, the mean neck pain score decreased to 20.7 months, favoring the arthroplasty group over the fusion group (see Appendix).

Fig. 4
Neck Disability Index (NDI) score. Significant improvement occurred from baseline in both groups. Furthermore, the difference in improvement between groups was significant at all time points. The results were durable for both groups and did not show significant change over time after three months.

Fig. 5
Neck Disability Index (NDI) success is based on the achievement of a 15% improvement from baseline. Both groups had a high level of success (>75%) after three months, and no deterioration was seen over time. The arthroplasty group was significantly better than the fusion group at all time points.
(95% CI: 17.0 to 24.4) and 30.6 (95% CI: 25.5 to 35.8), respectively. Significant improvement in the neck pain score occurred by six weeks and was maintained at forty-eight months for both groups. The improvement was significantly greater in the arthroplasty group at all time points (Fig. 6).

**SF-36 Summary Scores**

Health-related quality of life was assessed by the SF-36 physical component and mental component summary scores. At forty-eight months, the mean postoperative SF-36 physical component and mental component scores had significantly improved.
for both treatment groups compared with preoperative levels. Furthermore, at forty-eight months, the mean SF-36 physical component score improvement was significantly better in the arthroplasty group compared with the fusion group (\(p = 0.007\)). The mean preoperative SF-36 physical component scores were 32.6 and 31.8 for the arthroplasty and fusion groups, respectively, increasing to 48.4 (95% CI: 46.8 to 49.9) and 44.9 (95% CI: 43.0 to 46.9) at forty-eight months (Fig. 7). The mean mental component score at the preoperative evaluation was 42.3 and 44.6 for the arthroplasty and fusion groups, respectively, increasing to 52.6 (95% CI: 51.1 to 54.0) and 51.9 (95% CI: 50.3 to 53.6) at forty-eight months.

The key pain and functional outcome scores at the preoperative evaluation and at two years and four years of follow-up are presented in Table II.

**Return to Work**

Preoperatively, 64.5% and 65% of the patients in the arthroplasty and fusion groups, respectively, were working. At six weeks after surgery, there were significantly more patients (49.2%) who had returned to work in the arthroplasty group than in the fusion group (39.4%). At forty-eight months, 74.7% and 67.9%, respectively, of the patients were working; the difference was not significant (see Appendix).

**Range of Motion**

The mean cervical spine motion in flexion-extension for the single-level arthroplasty group increased from 6.5° (95% CI: 6.0° to 6.9°) at baseline to 8.0° at twenty-four months and 8.5° (95% CI: 7.7° to 9.2°) at forty-eight months. This increase from baseline was significant at all time points after three months (\(p < 0.05\)). The fusion group showed a mean decrease of motion from 8.4° to 1.1° at forty-eight months.

**Adverse Events**

Adverse events that occurred up to two years postoperatively have been previously reported. For this study, we report only the more severe WHO grade-3 and 4 complications that occurred after twenty-four months and up to the forty-eight-month evaluation window. Forty-four patients in the arthroplasty

<table>
<thead>
<tr>
<th>TABLE II Key Pain and Functional Scores Over Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Neck Disability Index score</td>
</tr>
<tr>
<td>Arm pain score</td>
</tr>
<tr>
<td>Neck pain score</td>
</tr>
<tr>
<td>Short Form-36 physical component</td>
</tr>
</tbody>
</table>

*The values are given as the mean, with the standard deviation in parentheses.

<table>
<thead>
<tr>
<th>TABLE III Number of Patients Who Had Secondary Procedures and Interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second Procedures</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Index level</td>
</tr>
<tr>
<td>Revisions</td>
</tr>
<tr>
<td>Removals</td>
</tr>
<tr>
<td>Supplemental</td>
</tr>
<tr>
<td>fixations</td>
</tr>
<tr>
<td>Reoperations</td>
</tr>
<tr>
<td>External bone</td>
</tr>
<tr>
<td>growth stimulators</td>
</tr>
<tr>
<td>Adjacent levels†</td>
</tr>
<tr>
<td>Other cervical</td>
</tr>
<tr>
<td>levels</td>
</tr>
</tbody>
</table>

*Two-sided \(p\) values were from Fisher exact test. †Some secondary procedures were involved with both index and adjacent levels, and they were counted for index level only in this table.
New technology such as disc arthroplasty requires long-term follow-up to assess durability, the biologic effects of wear, and the response of the prosthesis to its environment. Failure of other joint arthroplasty prostheses does not typically occur until at least five to ten years postoperatively. Spinal arthroplasties similarly need to have serial assessments to determine whether complications such as wear-related failures, device fatigue, or spinal instability have developed. The authors support the FDA requirement that the study sponsor be responsible for attempting to follow all willing study participants for up to eight years after the surgery.

Four years postoperatively, the arthroplasty prosthesis has proven quite durable, with few failures or explants and no change in neck motion over time. Few adverse events occurred in either group after twenty-four months. Spine or device-related events were primarily related to pain at the treated or adjacent disc levels. In the arthroplasty group, explantation occurred in one patient for continued pain. No arthroplasty device required removal for wear or wear-related failure. No new serious adverse event related to the device occurred. In the fusion group, reoperations occurred because of persistent pseudarthrosis and adjacent segment disc disease. Overall, the rates of reoperation are low in both the arthroplasty and fusion groups, with no significant difference detected.

The limitations of this four-year study mainly revolve around the relatively low rate of follow-up compared with the two-year study. At twenty-four months, 230 patients (95%) in the arthroplasty group were evaluated, while at forty-eight months, 181 (75%) were available. For the fusion group, 38% of the 221 enrolled patients failed to return for follow-up at forty-eight months. The lower rate may lead to attrition bias and affect the validity of our results. The actual results may therefore be at variance with the ones reported in this study.

This lower follow-up rate was caused by the original study design, which was set for only two years. A longer follow-up period was requested by FDA regulators; however, this required institutional review board approval at each treatment center and renewed consent and authorization from each patient. As a result, not all centers participated in the longer follow-up study and not all patients consented to the longer follow-up period. Additionally, even for the patients and sites who wished to participate in the longer follow-up period, logistical issues prevented many patients from completing the forty-eight-month follow-up. The follow-up rate was lower at forty-eight months because of the timing of FDA and institutional review board approvals. The forty-eight-month follow-up period was added to the protocol by an amendment. This required time to obtain FDA approval. After FDA approval, the sites had to submit to their institutional review boards, which often took much more time. During this time, the follow-up period for thirty-one patients had ended, and by the time approvals were obtained, the patients were in their sixtieth month of follow-up. Although it is a minor change in the wording of the initial study protocol, the stipulated period of allowed follow-up and the duration of informed consent should either be open-ended or at least be indicated as eight to ten years. This
would preempt the barriers with regard to institutional review boards and informed consent faced in this longer-term follow-up effort.

No deterioration of outcomes after anterior cervical plate stabilization and bone allograft was noticed at four years postoperatively. Clinical improvement, however, continued to be significantly better in the arthroplasty group compared with the fusion group in the primary outcome variables at the time of the four-year follow-up. An additional advantage to the arthroplasty group is that these benefits were obtained while preserving cervical spine motion. As for any motion-sparing device, however, longer-term follow-up is necessary for assessment of potential problems related to bearing surface wear.

Appendix

A table showing a summary of demographic data and figures showing data on neurological success, arm pain score, and the percentage of patients working are available with the online version of this article as a data supplement at jbjs.org.

Note: This study was possible only by the work of the following surgeons and research coordinators.

**Surgeons:** Joseph Alexander, MD, Charles Branch, MD, Frederick Brown, MD, Joseph Cauthen, MD, Jeffrey Cox, MD, Domagala Coric, MD, Richard Cunningham, MD, William Dobkin, MD, Scott Dul, MD, Richard Fessler, MD, Timothy Garvey, MD, Scott Gingold, MD, Robert Hacker, MD, Donald Johnson, MD, J. Patrick Johnson, MD, Mark Kinlock, MD, Allen Let, MD, James Lynch, MD, Patrick McCormick, MD, Luis Mignucci, MD, Paul Nottingham, MD, Glenn T. Palt, MD, Stephen Papadopoulos, MD, Daniel Resnick, MD, John Rhee, MD, K. Daniel Riew, MD, Richard Rovin, MD, Rick Sasso, MD, Michael Smith, MD, Matthew Songer, MD, Brian Sullivan, MD, Lee Thibeau, MD, Donald Whiting, MD, Jeffrey Winfield, MD, and Seth Zeidman, MD.

**Research Coordinators:** Heather Allerton, Anne Anderson, Lisa Armstrong, Rebecca Babcock, Lee Thibeau, MD, Donald Whiting, MD, Jeffrey Winfield, MD, and Seth Zeidman, MD.

**Research Assistants:** Joseph Alexander, MD, Charles Branch, MD, Frederick Brown, MD, Joseph Cauthen, MD, Jeffrey Cox, MD, Domagala Coric, MD, Richard Cunningham, MD, William Dobkin, MD, Scott Dul, MD, Richard Fessler, MD, Timothy Garvey, MD, Scott Gingold, MD, Robert Hacker, MD, Donald Johnson, MD, J. Patrick Johnson, MD, Mark Kinlock, MD, Allen Let, MD, James Lynch, MD, Patrick McCormick, MD, Luis Mignucci, MD, Paul Nottingham, MD, Glenn T. Palt, MD, Stephen Papadopoulos, MD, Daniel Resnick, MD, John Rhee, MD, K. Daniel Riew, MD, Richard Rovin, MD, Rick Sasso, MD, Michael Smith, MD, Matthew Songer, MD, Brian Sullivan, MD, Lee Thibeau, MD, Donald Whiting, MD, Jeffrey Winfield, MD, and Seth Zeidman, MD.

**Reference:**