Comparing 4-Year Changes in Patient-Reported Outcomes Following Ankle Arthroplasty and Arthrodesis

Bruce J. Sangeorzan, MD, William R. Ledoux, PhD, Jane B. Shofer, MS, James Davitt, MD, John G. Anderson, MD, Donald Bohay, MD, J. Chris Coetzee, MD, John Maskill, MD, Michael Brage, MD, and Daniel C. Norvell, PhD

Investigation performed at the Center for Limb Loss and MoBility (CLiMB), VA Puget Sound Health Care System, Seattle, Washington

Background: The rate of total ankle arthroplasty (TAA) is increasing relative to ankle arthrodesis (AA) for patients seeking surgical treatment for end-stage ankle arthritis. Patients and providers would benefit from a more complete understanding of the rate of improvement, the average length of time to achieve maximal function and minimal pain, and whether there is a greater decline in function or an increase in pain over time following TAA compared with AA. The objectives of this study were to compare treatment changes in overall physical and mental function and ankle-specific function, as well as pain intensity at 48 months after TAA or AA in order to determine if the improvements are sustained.

Methods: This was a multisite prospective cohort study that included 517 participants (414 TAA and 103 AA) who presented for surgical treatment. Participants were compared 48 months after surgery using the Foot and Ankle Ability Measure (FAAM) Activities of Daily Living and Sports subscales (0 to 100 points), the Short Form-36 (SF-36) Physical and Mental Component Summary (PCS and MCS) scores (0 to 100 points), and pain scores (0 to 10 points).

Results: Both groups achieved significant improvement in the 2 FAAM measures, the SF-36 PCS score, and all of the pain measures at 48 months after surgey ($p < 0.001$). Mean improvements from baseline in patients undergoing TAA for the FAAM Activities of Daily Living, FAAM Sports, and SF-36 scores were at least 9 points, 8 points, and 3.5 points, respectively, which were higher than in those undergoing AA. Mean improvements in worst and average pain were at least 0.9 point higher in patients undergoing TAA than in those undergoing AA at 12, 24, and 36 months. These differences were attenuated by 48 months. For both treatments, all improvements from baseline to 24 months had been maintained at 48 months.

Conclusions: When both procedures are performed by the same group of surgeons, patients who undergo TAA or AA for end-stage ankle arthritis have significant improvement in overall function, ankle-specific function, and pain at 48 months after surgery, with better functional improvement in the TAA group.

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

Patients seeking surgical treatment for end-stage ankle arthritis have 2 treatment options: ankle arthrodesis (AA) (Fig. 1-A) or total ankle arthroplasty (TAA) (Fig. 1-B). A study evaluating the National Inpatient Sample database reported that the proportion of TAA procedures increased from 14% in 2007 to 45% in 2013[1]. There are dozens

Disclosure: This project was funded by National Institutes of Health (NIH) grant number R01 AR056316. The NIH played no role in the investigation. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work and "yes" to indicate that the author had other relationships or activities that could be perceived to influence, or have the potential to influence, what was written in this work (http://links.lww.com/JBJS/G380).

A **data-sharing statement** is provided with the online version of the article (http://links.lww.com/JBJS/G381).

of published case series on TAA and AA, but there is no strong, consistent evidence regarding long-term comparative effectiveness or safety. Given these concerns, numerous researchers have called for prospective studies that directly compare TAA with AA[2-6]. Several studies with short-term follow-up have demonstrated either equality between the treatments or greater function but higher revision rates after TAA[7-14]. However, studies are limited by a lack of long-term follow-up, with most extending for ≤2 years. A recent meta-analysis[15] that evaluated 10 studies concluded that these 2 treatments lead to similar outcomes but higher rates of complications following TAA in the short term; the authors recommended additional high-quality research comparing these treatments over a longer follow-up period.

A prospective cohort study with 2 years of follow-up that was performed by our investigators demonstrated a superiority in patient-reported outcomes in the TAA group and equality in ankle-specific complications, including revisions, in both groups[9]. The results included a more rapid improvement in patient-reported outcomes in the TAA group. Understanding whether there is a decline in function or an increase in pain over time with either treatment option may help both surgeons and patients make educated decisions since it appears that both treatments lead to improvement.

The goal of this study was to evaluate outcomes in the same cohort of patients at 4 years after surgery to determine if improvements and differences are sustained. This study addressed the following objectives for TAA and AA, from preoperatively to 48 months after surgery: (1) comparing the changes in overall physical and mental function, ankle-specific function, and pain intensity, (2) evaluating whether improvements that had been achieved at 24 months after surgery were sustained at 48 months, and (3) assessing patient satisfaction with the surgery at 48 months after surgery.

Materials and Methods

Study Design

Six sites prospectively recruited patients from May 2012 to May 2015. The study was approved by the human subjects review board at each participating institution and was initially designed as a randomized controlled trial (RCT) with a patient-preference cohort.

Participants

Initially, 812 patients were screened and were found to be candidates for both TAA and AA. One hundred and sixty-one patients were excluded based on criteria that included factors that affected ambulatory function or an inability to

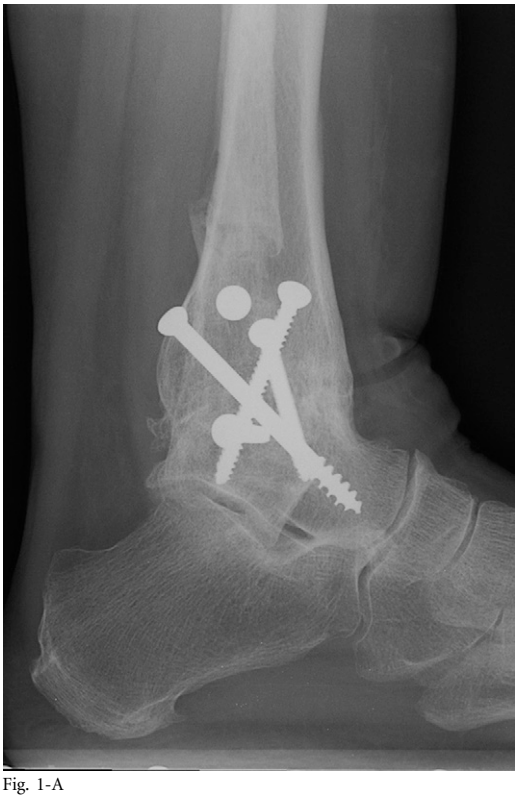

Fig. 1-A

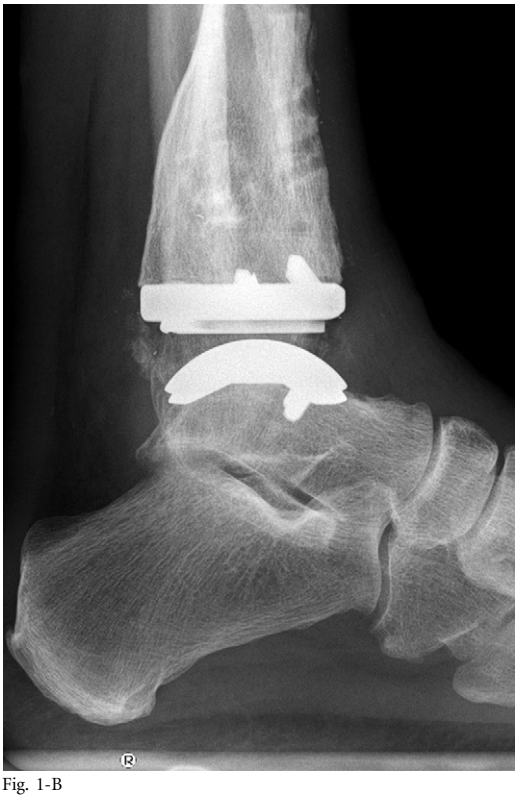

Fig. 1-B

Figs. 1-A and 1-B Radiographs. **Fig. 1-A** An AA was performed for posttraumatic arthritis. The screws were used to stabilize the bones during healing. There is no additional motion of the ankle joint, and no joint space remains. The remaining motion occurs via the surrounding joints. **Fig. 1-B** A TAA was performed for posttraumatic arthritis. The bone on either side of the joint was removed and replaced with metal. The motion is facilitated by a polymer spacer between the metal parts.

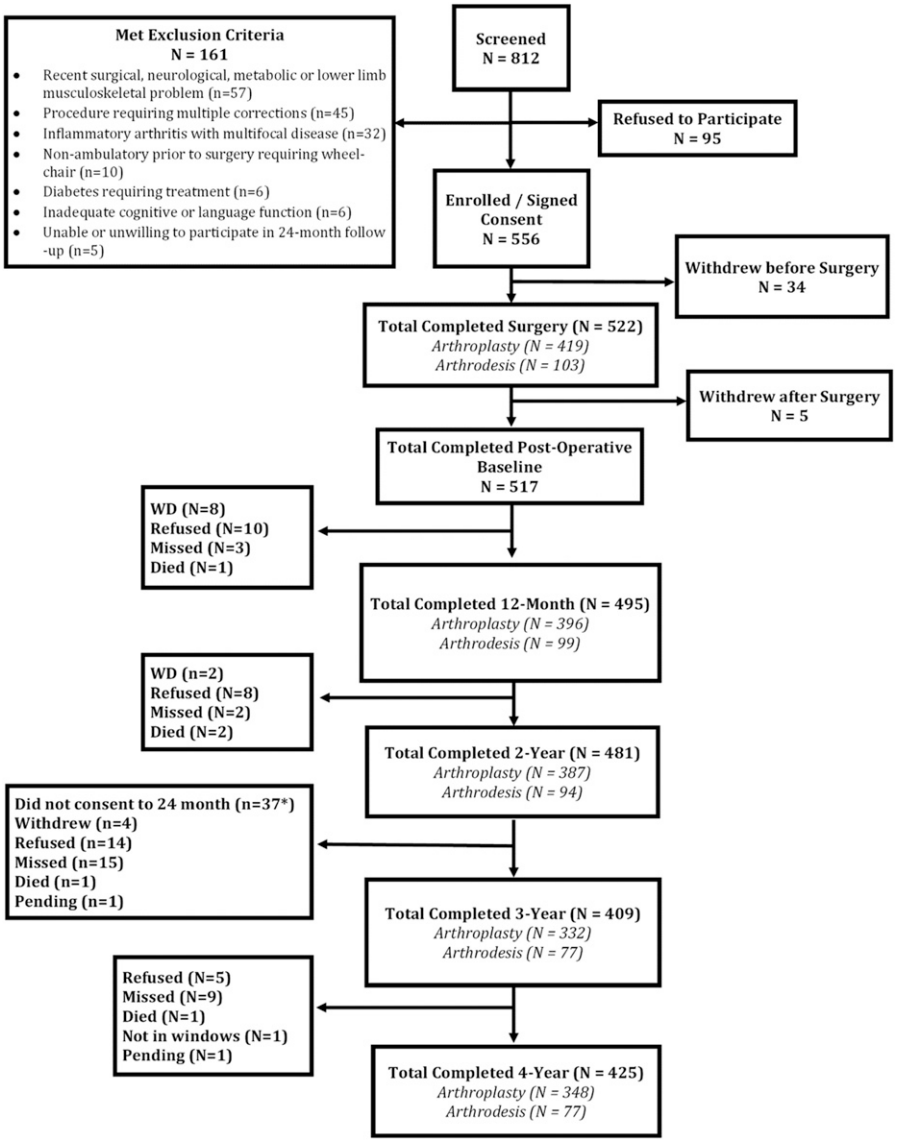

Fig. 2

STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) diagram summarizing the patients who were screened and enrolled, and those who completed each follow-up. WD = withdrew.

participate in patient reporting of outcomes, and 95 declined to participate. The specific inclusion and exclusion criteria were reported in a prior publication in 2019[9]. Thirty-four patients who had enrolled withdrew prior to surgery, leaving 522 patients who agreed to participate in the study described herein. The reasons for noninclusion are listed in Figure 2.

Risk Factors and Interventions

Preoperative patient characteristics were obtained through interviews and medical record review as outlined in the prior publication by Norvell et al.[9] (Table I). To be eligible to participate, surgeons were required to have performed a minimum of 30 arthrodeses and 30 arthroplasties in order to ensure proficiency, per prior recommendations[16].

Outcomes

The following outcomes were measured at 12, 24, 36, and 48 months after surgery. The Foot and Ankle Ability Measure (FAAM)[17] was selected to measure patient-reported ankle-specific function. This measure generates 2 separate subscales: Activities of Daily Living and Sports (0 to 100 points for each subscale, with higher scores representing greater function). The minimal clinically important difference (MCID) is 9 points[18]. Overall physical and mental function were measured with the Short Form-36 (SF-36) Physical Component Summary (PCS)

TABLE I Comparison of Baseline Characteristics Between TAA and AA Groups[9]

Characteristic	TAA* (N = 414)	AA (N = 103)	P Value	Total (N = 517)
Male sex†	237 (57)	61 (59)	0.72	298 (58)
Age‡ *(yr)*	63.2 ± 9.7	54.2 ± 12.7	<0.01	61.4 ± 10.9
BMI‡ *(kg/m²)*	29.9 ± 5.5	33.1 ± 7.2	<0.01	30.5 ± 6.0
Race†			0.78	
White/Caucasian	403 (97)	100 (97)		503 (97)
Non-white	11 (3)	3 (3)		14 (3)
Marital status† (married)	327 (79)	71 (69)	0.08	398 (77)
College graduate†	246 (59)	58 (56)	0.53	304 (59)
Full-time employment†	155 (37)	53 (51)§	0.01	208 (40)
Income† (≥$75,001)	193 (47)	33 (32)	<0.01	226 (44)
Cause of end-stage ankle arthritis†			0.05	
Posttraumatic	213 (51)	70 (68)		283 (55)
Recurrent sprains	53 (13)	10 (10)		63 (12)
Degenerative	71 (17)	10 (10)		81 (16)
Instability	50 (12)	6 (6)		56 (11)
Misalignment	16 (4)	3 (3)		19 (4)
Other	11 (3)	4 (4)		15 (3)
Previous foot/ankle surgery†	237 (57)	77 (75)	<0.01	314 (61)
Radiographic findings				
Osteoarthritis severity grade†			0.13	
0-1	4 (1)	2 (2)		6 (1)
2	21 (5)	9 (9)		30 (6)
3	93 (22)	29 (28)		122 (24)
4	296 (71)	63 (61)		359 (69)
Alignment‡ *(°)*	8.7 ± 8.8	9.2 ± 9.3	0.64	8.8 ± 8.9
Subluxation‡ *(°)*	16.4 ± 18.6#	13.5 ± 17.8	0.15	15.8 ± 18.4
Medical history/comorbidities				
Osteoporosis†	43 (10)	5 (5)	0.08	48 (9)
Depression and/or anxiety†	31 (7)	22 (21)	<0.01	53 (10)
Degenerative disc disease†	75 (18)	23 (22)	0.33	98 (19)
FCI‡	2.7 ± 1.5	3.4 ± 2.1	<0.01	2.9 ± 1.7
Current smoker†	8 (2)	10 (10)	<0.01	18 (3)
Current alcohol use ≥6 times/wk†	61 (15)	11 (11)	0.19	72 (14)

*Of the 414 TAAs, 211 (51.0%) were a Salto Talaris Ankle (Integra LifeSciences); 174 (42.0%), an INBONE Total Ankle System (Wright Medical); 23 (5.6%), a STAR (Scandinavian Total Ankle Replacement) (Stryker); 5 (1.2%), a Trabecular Metal Total Ankle (Zimmer Biomet); and 1 (0.2%), other. †The values are given as the number with the percentage in parentheses. ‡The values are given as the mean and standard deviation. §Data missing for 1 patient. #Data missing for 2 patients. (Reproduced from: Norvell DC, Ledoux WR, Shofer JB, Hansen ST, Davitt J, Anderson JG, Bohay D, Coetzee JC, Maskill J, Brage M, Houghton M, Sangeorzan BJ. Effectiveness and safety of ankle arthrodesis versus arthroplasty: a prospective multicenter study. J Bone Joint Surg Am. 2019 Aug 21;101[16]:1485-94.)

and Mental Component Summary (MCS) scores (0 to 100 points). Pain was assessed using 4 questions from the Chronic Pain Grade (CPG) questionnaire[19], which included the intensity of present and worst ankle pain, average ankle pain, and overall pain in the past 6 months (0 to 10 points each).

Patient satisfaction was assessed with 3 standard questions: (1) If you had to do it over again, would you have surgery on your ankle (yes or no)? (2) Have your expectations been met (yes or no)? (3) How would you rate your overall satisfaction with the results of your ankle surgery (a: very satisfied, b: satisfied, c: uncertain, or d: dissatisfied)? These 3 questions have been correlated with patient expectations and patient function[20], and the overall satisfaction question was validated against the Nottingham Health Profile, the SF-36, the SF-12, the Oxford-12, and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)[21,22].

TABLE II Change in Outcome Measures from Preoperatively to 12, 24, 36, and 48 Months and Comparison of These Changes Between the TAA and AA Groups*

	TAA (N = 414)	AA (N = 103)	TAA Minus AA
FAAM activities of daily living			
Preop.	46.4 ± 1.3	48.2 ± 2.1	
12 mo – preop.	**34.7 ± 0.9 (32.5, 36.9)**	**23.3 ± 1.8 (18.7, 27.9)**	**11.4 ± 2.1 (6.2, 16.6)**
24 mo – preop.	**35.2 ± 0.9 (33.0, 37.4)**	**26.0 ± 1.8 (21.3, 30.7)**	**9.2 ± 2.1 (3.9, 14.5)**
36 mo – preop.	**34.2 ± 0.9 (31.9, 36.5)**	**22.9 ± 2.0 (17.9, 27.9)**	**11.3 ± 2.2 (5.7, 16.9)**
48 mo – preop.	**33.8 ± 0.9 (31.5, 36.0)**	**23.5 ± 2.0 (18.5, 28.5)**	**10.3 ± 2.2 (4.7, 15.9)**
P value	<0.0001	<0.0001	<0.0001
FAAM Sports			
Preop.	19.8 ± 2.0	21.9 ± 3.0	
12 mo – preop.	**39.4 ± 1.3 (36.1, 42.6)**	**23.4 ± 2.8 (16.4, 30.5)**	**15.9 ± 3.1 (8.0, 23.9)**
24 mo – preop.	**39.9 ± 1.3 (36.6, 43.2)**	**30.9 ± 2.8 (23.8, 37.9)**	**9.0 ± 3.2 (1.0, 17.0)**
36 mo – preop.	**38.6 ± 1.4 (35.1, 42.1)**	**26.9 ± 3.0 (19.4, 34.5)**	**11.7 ± 3.4 (3.1, 20.2)**
48 mo – preop.	**37.4 ± 1.4 (34.0, 40.9)**	**29.4 ± 3.0 (21.8, 36.9)**	8.1 ± 3.4 (−0.5, 16.6)
P value	<0.0001	<0.0001	<0.0001
SF-36 PCS			
Preop.	34.1 ± 0.4	36.0 ± 0.9	
12 mo – preop.	**12.7 ± 0.4 (11.6, 13.8)**	**8.3 ± 0.9 (6.0, 10.6)**	**4.4 ± 1.0 (1.8, 7.1)**
24 mo – preop.	**12.3 ± 0.4 (11.2, 13.4)**	**8.3 ± 0.9 (5.9, 10.7)**	**4.0 ± 1.1 (1.4, 6.7)**
36 mo – preop.	**11.9 ± 0.5 (10.8, 13.1)**	**7.0 ± 1.0 (4.5, 9.5)**	**4.9 ± 1.1 (2.1, 7.8)**
48 mo – preop.	**11.0 ± 0.4 (9.9, 12.1)**	**7.5 ± 1.0 (5.0, 10.0)**	**3.5 ± 1.1 (0.7, 6.3)**
P value	<0.0001	<0.0001	0.0001
SF-36 MCS			
Preop.	55.5 ± 0.4	54.1 ± 0.9	
12 mo – preop.	1.0 ± 0.4 (0.0, 2.1)	0.9 ± 0.9 (−1.4, 3.1)	0.2 ± 1.0 (−2.3, 2.7)
24 mo – preop.	0.8 ± 0.4 (−0.2, 1.8)	1.9 ± 0.9 (−0.3, 4.2)	−1.2 ± 1.0 (−3.7, 1.4)
36 mo – preop.	0.3 ± 0.4 (−0.8, 1.4)	0.8 ± 0.9 (−1.5, 3.2)	−0.5 ± 1.1 (−3.2, 2.2)
48 mo – preop.	0.4 ± 0.4 (−0.7, 1.5)	1.0 ± 0.9 (−1.4, 3.4)	−0.6 ± 1.1 (−3.3, 2.2)
P value	0.65	0.014	0.71
Present pain			
Preop.	5.1 ± 0.1	5.5 ± 0.2	
12 mo – preop.	**−3.8 ± 0.1 (−4.1, −3.5)**	**−3.7 ± 0.3 (−4.4, −3.1)**	−0.1 ± 0.3 (−0.9, 0.6)
24 mo – preop.	**−3.8 ± 0.1 (−4.1, −3.5)**	**−3.6 ± 0.3 (−4.3, −2.9)**	−0.2 ± 0.3 (−0.9, 0.6)
36 mo – preop.	**−3.6 ± 0.1 (−4.0, −3.3)**	**−3.3 ± 0.3 (−4.0, −2.6)**	−0.4 ± 0.3 (−1.2, 0.5)
48 mo – preop.	**−3.5 ± 0.1 (−3.8, −3.2)**	**−3.7 ± 0.3 (−4.4, −3.0)**	0.2 ± 0.3 (−0.6, 1.0)
P value	<0.0001	<0.0001	1.0
Worst pain			
Preop.	8.5 ± 0.2	8.4 ± 0.3	
12 mo – preop.	**−5.0 ± 0.1 (−5.3, −4.6)**	**−4.1 ± 0.3 (−4.8, −3.3)**	**−0.9 ± 0.3 (−1.7, −0.1)**
24 m0 – preop.	**−5.5 ± 0.1 (−5.8, −5.1)**	**−4.3 ± 0.3 (−5.1, −3.6)**	**−1.1 ± 0.3 (−2.0, −0.3)**
36 mo – preop.	**−5.4 ± 0.1 (−5.8, −5.0)**	**−4.3 ± 0.3 (−5.1, −3.5)**	**−1.1 ± 0.4 (−2.0, −0.2)**
48 mo – preop.	**−5.3 ± 0.1 (−5.7, −5.0)**	**−4.6 ± 0.3 (−5.4, −3.8)**	−0.7 ± 0.4 (−1.6, 0.2)
P value	<0.0001	<0.0001	0.0035

continued

TABLE II (*continued*)

	TAA (N = 414)	AA (N = 103)	TAA Minus AA
Average pain			
Preop.	6.3 ± 0.1	6.0 ± 0.2	
12 mo – preop.	**−4.2 ± 0.1 (−4.5, −3.9)**	**−3.3 ± 0.2 (−3.9, −2.7)**	**−0.9 ± 0.3 (−1.6, −0.2)**
24 mo – preop.	**−4.4 ± 0.1 (−4.7, −4.1)**	**−3.5 ± 0.2 (−4.1, −2.9)**	**−0.9 ± 0.3 (−1.6, −0.2)**
36 mo – preop.	**−4.4 ± 0.1 (−4.7, −4.1)**	**−3.3 ± 0.3 (−3.9, −2.6)**	**−1.1 ± 0.3 (−1.8, −0.4)**
48 mo – preop.	**−4.3 ± 0.1 (−4.6, −4.0)**	**−3.6 ± 0.3 (−4.3, −3.0)**	−0.6 ± 0.3 (−1.4, 0.1)
P value	<0.0001	<0.0001	0.0006

*From linear mixed-effects regression of outcome on study visit by surgery type interaction. All models included confounders of age, sex, and BMI. Additional confounders include previous surgery history for FAAM Activities of Daily Living; depression and Functional Comorbidity Index (FCI) for FAAM Sports and SF-36 PCS; end-stage ankle arthritis cause, depression and anxiety history, FCI, and current smoking for SF-36 MCS; and anxiety history and smoking use for current pain, employment for worst pain, and employment and current smoking for average pain. Site and patient within site were modeled as random. The scores are given as the mean and standard deviation, with or without the 95% CI in parentheses. Differences in bold are significant. During the 4-year period, there were 36 (8.7%) revisions in the TAA group and 18 (17.5%) revisions in the AA group. These participants were not excluded from the analysis; therefore, their data contributed to results in this table.

Statistical Analysis

Differences in preoperative characteristics by surgery type were assessed using 2-sample t tests and chi-square tests. The traditional confounders of age, sex, and body mass index (BMI) were included as covariates in all analyses. Additional potential confounders included the cause of arthritis, prior surgery, employment status, depression, anxiety, and current smoking.

Linear mixed-effects regression was used to determine if there were differences in the postoperative improvement in each continuous outcome by surgical procedure, and if improvement that had been found at 24 months had been maintained at 48 months. Patient-reported outcomes (the FAAM Activities of Daily Living and Sports subscales, SF-36 PCS, SF-36 MCS, and pain questions) were the dependent variables. Study visit (baseline, 12, 24, 36, and 48 months after treatment), surgical procedure, and potential confounders were the independent fixed main effects. All of the models included interactions between the main effect variables and

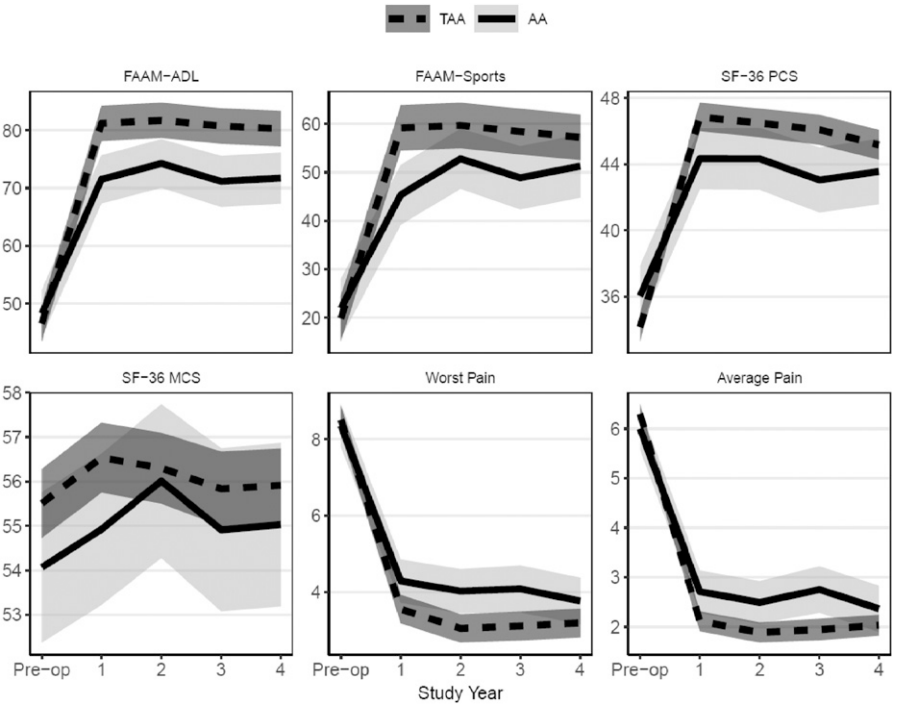

Fig. 3

Mean outcomes and 95% confidence intervals by study year and surgical procedure estimated from linear mixed-effects regression of outcome on study year-by-surgical procedure interaction and age, sex, BMI, and outcome-specific confounders as covariates, with the patient as a random effect. ADL = activities of daily living.

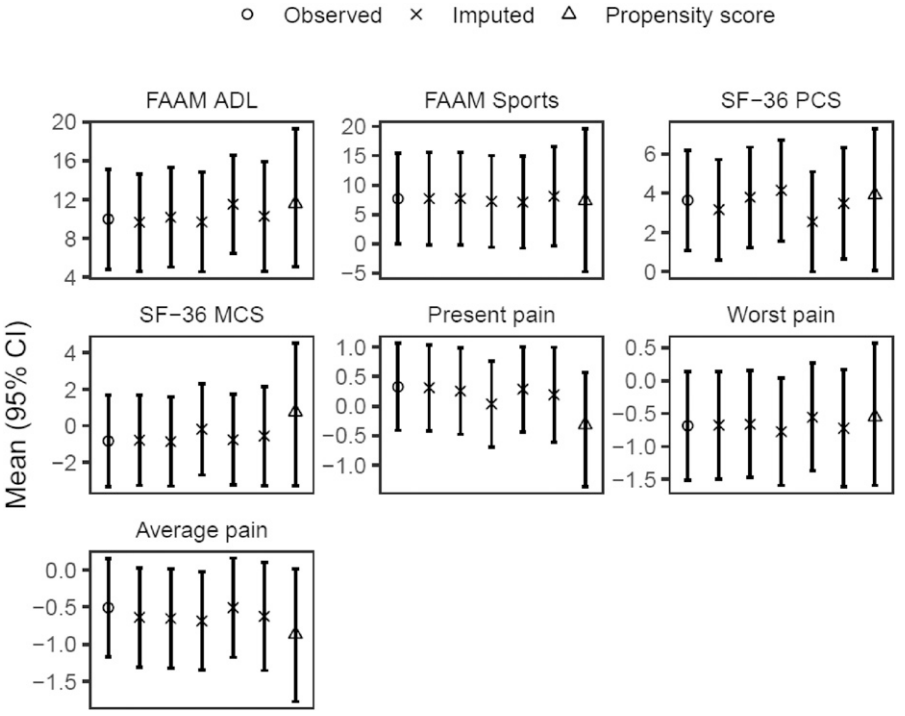

Fig. 4

Mean difference (y axis) in 48-month improvement by surgery group (TAA minus AA) from linear mixed-effects models of outcomes on study visit-by-surgery group interaction as performed on the observed data, the 5 imputed data sets, and the observed data using propensity score weights. ADL = activities of daily living.

study visit to estimate the difference in postoperative improvement by surgical procedure, adjusting for improvement due to the potential confounders. Surgical institution and patient were random effects. Means and 95% confidence intervals (CIs) for improvement stratified by surgical procedure and differences in improvement from baseline and from 24 months to 48 months by surgical procedure were estimated using simultaneous inference[23].

As a sensitivity analysis, to control for confounding by indication, we also employed a propensity score analysis. First, propensity scores for treatment assignment were estimated using logistic regression of treatment assignment on age, BMI, sex, site, and the confounders listed above. Second, these scores were used to estimate inverse probability weights that were then applied to the linear mixed-effects regressions described above. This 2-step procedure was bootstrapped to obtain estimates of variability, accounting for the error in propensity score estimates. Efficacy of the weights was assessed by computing standardized differences between treatment groups, with the aim to have all standardized differences be 0.1, as recommended by Austin and Stuart[24]. In addition, multiple imputation of missing follow-up outcomes was carried out to test the robustness of findings due to the influence of incomplete follow-up of study participants using the MICE (multiple imputation by chained equations) algorithm[25]. Five imputed data sets were generated and checked for algorithm convergence, and then the models were reapplied. Analyses were performed using R version 3.6.1 (R Foundation for Statistical Computing) with the tidyverse, lme4, emmeans, and mice packages.

Results

Participants

Four hundred and nineteen participants underwent TAA, and 103 underwent AA. Five TAA patients withdrew immediately after surgery, leaving 517 to complete the baseline assessment. Postoperative scores were obtained in 495 (96%), 481 (93%), 409 (79%), and 425 (82%) patients at 12, 24, 36, and 48 months, respectively, with 348 (84%) in the TAA group and 77 (75%) in the AA group reaching 48 months (Fig. 2).

Risk Factors

Prior to treatment, the 2 groups were not different with regard to sex, race, marital status, education, severity of osteoarthritis, alignment, osteoporosis, degenerative disc disease, or alcohol use (all $p \geq 0.08$; Table I), and there was no difference in the baseline outcome measures by group (Table II). However, patients who underwent TAA were older and weighed less. They were less likely to be employed, had higher incomes, and were less likely to have posttraumatic arthritis or have had previous ankle surgery. Also, they were less likely to have had depression/and or anxiety and had lower functional comorbidity scores (Functional Comorbidity Index, FCI), and they were less likely to smoke. These variables were accounted for in the regression models as described above.

Outcomes

Both groups achieved significant improvement in the 2 FAAM measures, the SF-36 PCS, and all of the pain measures ($p <$

0.001) (Fig. 3, Table II). Compared with AA, those undergoing TAA had greater improvement from baseline in all of these measures ($p \leq 0.004$) except for present pain ($p = 1.0$). In the arthroplasty patients, mean improvements (and standard deviation) from baseline for FAAM Activities of Daily Living, FAAM Sports, and SF-36 PCS scores were at least 9 ± 2 points, 8 ± 3 points, and 3.5 ± 1.1 points, respectively, which were higher than in those undergoing AA at all postoperative study visits. Mean improvements in worst and average pain were at least 0.9 ± 0.3 point higher for arthroplasty than for arthrodesis patients at 12, 24, and 36 months, but these differences were attenuated by 48 months.

Mean improvement from baseline in SF-36 MCS scores was <2 points for both groups at all of the follow-ups, and there was no difference in improvement based on treatment type ($p = 0.71$). For both treatments, all improvements from baseline to 24 months were maintained at 48 months. Mean changes between 24 and 48 months for either surgery were within 2.5 points for the FAAM measures, within 1.3 points for the SF-36 measures, and within 0.3 point for the pain measures. The propensity score analysis yielded findings similar to those presented in Table II. For example, compared with AA, the patients who underwent TAA had a greater mean 48-month improvement in FAAM Activities of Daily Living of 12 points (95% CI, 5 to 19) and in average pain of 0.9 point (95% CI, 0.0 to 1.8) when the propensity score analysis was used. Models applied to 5 data sets with imputed missing outcomes yielded findings similar in magnitude (data not shown). Estimates for the 48-month mean improvement in the FAAM Activities of Daily Living for patients who underwent TAA and those who underwent AA ranged from 32.9 to 33.8 and 22.3 to 24.1, respectively, with the mean difference in improvement between the 2 groups ranging from 8.8 to 11.5. Differences in the mean 48-month improvement in average pain by group ranged from 0.5 to 0.7. Results from both the propensity score analysis and the multiple imputation are summarized in Figure 4.

Of those with 48 months of follow-up, 78% of arthroplasty patients reported that they were completely satisfied with the results of their surgery versus 60% of arthrodesis patients ($p = 0.003$). More than 96% of both groups reported that they would repeat the surgery, and expectations had been met for 87% and 85% of arthroplasty and arthrodesis patients, respectively. There were 36 (8.7%) revisions in the TAA group and 18 (17.5%) revisions in the AA group. These participants were not excluded from the analysis.

Discussion

This study demonstrates that patients who have undergone TAA or AA for end-stage ankle arthritis have improved significantly in overall function, ankle-specific function, and pain at 48 months after surgery. Improvements that had been achieved at 24 months were maintained through 48 months. Most patients were completely satisfied with their results, would repeat the surgery again, and had had their expectations met. The improvements were significantly greater after TAA in nearly every patient-reported measure. These differences are clinically important, particularly in the ankle-specific measures. Furthermore, 28% more patients who underwent TAA were completely satisfied than those who underwent AA. At the 48-month follow-up, revision rates were 8.7% and 17.5% in the TAA and AA groups, respectively. Future analyses will include a detailed time-to-event evaluation of comparative revision rates and risk factors that are associated with revision.

The interpretation of these findings is highly relevant to patients who experience end-stage ankle arthritis. During the planning phase, we collaborated with the American Orthopaedic Foot & Ankle Society (AOFAS) using the McMaster Toronto Arthritis (MACTAR) patient-preference disability questionnaire to evaluate patient outcome preferences in 235 foot and ankle patients[26]. After review of dozens of measures, we concluded that the FAAM[17,18] best represented the outcomes that were most important to patients, provided more granular evaluations of each of these general preferences (e.g., walking on all types of surfaces, walking short distances, walking long distances), and met psychometric requirements, including excellent responsiveness and having an achievable MCID.

To our knowledge, only 2 prospective cohort studies comparing TAA with AA and evaluating outcomes at ≥4 years after surgery have been published. Daniels et al. reported improvements in Ankle Osteoarthritis Scale (AOS) scores in both groups between baseline and the 4-year follow-up, but neither reached the MCID[27]. There were no significant differences in AOS or SF-36 scores between the groups, but there was a higher rate of revision in the TAA group. Wąsik et al. similarly reported improvement in both groups, with little difference found in pain, AOFAS scores, or health-related quality of life scores[23]. However, the outcome measures that were used in those studies have come under scrutiny. The AOFAS measure has an objective component that makes it difficult to perform and has not been shown to be reliable[28].

Between 1995 and 2004, SooHoo et al. reported higher revision rates in patients treated by TAA compared with AA[29]. However, when the same data source was evaluated using 2005 to 2010 data, the short-term complication risk was low for both procedures, and patients who underwent TAA had significantly lower rates of readmission ($p < 0.0001$) and periprosthetic joint infection or wound infection ($p = 0.02$) compared with patients who underwent AA[30]. Saltzman et al. reported a higher rate of reoperations after TAA compared with AA[11]. Their study included a "continuing access group," with longer enrollment and follow-up of the TAA group alone; the rate of adverse events in the TAA group declined by 40%. In spite of limited long-term follow-up data comparing the safety and efficacy of these 2 procedures, there has been substantial growth in the number of TAA procedures[31].

This study has limitations. It was initially designed as an RCT with a patient-preference cohort; however, an insufficient

number of patients volunteered for the randomization arm, despite a standardized explanation by the research staff that neither procedure was superior. Therefore, results regarding the comparative effectiveness of these 2 treatments were subject to bias due to confounding. Some of the baseline differences do suggest that there was some selection bias. To account and control for these baseline differences, we paid careful attention to potential confounding due to selection bias through a rigorous evaluation of each unevenly distributed characteristic and its association with each outcome. This process led to our final multivariable models. As a sensitivity analysis, we employed propensity scoring, which rendered equivalent results. The selection into groups by the surgeons may have indicated a bias; however, the results of the analysis accounting for these differences, both in multivariable modeling and in propensity scoring, suggest that selection bias did not influence the observed treatment effect. This provided additional confidence that the comparative improvements were not likely to be confounded by specific baseline characteristics, and that the findings can be applied to most patients who are surgical candidates for end-stage ankle arthritis. Additionally, a multiple imputation analysis was performed to account for loss to follow-up; this also demonstrated equivalent results. Data were collected with the integrity of an RCT and under the rigorous supervision of a data safety monitoring board. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Observational Studies Task Force published the following statement: "It is the position of this Task Force that rigorous well-designed and well-executed observational studies can provide evidence of causal relationships."[32] Several reputable journals have published reviews comparing RCTs to observational studies, which found similar results for the 2 study types in several medical specialty areas[33-35]. A previous study found that the survival rates after TAA were substantially higher after the first 30 procedures had been performed by the surgeon[16]. To minimize potential treatment outcome bias, we included only surgeons with experience in both treatments. Surgeons were required to have had a minimum of 5 years of experience and to have performed at least 30 TAA and 30 AA procedures and a minimum of 100 joint replacements in weight-bearing joints to ensure that the surgeons understood both the principles of joint replacement and the specifics of ankle replacement.

In conclusion, both established treatments for end-stage ankle arthritis are effective at pain relief and improved patient-reported outcomes; however, it appears that TAA leads to greater improvement in most patient-reported outcome measures at 48 months after surgery. ■

NOTE: Dr. Sangeorzan, Dr. Norvell, and Jane Shofer had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The study team thanks the research coordinators who were responsible for obtaining the follow-up patient-reported outcomes for this study: Lyndsey Behrend, Amber Curtis, Jenn Hicks, Michelle Padley, Jacky Shelton, Becky Stone, and Erin Zimmerman. We also thank Ian Ellis who was responsible for data management.

Bruce J. Sangeorzan, MD[1,2]
William R. Ledoux, PhD[1,2]
Jane B. Shofer, MS[1]
James Davitt, MD[3]
John G. Anderson, MD[4]
Donald Bohay, MD[4]
J. Chris Coetzee, MD[5]
John Maskill, MD[4]
Michael Brage, MD[2]
Daniel C. Norvell, PhD[1]

[1]VA Puget Sound Health Care System, Seattle, Washington

[2]Departments of Orthopaedics and Sports Medicine (B.J.S., W.R.L., and M.B.) and Mechanical Engineering (W.R.L.), University of Washington, Seattle, Washington

[3]Orthopedic + Fracture Specialists, Portland, Oregon

[4]Orthopaedic Associates of Michigan, Grand Rapids, Michigan

[5]Twin Cities Orthopedics, Edina, Minnesota

Email address for D.C. Norvell: daniel.norvell@va.gov

ORCID iD for B.J. Sangeorzan: 0000-0002-4665-2579
ORCID iD for W.R. Ledoux: 0000-0003-4982-7714
ORCID iD for J.B. Shofer: 0000-0001-6354-2817
ORCID iD for J. Davitt: 0000-0001-6309-5308
ORCID iD for J.G. Anderson: 0000-0001-7877-9972
ORCID iD for D. Bohay: 0000-0001-6273-4833
ORCID iD for J C. Coetzee: 0000-0001-6822-9512
ORCID iD for J. Maskill: 0000-0002-7221-3342
ORCID iD for M. Brage: 0000-0002-6440-4747
ORCID iD for D.C. Norvell: 0000-0002-4386-2677

References

1. Vakhshori V, Sabour AF, Alluri RK, Hatch GF 3rd, Tan EW. Patient and practice trends in total ankle replacement and tibiotalar arthrodesis in the United States from 2007 to 2013. J Am Acad Orthop Surg. 2019 Jan 15;27(2):e77-84.

2. Glazebrook M, Daniels T, Younger A, Foote CJ, Penner M, Wing K, Lau J, Leighton R, Dunbar M. Comparison of health-related quality of life between patients with end-stage ankle and hip arthrosis. J Bone Joint Surg Am. 2008 Mar;90(3):499-505.

3. Haddad SL, Coetzee JC, Estok R, Fahrbach K, Banel D, Nalysnyk L. Intermediate and long-term outcomes of total ankle arthroplasty and ankle arthrodesis. A systematic review of the literature. J Bone Joint Surg Am. 2007 Sep;89(9):1899-905.

4. Guyer AJ, Richardson G. Current concepts review: total ankle arthroplasty. Foot Ankle Int. 2008 Feb;29(2):256-64.

5. Gougoulias N, Khanna A, Maffulli N. How successful are current ankle replacements?: a systematic review of the literature. Clin Orthop Relat Res. 2010 Jan; 468(1):199-208. Epub 2009 Jul 18.

6. Agel J, Coetzee JC, Sangeorzan BJ, Roberts MM, Hansen ST Jr. Functional limitations of patients with end-stage ankle arthrosis. Foot Ankle Int. 2005 Jul;26(7): 537-9. Erratum in: Foot Ankle Int. 2005 Sep;26(9):table of contents

7. Benich MR, Ledoux WR, Orendurff MS, Shofer JB, Hansen ST, Davitt J, Anderson JG, Bohay D, Coetzee JC, Maskill J, Brage M, Houghton M, Sangeorzan BJ. Comparison of treatment outcomes of arthrodesis and two generations of ankle replacement implants. J Bone Joint Surg Am. 2017 Nov 1;99(21):1792-800.

8. Esparragoza L, Vidal C, Vaquero J. Comparative study of the quality of life between arthrodesis and total arthroplasty substitution of the ankle. J Foot Ankle Surg. 2011 Jul-Aug;50(4):383-7. Epub 2011 May 4.

9. Norvell DC, Ledoux WR, Shofer JB, Hansen ST, Davitt J, Anderson JG, Bohay D, Coetzee JC, Maskill J, Brage M, Houghton M, Sangeorzan BJ. Effectiveness and safety of ankle arthrodesis versus arthroplasty: a prospective multicenter study. J Bone Joint Surg Am. 2019 Aug 21;101(16):1485-94.

10. Norvell DC, Shofer JB, Hansen ST, Davitt J, Anderson JG, Bohay D, Coetzee JC, Maskill J, Brage M, Houghton M, Ledoux WR, Sangeorzan BJ. Frequency and impact of adverse events in patients undergoing surgery for end-stage ankle arthritis. Foot Ankle Int. 2018 Sep;39(9):1028-38. Epub 2018 May 31.

11. Saltzman CL, Mann RA, Ahrens JE, Amendola A, Anderson RB, Berlet GC, Brodsky JW, Chou LB, Clanton TO, Deland JT, Deorio JK, Horton GA, Lee TH, Mann JA, Nunley JA, Thordarson DB, Walling AK, Wapner KL, Coughlin MJ. Prospective controlled trial of STAR total ankle replacement versus ankle fusion: initial results. Foot Ankle Int. 2009 Jul;30(7):579-96.

12. Shofer JB, Ledoux WR, Orendurff MS, Hansen ST, Davitt J, Anderson JG, Bohay D, Coetzee JC, Houghton M, Norvell DC, Sangeorzan BJ. Step activity after surgical treatment of ankle arthritis. J Bone Joint Surg Am. 2019 Jul 3;101(13):1177-84.

13. Singer S, Klejman S, Pinsker E, Houck J, Daniels T. Ankle arthroplasty and ankle arthrodesis: gait analysis compared with normal controls. J Bone Joint Surg Am. 2013 Dec 18;95(24):e191.1-10.

14. Slobogean GP, Younger A, Apostle KL, Marra CA, Wing K, Penner M, Daniels T, Glazebrook M. Preference-based quality of life of end-stage ankle arthritis treated with arthroplasty or arthrodesis. Foot Ankle Int. 2010 Jul;31(7):563-6.

15. Kim HJ, Suh DH, Yang JH, Lee JW, Kim HJ, Ahn HS, Han SW, Choi GW. Total ankle arthroplasty versus ankle arthrodesis for the treatment of end-stage ankle arthritis: a meta-analysis of comparative studies. Int Orthop. 2017 Jan;41(1):101-9. Epub 2016 Oct 7.

16. Henricson A, Skoog A, Carlsson A. The Swedish Ankle Arthroplasty Register: an analysis of 531 arthroplasties between 1993 and 2005. Acta Orthop. 2007 Oct; 78(5):569-74.

17. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). Foot Ankle Int. 2005 Nov; 26(11):968-83.

18. Kivlan BR, Martin RL, Wukich DK. Responsiveness of the Foot and Ankle Ability Measure (FAAM) in individuals with diabetes. Foot. 2011 Jun;21(2):84-7. Epub 2011 May 7.

19. Smith BH, Penny KI, Purves AM, Munro C, Wilson B, Grimshaw J, Chambers WA, Smith WC. The Chronic Pain Grade questionnaire: validation and reliability in postal research. Pain. 1997 Jun;71(2):141-7.

20. Mancuso CA, Salvati EA, Johanson NA, Peterson MG, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. J Arthroplasty. 1997 Jun; 12(4):387-96.

21. Robertsson O, Dunbar M, Pehrsson T, Knutson K, Lidgren L. Patient satisfaction after knee arthroplasty: a report on 27,372 knees operated on between 1981 and 1995 in Sweden. Acta Orthop Scand. 2000 Jun;71(3):262-7.

22. Robertsson O, Dunbar MJ. Patient satisfaction compared with general health and disease-specific questionnaires in knee arthroplasty patients. J Arthroplasty. 2001 Jun;16(4):476-82.

23. Wasik J, Stołtny T, Pasek J, Szyluk K, Pyda M, Ostałowska A, Kasperczyk S, Koczy B. Effect of total ankle arthroplasty and ankle arthrodesis for ankle osteoarthritis: a comparative study. Med Sci Monit. 2019 Sep 10;25:6797-804.

24. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015 Dec 10;34(28): 3661-79. Epub 2015 Aug 3.

25. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputations by chained equations in R. J Stat Softw. 2011;45(3):1-67.

26. Sangeorzan BJ, Beskin JL, Britt SM, Brage M, Holt S, Kadel NJ, Legro MW, Sands AK. SooHoo NF, Stroud CC; Outcomes Committee of the AOFAS. Issues of importance to patients seeking care from members of the AOFAS: a preliminary report of the Outcomes Committee of the AOFAS. Foot Ankle Int. 2005 Aug;26(8):638-44.

27. Daniels TR, Younger AS, Penner M, Wing K, Dryde PJ, Wong H, Glazebrook M. Intermediate-term results of total ankle replacement and ankle arthrodesis: a COFAS multicenter study. J Bone Joint Surg Am. 2014 Jan 15;96(2):135-42.

28. Madeley NJ, Wing KJ, Topliss C, Penner MJ, Glazebrook MA, Younger AS. Responsiveness and validity of the SF-36, Ankle Osteoarthritis Scale, AOFAS Ankle Hindfoot Score, and Foot Function Index in end stage ankle arthritis. Foot Ankle Int. 2012 Jan;33(1):57-63.

29. SooHoo NF. Zingmond DS, Ko CY. Comparison of reoperation rates following ankle arthrodesis and total ankle arthroplasty. J Bone Joint Surg Am. 2007 Oct; 89(10):2143-9.

30. Stavrakis AI. SooHoo NF. Trends in complication rates following ankle arthrodesis and total ankle replacement. J Bone Joint Surg Am. 2016 Sep 7;98(17): 1453-8.

31. Pugely AJ, Lu X, Amendola A, Callaghan JJ, Martin CT, Cram P. Trends in the use of total ankle replacement and ankle arthrodesis in the United States Medicare population. Foot Ankle Int. 2014 Mar;35(3):207-15. Epub 2013 Oct 31.

32. Berger ML, Dreyer N, Anderson F, Towse A, Sedrakyan A, Normand SL. Prospective observational studies to assess comparative effectiveness: the ISPOR Good Research Practices Task Force report. Value Health. 2012 Mar-Apr;15(2): 217-30.

33. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med. 2000 Jun 22;342(25):1878-86.

34. Concato J, Lawler EV, Lew RA, Gaziano JM, Aslan M, Huang GD. Observational methods in comparative effectiveness research. Am J Med. 2010 Dec;123(12): e16-23.

35. Kitsios GD, Dahabreh IJ, Abu Dabrh AM, Thaler DE, Kent DM. Patent foramen ovale closure and medical treatments for secondary stroke prevention: a systematic review of observational and randomized evidence. Stroke. 2012 Feb;43(2):422-31. Epub 2011 Dec 15.