See the corresponding editorial in this issue, pp 213–215.

J Neurosurg Spine 16:216–228, 2012

Analysis of the three United States Food and Drug Administration investigational device exemption cervical arthroplasty trials

Clinical article

***Cheerag D. Upadhyaya, M.D., M.Sc.,[1] Jau-Ching Wu, M.D.,[1,2] Gregory Trost, M.D.,[7] Regis W. Haid, M.D.,[3] Vincent C. Traynelis, M.D.,[4] Bobby Tay, M.D.,[5] Domagoj Coric, M.D.,[6] and Praveen V. Mummaneni, M.D.[1]**

Departments of [1]Neurological Surgery and [5]Orthopedic Surgery, University of California, San Francisco, California; [2]Department of Neurosurgery, Neurological Institute, Taipei Veterans General Hospital, National Yang-Ming University, Taipei, Taiwan; [3]Atlanta Brain & Spine, Atlanta, Georgia; [4]Department of Neurosurgery, Rush University Medical Center, Chicago, Illinois; [6]Carolina Neurosurgery and Spine Associates, Charlotte, North Carolina; and [7]Department of Neurological Surgery, University of Wisconsin Hospital & Clinics, Madison, Wisconsin

Object. There are now 3 randomized, multicenter, US FDA investigational device exemption, industry-sponsored studies comparing arthroplasty with anterior cervical discectomy and fusion (ACDF) for single-level cervical disease with 2 years of follow-up. These 3 studies evaluated the Prestige ST, Bryan, and ProDisc-C artificial discs. The authors analyzed the combined results of these trials.

Methods. A total of 1213 patients with symptomatic, single-level cervical disc disease were randomized into 2 treatment arms in the 3 randomized trials. Six hundred twenty-one patients received an artificial cervical disc, and 592 patients were treated with ACDF. In the three trials, 94% of the arthroplasty group and 87% of the ACDF group have completed 2 years of follow-up. The authors analyzed the 2-year data from these 3 trials including previously unpublished source data. Statistical analysis was performed with fixed and random effects models.

Results. The authors' analysis revealed that segmental sagittal motion was preserved with arthroplasty (preoperatively 7.26° and postoperatively 8.14°) at the 2-year time point. The fusion rate for ACDF at 2 years was 95%. The Neck Disability Index, 36-Item Short Form Health Survey Mental, and Physical Component Summaries, neck pain, and arm pain scores were not statistically different between the groups at the 24-month follow-up. The arthroplasty group demonstrated superior results at 24 months in neurological success (RR 0.595, $I^2 = 0\%$, $p = 0.006$). The arthroplasty group had a lower rate of secondary surgeries at the 2-year time point (RR 0.44, $I^2 = 0\%$, $p = 0.004$). At the 2-year time point, the reoperation rate for adjacent-level disease was lower for the arthroplasty group when the authors analyzed the combined data set using a fixed effects model (RR 0.460, $I^2 = 2.9\%$, $p = 0.030$), but this finding was not significant using a random effects model. Adverse event reporting was too heterogeneous between the 3 trials to combine for analysis.

Conclusions. Both anterior cervical discectomy and fusion as well as arthroplasty demonstrate excellent 2-year surgical results for the treatment of 1-level cervical disc disease with radiculopathy. Arthroplasty is associated with a lower rate of secondary surgery and a higher rate of neurological success at 2 years. Arthroplasty may be associated with a lower rate of adjacent-level disease at 2 years, but further follow-up and analysis are needed to confirm this finding.
(DOI: 10.3171/2011.6.SPINE10623)

Key Words • **degenerative disc disease** • **cervical arthroplasty** • **clinical trial** • **anterior cervical discectomy and fusion** • **cervical**

Anterior cervical discectomy and fusion is an effective surgical treatment for symptomatic cervical radiculopathy and is associated with very high patient satisfaction scores and a 95% arthrodesis rate.[2,13,22] However, the drawbacks of ACDF include a loss of segmental motion and the need for reoperations at either the index level (for pseudarthrosis treatment) or at the adjacent level (for adjacent-segment disease treatment). These issues led to the development of cervical arthroplasty.[6,13,14,17]

The first 3 large, multicenter, noninferiority, prospective, randomized US FDA-approved IDE studies have been published comparing cervical arthroplasty with ACDF.[5,12,15] Each of these 3 trials compared an artificial disc (Bryan, Prestige ST, or ProDisc-C) to a single-level

Abbreviations used in this paper: ACDF = anterior cervical discectomy and fusion; BMI = body mass index; IDE = investigational device exemption; MCS = Mental Component Summary; NDI = neck disability index; NNT = number needed to treat; PCS = Physical Component Summary; ROM = range of motion; RR = relative risk; SF-36 = 36-Item Short Form Health Survey; VAS = visual analog scale; WMD = weighted mean difference.

* Drs. Upadhyaya and Wu contributed equally to this work

ACDF with allograft and plating. All 3 trials applied similar inclusion and exclusion criteria. All trials used similar clinical outcome measurements, and only these trials have published at least 24 months of follow-up data.

The study objective was to combine the currently available data from these 3 IDE trials.[5,12,15] A combined analysis of these randomized, controlled trials has several potential advantages including a higher statistical power to detect an effect than an individual study alone, superior generalization to the population, and improved ability to control study variations.

Methods

Only randomized, controlled, multicenter, prospective US IDE trials comparing single-level cervical arthroplasty with ACDF that included more than 100 patients in each study arm and had at least 24 months of follow-up published in peer-reviewed journals were included in this analysis.[5,12,15] We chose US FDA IDE studies as they were relatively homogeneous in their inclusion and exclusion criteria and in their outcome measures and follow-up parameters. We did not include unpublished abstracts, underpowered small series, non-FDA studies, or reports of interim (< 2 years) results from selected institutions in this analysis.[1,3,4,16,18–21,23]

The 3 trials that satisfied our criteria were the Prestige ST trial (Medtronic),[12] the Bryan disc trial (Medtronic),[5] and the ProDisc-C trial (Synthes).[15] One or more coauthors from each of these trials participated in this analysis. The coauthors had access to unpublished as well as published summary data from each of these trials. We obtained unpublished, original summary data from each of the trials to complete our analysis.

Statistics were calculated utilizing fixed effects and random effects assumptions to calculate a pooled RR for categorical variables and WMD for continuous variables. A random effects model typically yields a more conservative effect estimate. Given that the trials were studying different devices, we thought it prudent to use both methods in our analysis. If the 2 approaches yielded similar results, we presented the fixed effects model only. However, we presented both models if there was a difference between the models. The I-squared statistic was calculated as an indicator of heterogeneity among studies. An I-squared value was considered as low heterogeneity if the value was between 0% and 40%. The statistics were processed by STATA SE, version 9.2 (StataCorp.).

Artificial Cervical Discs and ACDF

The Prestige ST Cervical Disc, the Bryan Cervical Disc, and the ProDisc-C were the 3 devices in the experimental groups, whereas the control group underwent an interbody fusion with allograft with plate fixation uniformly in the studies.

Study Designs

All studies were prospective, randomized, and multicenter noninferiority studies conducted under an FDA-approved IDE to assess the safety and effectiveness of each artificial disc (compared with single-level ACDF). In the ProDisc-C study, the last observation was not carried forward. In the Prestige ST and Bryan studies, the last observation carried forward was used for patients who underwent secondary surgery that was defined as a treatment failure. The last observation carried forward was not used for lost to follow-up.

Inclusion and Exclusion Criteria

The Prestige ST study included patients older than 18 years with single-level symptomatic cervical spondylosis between the C-3 and C-7 levels who presented with intractable radiculopathy or myelopathy or both. The Bryan study included patients older than 21 years with cervical disc herniation between C-3 and C-7 with intractable radiculopathy or myelopathy. The ProDisc-C studied included patients between 18 and 60 years of age with symptomatic cervical spondylosis between C-3 and C-7 resulting in intractable radiculopathy, myelopathy, and neck pain.

Exclusion criteria were nearly identical among the studies. The studies excluded patients with multilevel cervical spondylosis, evidence of cervical instability (sagittal plane translation of greater than 3.0–3.5 mm or sagittal plane angulation of greater than 20° at a single level), severe spondylosis, marked reduction or absence of motion, collapse of the intervertebral disc space of greater than 50% of its normal height, and metabolic bone diseases.

Clinical Outcome Assessments

The data were extracted for pooling by 2 independent coauthors (C.D.U. and J.C.W.). These independent coauthors were not coinvestigators in any of the 3 trials. The time points extracted included data prior to the operation and data at 24 months postsurgery. The NDI, SF-36 (including MCS and PCS), and scores for neck and arm pain, including frequency and intensity, were pooled for analysis. The Prestige ST and Bryan studies used a numerical rating scale from 0 to 10. The ProDisc-C study used a 100-mm VAS. The arm and neck composite scores were calculated differently between the Prestige ST and Bryan studies. The Prestige ST trial reported a composite score by multiplying the intensity and frequency scores, thereby having a composite score with a value between 0 and 100. The Bryan study summed the scores, divided by 20, and then multiplied by 100%. Therefore, the scores could range from 0% to 100%. Finally, the ProDisc-C study reported VAS intensity and frequency scores separately. Per the FDA's parameters, in all 3 FDA IDE trials, the "neurological success" was defined as the maintenance or improvement in each of the neurological evaluations including sensory, motor, and reflex functions. Failure in any one of the evaluations deemed the case a neurological failure for that time point.

Surgical outcome data were also evaluated. Specifically, the number of supplemental, revision, replacement, removal, and adjacent-segment level surgeries were evaluated and compared. The secondary surgical procedure classification used the US FDA criteria, which defined such surgeries as follows: 1) revision: adjusts or modifies the original implant configuration (such as change of

screw length); 2) removal: removes one or more components of the original implant, replacing with a different type of implant (such as removal of an artificial disc and replacement with a fusion with plating); 3) supplemental fixation: additional spinal devices not approved as part of the protocol are placed (for example, posterior laminar wiring); and 4) reoperation: any surgical procedure at the treated level that does not remove, modify, or add any components (for example, a cervical foraminotomy). We elected to perform a combined analysis of revision surgery, removal surgery, supplemental surgery and reoperation since some categories would be more specific to one arm or another (for example, it would be unlikely for a patient receiving an arthroplasty to undergo supplemental fixation). While the FDA considered an external bone stimulator a form of supplemental fixation, we did not include patients receiving this therapy in this analysis.

In FDA IDE trials, adverse events are any adverse change in the condition of the patient involved in the clinical trial. Adverse events are classified in multiple ways including severity and relatedness. In the ProDisc-C trial the surgical investigator determined the relatedness of the adverse event to the surgery or device. In the Prestige ST and Bryan studies, there was a review process by 2 separate teams that determined the severity, relatedness, and so on of the adverse event. It is possible that the same adverse event could then have been differently categorized between the trials. We were unable to obtain and homogenize the actual adverse events for the 1213 patients in the trials. Therefore, due to the heterogeneity in adverse event reporting, we could not perform a combined analysis of adverse events. This issue regarding the heterogeneity of adverse events has been reported in the medical literature.[7,8,10]

Return to Work

The Prestige ST and the Bryan studies collected information regarding time to return to work. The ProDisc-C study collected return-to-work status information at follow-up but did not obtain the exact number of days postoperatively that the patient returned to work. Therefore, a patient who was working at the 6-month follow-up appointment could have started working at 3.5 or 5.5 months. This difference in reporting in the trials made it difficult for us to analyze the return-to-work data.

Radiographic Outcome Measurement

In the trials, plain radiographic studies were obtained preoperatively and postoperatively for comparison. Neutral and dynamic (flexion and extension) lateral radiographs were used to evaluate for segmental motion of the cervical spine as well as device function. Radiographic outcomes reported at 24 months were combined for analysis.

The ProDisc-C trial had a somewhat stricter definition of fusion, requiring less than 2° of motion on flexion-extension radiographs and no implant loosening (halo/radiolucency). The Prestige ST and Bryan studies accepted less than 4° of motion on flexion-extension radiographs. Both are accepted radiographic assessments of solid fusion in the cervical spine.

Patient Demographic Data

By pooling the 3 clinical trials, a total of 1213 patients (621 in the arthroplasty groups and 592 in the ACDF groups) were analyzed. The follow-up rates at 24 months were 94% in the arthroplasty groups and 87% in the ACDF groups (Table 1). Patient demographics were not clinically different between the ACDF and arthroplasty groups. The mean age in years of the arthroplasty group among the Prestige ST, Bryan, and ProDisc-C studies was 43.3 ± 7.8, 44.4 ± 8.8, and 43.5 ± 7.1, respectively. The mean age in years of the ACDF group was 43.9 ± 8.5, 44.7 ± 6.8, and 42.1 ± 8.4, respectively. The mean BMI in the arthroplasty group among the Prestige ST, Bryan, and ProDisc-C studies was 28.1 ± 5.6, 26.6 ± 4.8, and 27.3 ± 5.5, respectively. The mean BMI in the ACDF group was 28.3 ± 5.1, 27.6 ± 5.0, and 26.4 ± 5.3, respectively. Means are presented as mean ± SD. The percent male population ranged from 44.7% to 51.1% across both arms of all studies.

In the Prestige ST study, the percentage of patients with workers' compensation claims was 11.6% of the arthroplasty group and 13.2% of the ACDF group. In the Bryan study, the percentage of patients with workers' compensation claims was 6.2% and 5.0% for the arthroplasty and ACDF groups, respectively. In the ProDisc-C study, the percentage of patients with workers' compensation claims was 14.6% and 8.5% in the arthroplasty and ACDF groups, respectively. The percentage of patients involved in litigation in the Prestige ST study was 10.9% (arthroplasty group) and 12.1% (ACDF group). The percentage of patients involved in litigation in the Bryan study was 2.5% (arthroplasty group) and 2.7% (ACDF group). Finally, the percentage of patients involved in litigation in the ProDisc-C trial was 2.9% (arthroplasty group) and 3.8% (ACDF group). Workers' compensation was defined in all studies as currently receiving workers' compensation. Litigation was defined as unresolved litigation.

Finally, the percentage of patients using tobacco in the Prestige ST study was 34.4% (arthroplasty group) and 34.7% (ACDF group). In the Bryan study, the percentage of patients using tobacco was 26% (arthroplasty group) and 24% (ACDF group). In the ProDisc-C study, the percentage of patients using tobacco was 34.9% (arthroplasty group) and 33.0% (ACDF group). Study data did not allow for an analysis of differences in fusion rates based on smoking status.

These demographic variables (including mean age, sex, tobacco use, patients receiving workers' compensation, patients with unresolved litigation, mean BMI, and preoperative and 24-month working status) were similar

TABLE 1: The 24-month follow-up rate

Trial	No. of Patients (%)	
	Arthroplasty	ACDF
Prestige ST	253 (91.7%)	220 (83.0%)
Bryan	230 (95.0%)	194 (87.8%)
ProDisc-C	101 (98.0%)	100 (94.3%)
combined	584 (94.0%)	514 (86.8%)

TABLE 2: Demographic comparisons of the studies*

Parameter	Prestige ST		Bryan		ProDisc-C	
	Arthroplasty	ACDF	Arthroplasty	ACDF	Arthroplasty	ACDF
mean age (yrs)	43.3	43.9	44.4	44.7	43.5	42.1
% male	46.4%	46.0%	45.5%	51.1%	46.2%	44.7%
workers' comp†	11.6%	13.2%	6.2%	5.0%	14.6%	8.5%
litigation‡	10.9%	12.1%	2.5%	2.7%	2.9%	3.8%
tobacco use	34.4%	34.7%	26%	24%	34.9%	33.0%
mean BMI	28.1	28.3	26.6	27.6	27.3	26.4
preop working status	65.9%	62.6%	64.5%	65.0%	82.5%	84.9%
24-mo working status	76.3%	75.9%	76.8%	73.6%	82.8%	80.0%

* Within each study there were no statistical demographic differences.
† Defined as currently receiving workers' compensation.
‡ Defined as unresolved litigation.

and are outlined in Table 2. Within each study there were no statistical demographic differences.

Results

Clinical Outcomes of the Combined Analysis

Neck Disability Index. The NDIs in the arthroplasty and the ACDF groups were reduced effectively at the 2-year follow-up compared with preoperative indices. At 24 months, while there appeared to be a trend favoring arthroplasty (Fig. 1), the results were not statistically significant with a WMD of –1.991 (95% CI –4.411 to 0.429, $p = 0.107$, $I^2 = 0\%$) (Table 3).

The SF-36 MCS and PCS Scores. The SF-36 MCS and PCS scores demonstrated significant improvements at the 2-year follow-up compared with preoperative scores. The MCS score had no significant difference between the arthroplasty and ACDF groups at 24 months with a pooled WMD of 0.485 (95% CI –0.865 to 1.834, $p = 0.502$, $I^2 = 0\%$) (Table 3). Likewise, the PCS demonstrated no significant difference between the 2 groups at 24 months with a pooled WMD of 0.103 (95% CI –1.268 to 1.474, $p = 0.153$, $I^2 = 46.8\%$; Fig. 2 and Table 3).

Neck Pain and Arm Pain Scores. The ProDisc-C study used VAS sores to assess neck and arm pain. The Prestige ST and Bryan studies used the numeric rating scale to assess neck and arm pain. There were significant differences in how the neck pain and arm pain scores were calculated and reported among the 3 studies. To overcome this issue, we obtained raw summary data from Medtronic and Synthes for neck and arm pain intensity and frequency scores and converted the scores into a homogenized form. The 24-month neck pain frequency trended toward significance favoring arthroplasty, but did not reach statistical significance, with a WMD of –3.736 (95% CI –7.767 to 0.295, $p = 0.069$, $I^2 = 0\%$; Fig. 3). The 24-month neck pain intensity was not statistically significantly different between groups with a WMD of –1.879 (95% CI –5.782 to 2.024, $p = 0.345$, $I^2 = 32.2\%$; Fig. 3).

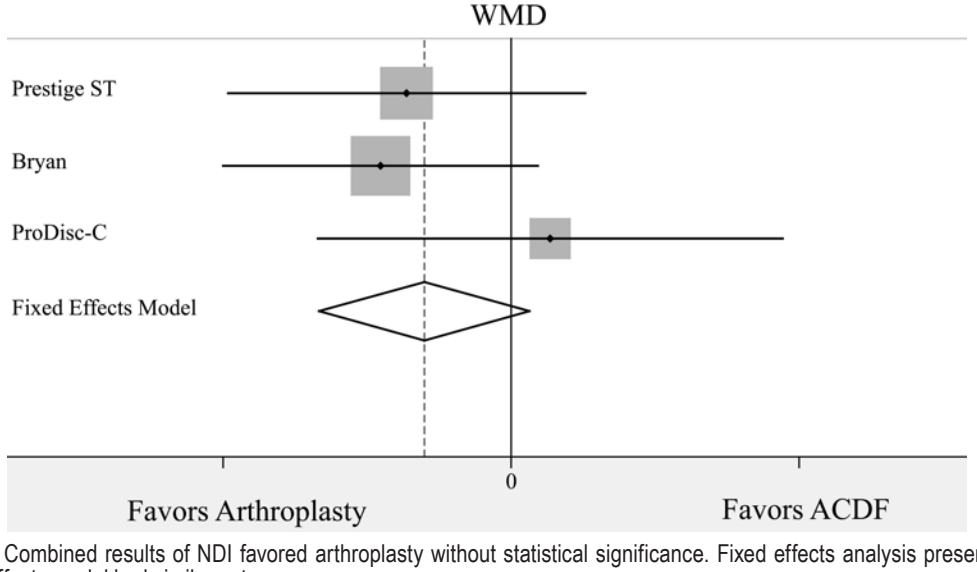

Fig. 1. Combined results of NDI favored arthroplasty without statistical significance. Fixed effects analysis presented. The random effects model had similar outcomes.

TABLE 3: Absolute numbers (means) of NDI, SF-36 MCS, SF-36 PCS, and neurological success rate

		NDI		SF-36 MCS		SF-36 PCS		Neurological Success Rate
Trial	Group	Preop	24 Mos	Preop	24 Mos	Preop	24 Mos	24 Mos
Prestige ST	arthroplasty	55.7	20.0	42.4	44.6	31.9	49.5	91.6%
	ACDF	56.4	22.4	42.7	44.4	32.0	50.2	83.6%
Bryan	arthroplasty	51.4	16.2	42.3	51.7	32.6	47.9	93.9%
	ACDF	50.2	19.2	44.6	51.7	31.8	46.3	90.2%
ProDisc-C	arthroplasty	53.9	21.4	40.6	48.7	34.6	48.2	90.9%
	ACDF	52.2	20.5	39.8	50.5	35.2	46.1	88.0%

The 24-month arm pain frequency also trended toward significance between the groups, favoring arthroplasty, but did not reach statistical significance, with a WMD of –2.798 (95% CI –6.601 to 1.006, p = 0.149, I^2 = 0%; Fig. 4). The 24-month arm pain intensity was not statistically significantly different between groups with a WMD of –0.084 (95% CI –3.441 to 3.273, p = 0.961, I^2 = 0%; Fig. 4 and Table 4).

Neurological Success. The neurological success indicated maintenance or improvement of neurological status assessed via motor, sensory, and reflex examination. The

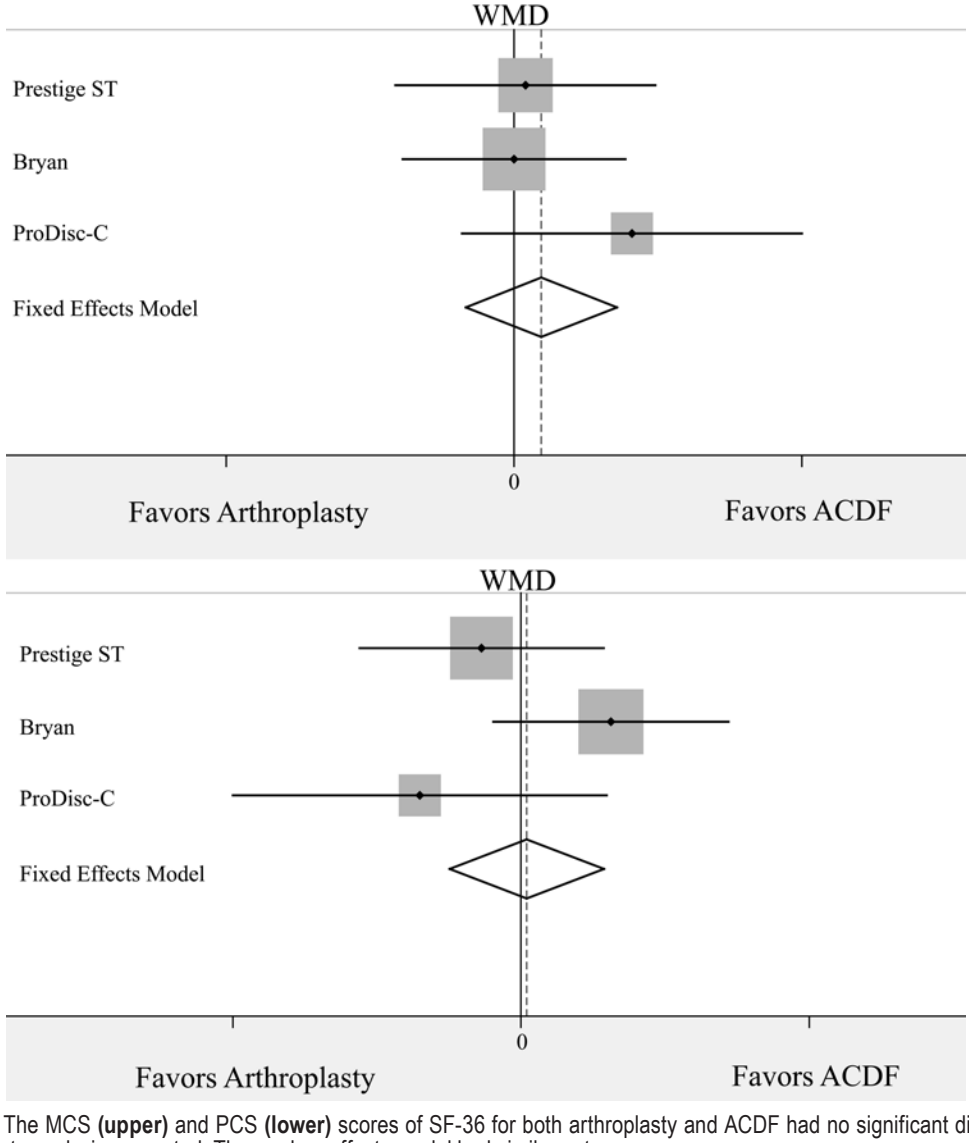

Fig. 2. The MCS **(upper)** and PCS **(lower)** scores of SF-36 for both arthroplasty and ACDF had no significant differences. Fixed effects analysis presented. The random effects model had similar outcomes.

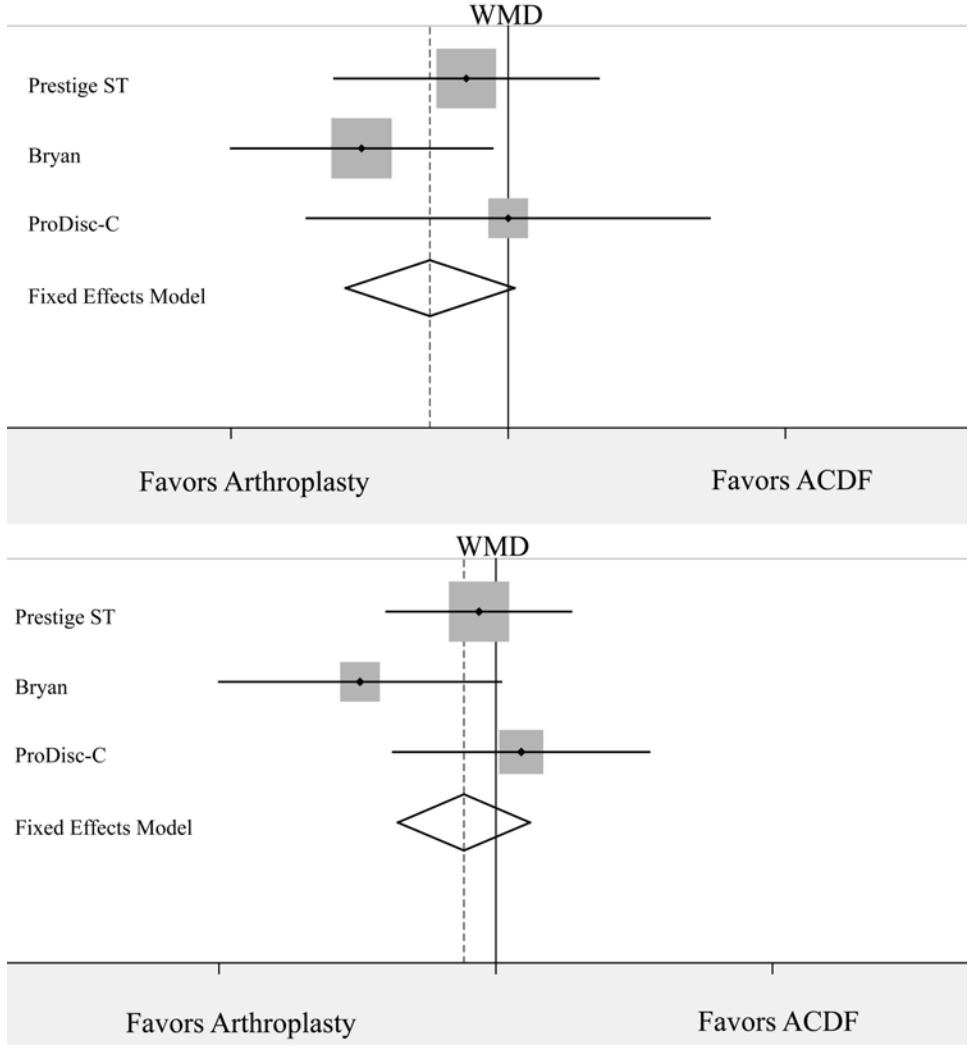

Fig. 3. Neck pain frequency **(upper)** and intensity **(lower)** had no significant differences between arthroplasty and ACDF. Fixed effects analysis presented. The random effects model had similar outcomes.

absolute numbers of neurological success rate in all 3 trials are reported in Table 3. At the 24-month follow-up, neurological success favoring arthroplasty was noted to be statistically significant with an RR of 0.595 (95% CI 0.411–0.862, p = 0.006, I^2 = 0%; Fig. 5) utilizing fixed and random effects models. Assuming a fixed effects model, the NNT was 19 (95% CI 11–66).

Combined Secondary Surgery. Secondary surgeries were defined as supplemental fixation, revision surgery, removal surgery, reoperation at the index level, or surgery for adjacent-level disease. We evaluated surgery for supplemental fixation, revision, removal, and reoperation together as all of these involved the index level (Table 5). Surgery for adjacent-level disease was evaluated separately. This combined analysis for secondary surgery at the index level did reach statistical significance favoring arthroplasty with an RR of 0.44 (95% CI 0.26–0.77, p = 0.004, I^2 = 0%) by fixed effect model and an RR of 0.45 (95% CI 0.26–0.79, p = 0.005, I^2 = 0%) when utilizing random effect model (Fig. 6).

Surgery for Adjacent-Level Disease. At the 24-month follow-up, the Prestige ST trial had 14 patients (5.1%) in the ACDF group and 4 patients (1.5%) in the arthroplasty group who required surgery for adjacent-level disease. In the Bryan study, 8 patients (3.6%) in the ACDF group and 7 patients (2.9%) in the arthroplasty group required surgery for adjacent-level disease at the final 24 months follow-up. Finally, in the ProDisc-C trial, no patients in the arthroplasty group and only a single patient in the ACDF group (0.94%) required surgery for adjacent-level disease. The number of patients requiring surgery due to the development of adjacent-level disease was combined, rather than the number of levels or number of surgeries (Table 6). Utilizing the fixed effects model, there was a statistically significant reduction in the adjacent-level reoperation risk favoring arthroplasty with an RR of 0.460 (95% CI 0.229–0.926, p = 0.030, I^2 = 2.9%; Fig. 7). However, utilizing a random effects model, while there was a trend toward significance favoring arthroplasty, the reduction of adjacent-level reoperation failed to reach statistical significance with an RR of 0.482 (95% CI 0.231–1.008, p =

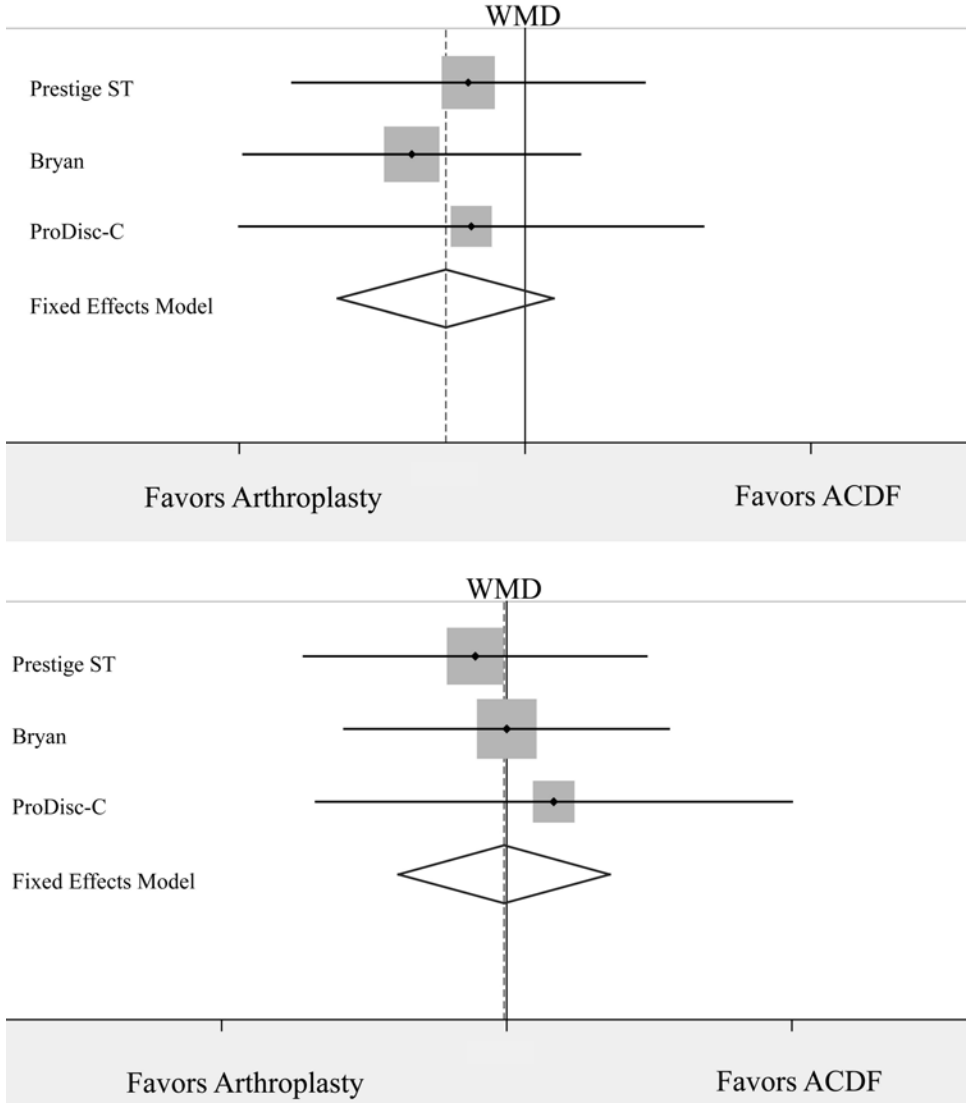

Fig. 4. Arm pain frequency **(upper)** and intensity **(lower)** had no significant differences between arthroplasty and ACDF. Fixed effects analysis presented. The random effects model had similar outcomes.

0.052, $I^2 = 2.9\%$; Fig. 7). Assuming a fixed effects model, the NNT would be 45 (95% CI 25–333).

Return to Work. The median return-to-work time for patients in the experimental group was 45 days in the Prestige ST trial, which was 16 days fewer than the control group. A similar result was observed in the Bryan trial (48 days for the experimental group and 61 days for the control group), and the ProDisc-C trial did not report the specific number of days to return to work. Instead, the ProDisc-C trial reported timeframe ranges for return to work.

Radiographic Outcomes

In the Prestige ST study, there were no reported cases of implant failures, migrations, or subsidence found in the artificial disc group at the 2-year time point. Only 1 patient was reported to develop heterotopic ossification. The arthroplasty patients in the Prestige ST study were prescribed NSAIDs for 2 weeks after surgery. The arthroplasty group in the Prestige ST study demonstrated an average angular motion of 7.55° ± 4.25° preoperatively and 7.73° ± 4.38° at 24 months postoperatively. The ACDF group was noted to have a fusion rate of 97.5% at 24 months.

In the Bryan clinical trial, the average angular motion was 6.45° ± 3.43° preoperatively and 8.1° ± 4.8° at 24 months postoperatively for the arthroplasty group. In the Bryan study, the arthroplasty patients were prescribed NSAIDs for 2 weeks after surgery, and no spontaneous fusions were reported at 24 months. The ACDF group demonstrated a 94.3% fusion rate.

In the ProDisc-C study, the arthroplasty group demonstrated an average flexion-extension ROM of 8.4° ± 4.9° preoperatively and 9.36° ± 5.95° postoperatively. The ACDF group achieved a 90.2% fusion rate. As previously noted, the ProDisc-C trial had a more stringent definition of fusion. Three (2.9%) of the patients in the arthroplasty group developed heterotopic ossification at the index level

TABLE 4: Absolute numbers (means) of neck and arm pain (intensity and frequency)*

		Neck Pain Intensity		Neck Pain Frequency		Arm Pain Intensity		Arm Pain Frequency	
Trial	Group	Preop	24 Mos	Preop	24 Mos	Preop	24 Mos	Preop	24 Mos
Prestige ST	arthroplasty	7.8	2.5	8.5	3.1	7.2	2.2	7.6	2.5
	ACDF	7.9	2.6	8.7	3.3	7.4	2.3	7.9	2.7
Bryan	arthroplasty	7.0	2.1	8.1	2.5	6.8	1.9	7.6	2.0
	ACDF	6.9	2.9	8.0	3.2	6.7	1.9	7.5	2.4
ProDisc-C	arthroplasty	73.0	25.7	85.9	34.0	63.9	20.2	70.6	20.9
	ACDF	65.7	24.2	79.3	30.6	61.0	17.6	70.5	22.8

* The Prestige ST and Bryan studies utilized a numerical rating scale (0–10), whereas the ProDisc-C study utilized the 100-mm VAS score.

(with resulting autofusion). Use of NSAIDs after cervical disc arthroplasty to minimize or prevent heterotopic ossification was optional, and investigators were allowed to follow their standard of care. The ProDisc-C patients who developed heterotopic ossification with autofusion of the index level did not exhibit any decline in functional outcomes compared with patients with well-functioning devices or successful fusions.

After pooling the data, the average angular motion for the arthroplasty levels in the 3 trials at the 24-month postoperative time point was 8.14° ± 4.86°. The overall fusion rate for the pooled ACDF patients was 95% (without any angular motion) (Table 7). These findings indicated that the artificial disc successfully maintained the physiological segmental ROM at the index level of surgery at 24 months and the ACDF did not (Fig. 8).

Discussion

Each of the 3 trials included in this analysis was industry supported and was exclusively focused on a single device (Prestige ST, Bryan, or ProDisc-C). The trials we analyzed were all FDA-regulated IDE studies of very similar designs. They were all prospective, randomized, and multicentered (32 centers for Prestige ST, 30 for Bryan, and 13 for ProDisc-C). We elected to use large (> 100 patients in each arm), English-language, randomized, multicentered, published studies with at least 24 months of follow-up and with similar outcome measures. Furthermore, all included studies were FDA IDE trials. We believed that this would allow for the best estimates of effect to compare arthroplasty with ACDF. While publication and language bias remain a potential consideration, studies that were not included were either very small, had short follow-up, were not Class I, or were not published in the English language.

The difference between fixed effects analysis and random effects analysis should be explained. The assumption underlying a fixed effects analysis is that the same (that is, "fixed") effect is present in all studies being evaluated. For example, in this analysis, a fixed effects model assumes that the RR of adjacent-segment disease is the same across all studies. Another way of looking at it is that if the studies were infinitely large, they would yield the identical result regarding adjacent-level disease. A random effects analysis assumes that the 3 trials are mea-

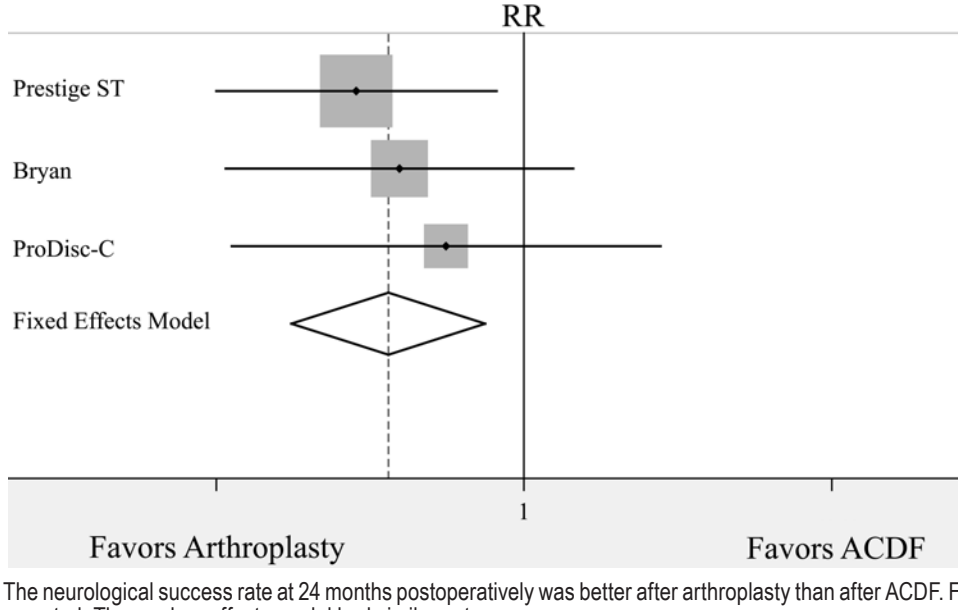

Fig. 5. The neurological success rate at 24 months postoperatively was better after arthroplasty than after ACDF. Fixed effects analysis presented. The random effects model had similar outcomes.

TABLE 5: Secondary surgery at the index level

Trial	Group	No. of Patients	Index-Level Secondary Op				Combined
			Revision	Supplemental*	Removal	Reoperation	
Prestige ST	arthroplasty	276	0	0	6	4	10
	ACDF	265	5	3	11	2	21
Bryan	arthroplasty	242	1	0	3	2	6
	ACDF	221	0	5	3	1	9
ProDisc-C	arthroplasty	103	0	0	2	0	2
	ACDF	106	5	3	0	1	9

* Supplemental fixation excluded the use of external bone stimulator.

suring different treatment effects, but that the treatment effects have a distribution around some central value. A random effects model is also likely to result in a more conservative estimate. While the 3 trials are structured similarly, they do have some differences and are using 3 different cervical arthroplasty devices. Therefore, we thought it was prudent to evaluate using fixed effects and random effects models, presenting both sets of data when significant differences were noted.

The artificial discs included in this analysis were the Prestige ST, Bryan, and ProDisc-C. There are differences in the design of these discs.[9] The Prestige ST is a metal-on-metal, ball-and-trough design made of stainless steel with 2 flanges for vertebral screw fixation to the cephalad and caudad vertebral bodies. The Bryan disc is a coupled titanium alloy surface/shell with a polyurethane "nucleus." ProDisc-C is manufactured with 2 cobalt-chrome endplates with sagittal keels for fixation in the cephalad and caudad vertebral bodies. The core artificial joint surface is made of a polished concave cephalad component that rides on a curved ultra–high molecular weight polyethylene insert fixed to the caudal part.[11]

All 3 studies used noninferiority as the primary hypothesis, in which the goal was to demonstrate the safety and efficacy of the arthroplasty devices. The arthroplasty group at 24 months was found to be noninferior to ACDF in safety and efficacy in the studies. An additional ad hoc hypothesis to testify the superiority of the device revealed the benefit of arthroplasty compared with ACDF in terms of overall success. Overall success, the primary end point, is defined as maintenance of neurological status, absence of adverse events (which included secondary surgeries and excluded adjacent-level surgery), and a 15-point reduction in NDI. Overall success is an FDA-required criterion for these 3 IDE trials. The overall success score is dependent on the adverse events reporting and could not be compared between the studies since we could not compare adverse events. The top 3 adverse events reported in each of the US FDA IDE trials are listed in Table 8.

The adverse events had variable definitions between the Medtronic-sponsored trials (Prestige ST and Bryan) and the Synthes-sponsored trial (ProDisc-C). In FDA IDE trials, adverse events are any adverse change in the condition of the patient involved in the clinical trial. Since any reported change or side effect is considered an adverse event (for example, headaches) during the period of the clinical trial, it is difficult to directly compare the incidence of clinically/surgically relevant adverse events

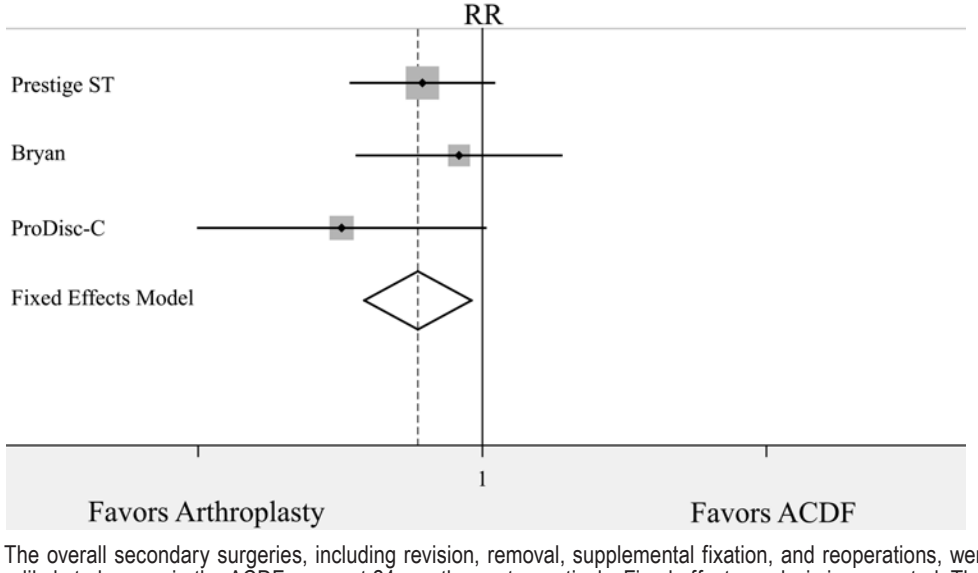

Fig. 6. The overall secondary surgeries, including revision, removal, supplemental fixation, and reoperations, were significantly more likely to happen in the ACDF group at 24 months postoperatively. Fixed effects analysis is presented. The random effects model had similar outcomes.

TABLE 6: Surgery for adjacent-level disease

Trial	Group	No. of Patients	No. of Patients Requiring Op for Adjacent-Level Disease
Prestige ST	arthroplasty	276	4
	ACDF	265	14
Bryan	arthroplasty	242	7
	ACDF	221	8
ProDisc-C	arthroplasty	103	0
	ACDF	106	1

TABLE 7: Radiographic results

Trial	Mean Preop ROM (°)	Mean 24-Mo ROM (°) (arthroplasty)	24-Mo Fusion Rate (ACDF)*
Prestige ST	7.55	7.73	215 (97.5%)
Bryan	6.45	8.10	183 (94.3%)
ProDisc-C†	8.40	9.36	90 (90.2%)
combined	7.26	8.14	488 (95.0%)

* Presented as the number of patients with rate in parentheses.
† The ProDisc-C trial definition of fusion required less than 2° of motion on flexion-extension radiographs. The Prestige and Bryan study definitions of fusion allowed for less than 4° of motion on flexion-extension radiographs.

between studies. Adverse events are classified in multiple ways including severity and relatedness. Categories such as "serious adverse events," "possibly device-related adverse events," and "serious, possibly device-related adverse events" are sometimes used in an attempt to present the information in a clinically/surgically relevant manner. These definitions, however, do not determine parameters such as degree/length of dysphagia that is considered to be serious. In the ProDisc-C trial, the surgical investigator determined the relatedness of the adverse event to the surgery or device. In the Prestige ST and Bryan studies, there was a review process by 2 separate teams that determined the severity, relatedness, and so on of the adverse event. It is possible that the same adverse event could then have been differently categorized among the trials. We were unable to standardize the actual adverse events for the 1213 patients in the trials. Therefore, because of the heterogeneity in adverse event reporting, we could not perform a combined analysis of adverse events.

Neurological success is another FDA parameter. The neurological success indicated maintenance or improvement of neurological status assessed via motor, sensory, and reflex examination of the patient. The neurological success scores were statistically greater for the arthroplasty patients at the 2-year time point. We suspect that the "neurological success" of the arthroplasty patients was better because there were fewer secondary surgeries in the arthroplasty group. When patients need secondary surgeries, their neurological examination typically reveals abnormal findings, and this is reflected in the patient's neurological examination, which determines the neurological success score.

Other differences existed between the 3 clinical trials. The Prestige ST and Bryan studies had more than twice the number of participants in each arm than the ProDisc-C study. The Bryan study had a minimum age limit inclusion criterion of 21 years old, while the Prestige ST and ProDisc-C studies had a minimum age limit of 18 years old. The Bryan study included only cervical disc herniation, whereas spondylosis was also included in the other trials. In the ProDisc-C study, patients were blinded until immediately postoperatively. In the Bryan study and Prestige ST study, there was no blinding. We were unable to obtain results to determine the number of patients who declined intervention after randomization in the Prestige ST study. The Bryan trial reported that 117 patients (37 patients randomized for arthroplasty and 80 randomized for ACDF) declined intervention after learning the results of randomization. Because of blinded study design, no patient in the ProDisc-C study declined intervention after randomization.

Perioperative NSAIDs were routinely prescribed in

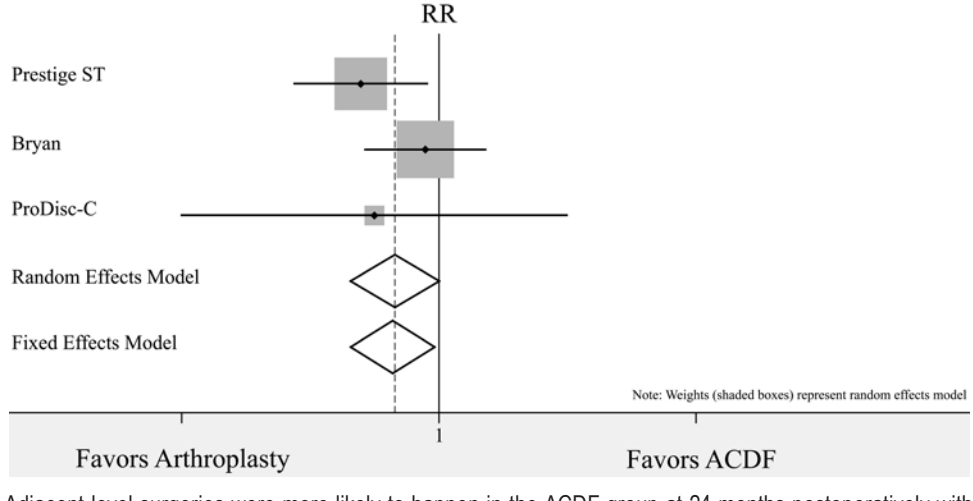

Fig. 7. Adjacent-level surgeries were more likely to happen in the ACDF group at 24 months postoperatively with statistical significance by the fixed effects model. When using the random effects model, there was a trend to favor arthroplasty. p = 0.052.

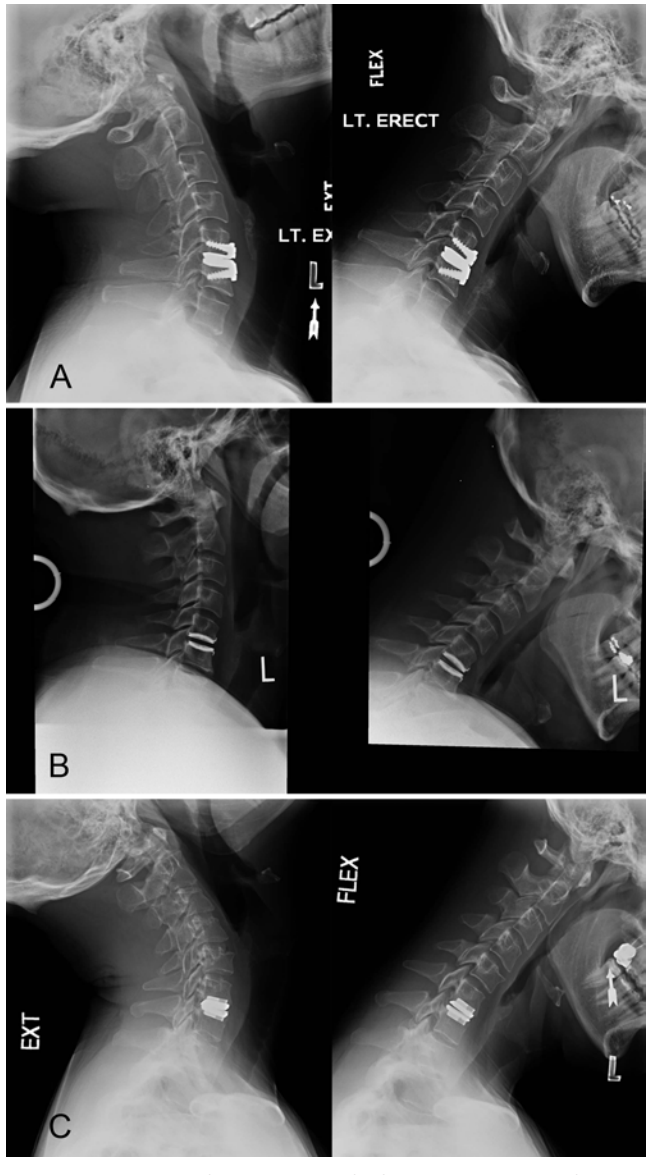

Fig. 8. Examples of maintenance of ROM at the index level for each artificial disc: Prestige ST **(A)**, Bryan **(B)**, and ProDisc-C **(C)**. Extension **(left)** and flexion **(right)** dynamic lateral radiographs of the cervical spine obtained during follow-up.

the Prestige and Bryan trials, but in the ProDisc-C study, the use of NSAIDs was left to the investigator's discretion. The ProDisc-C study reported a small number of heterotopic ossifications with 3 autofusions at the arthroplasty level.[15] The radiographic definition of fusion for the ACDF patients was motion less than 2° in the ProDisc-C trial versus 4° for the other 2 trials. Thus, the ProDisc-C had the lowest fusion rate at 90.2% for the control group among the trials.

The ProDisc-C had the highest follow-up rate at 24 months of 98% (device) and 94.8% (control), Bryan had 95% (device) and 87.8% (control), and Prestige ST had 91% (device) and 83% (control). Cervical arthroplasty demonstrated significantly better results compared with ACDF at the 2-year time point in the following parameters: neurological success and combined secondary

TABLE 8: Top 3 adverse events reported in each of the US FDA IDE trials*

Trial & AEs	Reported % of Total AEs	
	Arthroplasty	ACDF
Prestige ST		
neck &/or arm pain	52.5	49.1
other†	29.3	32.1
other pain‡	29.3	23.8
Bryan		
neck &/or arm pain	45.5	40.3
other†	32.2	23.1
neurological§	22.7	22.6
ProDisc-C		
headache	17.5	
neck pain		20.8
musculoskeletal	17.5	15.1
other op¶		19.8
gastrointestinal	15.5	

* AE = adverse event.
† The total of miscellaneous, infrequent events (for example, nonspecific chest discomfort, eye irritation, and goiter).
‡ All types of nonspinal and nonradicular pain (for example, headache).
§ All types of events related to the nervous system, including numbness, tingling, and other disorders of the nervous system (for example, Bell palsy or Parkinson disease).
¶ Surgery not related to the cervical spine (for example, knee arthroscopy or cesarean section).

surgeries (Table 9). Other outcome measures, including NDI, SF-36 MCS, SF-36 PCS, and neck and arm pain frequency and intensity scores were not significant (Table 9). The higher reoperation rate in the fusion group than in the arthroplasty group could be explained by the inherently higher rate of nonunion in the fusion group (requiring additional surgery) than in the disc arthroplasty group (in which the devices are inherently stable at the time of implantation). In addition, inability of any of these studies to truly blind the surgeon and patient to the surgery that was done introduces the possibility of the existence of

TABLE 9: Summary of combined results

Item	Favor
NDI	not significant
SF-36 MCS	not significant
SF-36 PCS	not significant
neck pain frequency score	not significant
neck pain intensity score	not significant
arm pain frequency score	not significant
arm pain intensity score	not significant
neurological success	arthroplasty, significantly
all secondary op	arthroplasty, significantly
op for ASD	arthroplasty, significance w/ fixed effects model only

significant surgeon bias or patient bias against reoperation in the arthroplasty group. Thus, although the result of the combined trials in regard to arthroplasty and reoperations is encouraging at this point, great care must be taken when assessing the true significance of this result.

Surgery for adjacent-level disease was significantly lower for arthroplasty when utilizing a fixed effects model, but significance was lost utilizing a more conservative random effects model. Assuming a fixed effects model, the NNT was 45 with a relatively broad 95% CI of 25–333. There are multiple contributors to adjacent-segment degeneration including genetics, the patient's activity level, and biomechanical alterations of the cervical motion segments. It is difficult to assess the relative contributions of each of these issues to adjacent-segment degeneration. Clinical trials with longer follow-up will be helpful to further explore this issue.

Conclusions

The currently available 2-year data suggest that cervical arthroplasty is a safe and effective alternative to ACDF to treat patients with single-level cervical disc disease meeting the FDA inclusion and exclusion criteria. Our analysis reveals that arthroplasty may decrease adjacent-level surgery at 2 years, but further follow-up and analysis are needed to confirm this finding. Our analysis also reveals that arthroplasty is associated with fewer secondary surgeries at the 2-year time point.

Disclosures

The following authors participated in the cervical arthroplasty FDA-IDE trials described in this study: Regis W. Haid, M.D., Prestige Trial (Medtronic); Vincent C. Traynelis, M.D., Prestige Trial (Medtronic); Bobby Tay, M.D., ProDisc-C Trial (Synthes); Domagoj Coric, M.D., Bryan Trial (Medtronic); Gregory Trost, M.D., Prestige Trial (Medtronic); and Praveen V. Mummaneni, M.D., Prestige Trial (Medtronic).

This study was not sponsored by industry. Dr. Upadhyaya and Dr. Wu have declared no conflicts of interest. Dr. Haid is a patent holder with Medtronic and receives royalties for Prestige. He is also a consultant for Nuvasive and Globus Medical. Dr. Traynelis is a consultant for and patent holder with Medtronic, and he receives royalties from Medtronic. He also receives royalties from Thieme Medical Publishers and Elsevier, Inc. At the time this paper was written, he was a consultant with United HealthCare. Dr. Tay is a consultant for Stryker Spine, Biomet Spine, and Scient'x, and he received clinical or research support for the study described from AO Spine, Omega, and OREF (educational grants). Dr. Coric is a consultant for and has direct stock ownership in SpinalMotion, Inc.; has direct stock ownership in and is a consultant for Pioneer Surgical; has direct stock ownership in and is a consultant for Spinal Wave; and is a consultant for DePuy Spine. Dr. Trost received a speaking honorarium from Medtronic. Dr. Mummaneni is a past consultant for DePuy Spine and Medtronic (not current). He receives a royalty from DePuy Spine (not related to this manuscript) and a royalty from Quality Medical Publishing.

Author contributions to the study and manuscript preparation include the following. Conception and design: Trost, Mummaneni. Acquisition of data: Wu, Upadhyaya, Coric, Mummaneni. Analysis and interpretation of data: Wu, Upadhyaya, Trost, Mummaneni. Drafting the article: Wu, Upadhyaya. Critically revising the article: Wu, Upadhyaya, Tay, Mummaneni. Statistical analysis: Wu, Upadhyaya. Administrative/technical/material support: Haid, Traynelis, Tay, Coric, Trost, Mummaneni.

References

1. Coric D, Finger F, Boltes P: Prospective randomized controlled study of the Bryan Cervical Disc: early clinical results from a single investigational site. **J Neurosurg Spine 4:** 31–35, 2006
2. Fraser JF, Härtl R: Anterior approaches to fusion of the cervical spine: a metaanalysis of fusion rates. **J Neurosurg Spine 6:**298–303, 2007
3. Goffin J, Casey A, Kehr P, Liebig K, Lind B, Logroscino C, et al: Preliminary clinical experience with the Bryan Cervical Disc Prosthesis. **Neurosurgery 51:**840–847, 2002
4. Goffin J, Van Calenbergh F, van Loon J, Casey A, Kehr P, Liebig K, et al: Intermediate follow-up after treatment of degenerative disc disease with the Bryan Cervical Disc Prosthesis: single-level and bi-level. **Spine (Phila Pa 1976) 28:**2673–2678, 2003
5. Heller JG, Sasso RC, Papadopoulos SM, Anderson PA, Fessler RG, Hacker RJ, et al: Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. **Spine (Phila Pa 1976) 34:**101–107, 2009
6. Hilibrand AS, Carlson GD, Palumbo MA, Jones PK, Bohlman HH: Radiculopathy and myelopathy at segments adjacent to the site of a previous anterior cervical arthrodesis. **J Bone Joint Surg Am 81:**519–528, 1999
7. Ioannidis JP: Adverse events in randomized trials: neglected, restricted, distorted, and silenced. **Arch Intern Med 169:** 1737–1739, 2009
8. Ioannidis JP, Mulrow CD, Goodman SN: Adverse events: the more you search, the more you find. **Ann Intern Med 144:** 298–300, 2006
9. Kang H, Park P, La Marca F, Hollister SJ, Lin CY: Analysis of load sharing on uncovertebral and facet joints at the C5-6 level with implantation of the Bryan, Prestige LP, or ProDisc-C cervical disc prosthesis: an in vivo image-based finite element study. **Neurosurg Focus 28(6):**E9, 2010
10. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al: The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. **BMJ 339:**b2700, 2009
11. Link HD, McAfee PC, Pimenta L: Choosing a cervical disc replacement. **Spine J 4 (6 Suppl):**294S–302S, 2004
12. Mummaneni PV, Burkus JK, Haid RW, Traynelis VC, Zdeblick TA: Clinical and radiographic analysis of cervical disc arthroplasty compared with allograft fusion: a randomized controlled clinical trial. **J Neurosurg Spine 6:**198–209, 2007
13. Mummaneni PV, Haid RW: The future in the care of the cervical spine: interbody fusion and arthroplasty. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. **J Neurosurg Spine 1:**155–159, 2004
14. Mummaneni PV, Robinson JC, Haid RW Jr: Cervical arthroplasty with the PRESTIGE LP cervical disc. **Neurosurgery 60 (4 Suppl 2):**310–315, 2007
15. Murrey D, Janssen M, Delamarter R, Goldstein J, Zigler J, Tay B, et al: Results of the prospective, randomized, controlled multicenter Food and Drug Administration investigational device exemption study of the ProDisc-C total disc replacement versus anterior discectomy and fusion for the treatment of 1-level symptomatic cervical disc disease. **Spine J 9:**275–286, 2009
16. Nabhan A, Ahlhelm F, Shariat K, Pitzen T, Steimer O, Steudel WI, et al: The ProDisc-C prosthesis: clinical and radiological experience 1 year after surgery. **Spine (Phila Pa 1976) 32:** 1935–1941, 2007
17. Park DH, Ramakrishnan P, Cho TH, Lorenz E, Eck JC, Humphreys SC, et al: Effect of lower two-level anterior cervical

fusion on the superior adjacent level. **J Neurosurg Spine 7:** 336–340, 2007
18. Pimenta L, McAfee PC, Cappuccino A, Bellera FP, Link HD: Clinical experience with the new artificial cervical PCM (Cervitech) disc. **Spine J 4 (6 Suppl):**315S–321S, 2004
19. Porchet F, Metcalf NH: Clinical outcomes with the Prestige II cervical disc: preliminary results from a prospective randomized clinical trial. **Neurosurg Focus 17(3):**E6, 2004
20. Riina J, Patel A, Dietz JW, Hoskins JS, Trammell TR, Schwartz DD: Comparison of single-level cervical fusion and a metal-on-metal cervical disc replacement device. **Am J Orthop 37:**E71–E77, 2008
21. Sekhon LH, Sears W, Duggal N: Cervical arthroplasty after previous surgery: results of treating 24 discs in 15 patients. **J Neurosurg Spine 3:**335–341, 2005
22. Tumialán LM, Pan J, Rodts GE, Mummaneni PV: The safety and efficacy of anterior cervical discectomy and fusion with polyetheretherketone spacer and recombinant human bone morphogenetic protein-2: a review of 200 patients. **J Neurosurg Spine 8:**529–535, 2008
23. Wang Y, Cai B, Zhang XS, Xiao SH, Wang Z, Lu N, et al: [Clinical outcomes of single level Bryan cervical disc arthroplasty: a prospective controlled study.] **Zhonghua Wai Ke Za Zhi 46:**328–332, 2008 (Chinese)

Manuscript submitted September 3, 2010.
Accepted June 6, 2011.
Please include this information when citing this paper: published online December 23, 2011; DOI: 10.3171/2011.6.SPINE10623.

Address correspondence to: Jau-Ching Wu, M.D., Department of Neurosurgery, Neurological Institute, Taipei Veterans General Hospital Room 509, 17F, No. 201, Shih-Pai Road, Sec. 2, Beitou, Taipei 11217, Taiwan. email: jauching@gmail.com.