


The Norwegian Cervical Arthroplasty Trial (NORCAT): 2-year clinical outcome after single-level cervical arthroplasty versus fusion—a prospective, single-blinded, randomized, controlled multicenter study

Jarle Sundseth^{1,2}  · Oddrun Anita Fredrikli^{3,4,5} · Frode Kolstad¹ · Lars Gunnar Johnsen^{4,5,6} · Are Hugo Pripp⁷ · Hege Andresen^{4,5} · Erling Myrseth⁸ · Kay Müller⁹ · Øystein P. Nygaard^{3,4,5} · John-Anker Zwart^{2,10} · On behalf of the NORCAT study group

Received: 2 September 2016/Revised: 5 December 2016/Accepted: 13 December 2016/Published online: 23 December 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Purpose Standard surgical treatment for symptomatic cervical disc disease has been discectomy and fusion, but the use of arthroplasty, designed to preserve motion, has increased, and most studies report clinical outcome in its favor. Few of these trials, however, blinded the patients. We, therefore, conducted the Norwegian Cervical Arthroplasty Trial, and present 2-year clinical outcome after arthroplasty or fusion.

Methods This multicenter trial included 136 patients with single-level cervical disc disease. The patients were randomized to arthroplasty or fusion, and blinded to the treatment modality. The surgical team was blinded to randomization until nerve root decompression was completed. Primary outcome was the self-rated Neck Disability Index. Secondary outcomes were the numeric rating scale

for pain and quality of life questionnaires Short Form-36 and EuroQol-5Dimension-3 Level.

Results There was a significant improvement in the primary and all secondary outcomes from baseline to 2-year follow-up for both arthroplasty and fusion ($P < 0.001$), and no observed significant between-group differences at any follow-up times. However, linear mixed model analyses, correcting for baseline values, dropouts and missing data, revealed a difference in Neck Disability Index ($P = 0.049$), and arm pain ($P = 0.027$) in favor of fusion at 2 years. The duration of surgery was longer ($P < 0.001$), and the frequency of reoperations higher ($P = 0.029$) with arthroplasty.

Conclusion The present study showed excellent clinical results and no significant difference between treatments at any scheduled follow-up. However, the rate of index level reoperations was higher and the duration of surgery longer with arthroplasty.

Trial registration <http://www.clinicaltrials.gov> NCT 0073 5176.19.

Ø. P. Nygaard and J.-A. Zwart contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s00586-016-4922-5) contains supplementary material, which is available to authorized users.

✉ Jarle Sundseth
jarle.sundseth@rikshospitalet.no

✉ John-Anker Zwart
j.a.zwart@medisin.uio.no

¹ Department of Neurosurgery, Oslo University Hospital Rikshospitalet, Postboks 4950 Nydalen, 0424 Oslo, Norway

² Faculty of Medicine, University of Oslo, Oslo, Norway

³ Department of Neurosurgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

⁴ National Advisory Unit on Spinal Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

⁵ Department of Neuroscience, Faculty of Medicine, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

⁶ Department of Orthopaedic Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

⁷ Oslo Centre of Biostatistics and Epidemiology, Research Support Services, Oslo University Hospital, Oslo, Norway

⁸ Department of Neurosurgery, Haukeland University Hospital, Bergen, Norway

⁹ Department of Neurosurgery, University Hospital of North Norway, Tromsø, Norway

¹⁰ Department of Neurology and FORMI, Oslo University Hospital Ullevaal, Oslo, Norway

Keywords Arthroplasty · Fusion · Cervical · Outcome

Introduction

Cervical radiculopathy is caused by encroachment of a cervical nerve root, usually by bone or disc herniation. Typical symptoms are neck pain radiating into the arm, possible loss of motor function and/or sensory loss [1]. The most common surgical treatment is discectomy and fusion [2] either as stand-alone implant surgery, or with the addition of anterior plating [3]. Anterior cervical discectomy is one of the most frequently performed spinal procedures. In the US, almost 550,000 patients were operated on between 2005 and 2008 [4]. Concern that fusion may cause adjacent segment disease [5] has given rise to motion preserving implants (arthroplasty). In the US, cervical arthroplasty surgery increased by 708% between 2005 and 2008 [4].

Multiple trials [6–14] and three recent meta-analyses [15–17] have compared the results of arthroplasty versus fusion. Most authors concluded with clinical outcome in favor of arthroplasty [6–12, 15–17]. However, few trials included blinding of patients [15], and blinding was only performed until just after the surgical procedure was completed [7, 11]. Only one study implemented blinding of the surgical team [14]. So far, no studies have demonstrated clinical outcome in favor of fusion.

The aim of the Norwegian Cervical Arthroplasty Trial (NORCAT) was to assess 2-year clinical outcome in patients operated for single-level cervical radiculopathy with either arthroplasty or fusion.

Methods

Study design

Patients with single-level radiculopathy were included from November 2008 to January 2013 at five neurosurgical departments in Norway. The surgical procedure was either arthroplasty or fusion. The randomization was stratified according to center, and blocked using the Unit of Applied Clinical Research website (<http://www.ntnu.edu/dmf/akf/randomisering>), to ensure equality in the groups. The study was designed to include 146 patients. Follow-up visits were scheduled at 3 months, 1 and 2 years. At 6 months, the patients answered the questionnaires by mail. Participating patients were blinded to the treatment until the last follow-up was completed.

The NORCAT received a grant from DePuy Synthes Spine (325 Paramount Drive Raynham, MA 02767). However, the sponsor was not involved in study design,

conducting the trial, writing or reviewing the manuscript. The grant was unrestricted and the sponsor had no right of refusal for publication of the data. The sponsor read the manuscript before submission.

Participants

Inclusion criteria were: age between 25 and 60 years, clinical C6 or C7 radiculopathy with corresponding radiological findings, Neck Disability Index (NDI) [18] $\geq 30\%$, no response to non-operative treatment, and no clinical improvement during the six weeks prior to surgery. Exclusion criteria were: significant spondylosis involving more than one level, adjacent level ankylosis, intramedullary changes on magnetic resonance imaging (MRI), and myelopathy. The complete list of inclusion and exclusion criteria is available in Table S1 in the supplementary appendix.

Study interventions

Discectomy via anterolateral approach was performed. The surgical team was blinded to the result of randomization until nerve root decompression was completed. Both arthroplasty [DISCOVER[®] prosthesis (DePuy Spine Inc., 325 Paramount Dr, Raynham, MA 02767, USA)], and fusion [CERVIOS[®] cage (Synthes GmbH, Eimattstrasse 3, 4436 Oberdorf, Switzerland)] implant systems were available in the operating theater.

Arthroplasty

The DISCOVER[®] prosthesis allows for unconstrained motion. Two titan plates are fixed to the endplates with a polyethylene inlay. Fluoroscopy was used to ensure that the prosthesis was placed in the midline and sufficiently towards the posterior edge of the vertebra. The appropriate size of implants was determined with templates.

Fusion

The CERVIOS[®] cage was used to achieve anterior cervical interbody fusion. The cage was preloaded with chronOS and the procedure was performed as stand-alone surgery.

Outcome measures

Primary outcome

The NDI is a self-rated questionnaire developed for patients with neck disability. The questionnaire is composed of ten items: seven related to activities of daily living (personal care, lifting, reading, work/daily activities,

driving, sleep, and recreation), two to pain (pain, headache), and one to concentration. Each item is rated from 0 to 5. The NDI summary score ranges from 0 to 50. We expressed the score as a percentage with lower scores indicating less severe symptoms. We used the validated Norwegian version [19].

Secondary outcomes

Secondary outcome measures were the Numeric Rating Scale (NRS 11) [20], the Short Form 36 (SF-36) [21], and the EuroQol-5 Dimension-3 Level (EQ-5D-3L) [22]. In addition, data regarding the surgical procedure (duration of surgery, duration of anesthesia, and total blood loss), perioperative major complications (dural tear, damage of n. laryngeus recurrens, index level nerve, esophagus, trachea or large vessel), Dysphagia Short Questionnaire [23], reoperations at index level within 2 years, and work status were recorded.

NRS 11 is a one-dimensional pain scale from 0 (“no pain at all”) to 10 (“worst imaginable pain”), used to evaluate arm and neck pain.

SF-36 is a generic questionnaire measuring health-related quality of life along eight dimensions (physical function, role limitations due to physical problems, bodily pain, general health, vitality, social function, role limitations due to emotional problems, and mental health) with two summary scores (physical component summary [PCS], mental component summary [MCS]). The score ranges from 0 to 100, with higher scores relating to better health. We used the validated Norwegian (chronic) version 2.0 [24].

The EQ-5D-3L is a generic quality of life questionnaire with five dimensions (mobility, self-care, activities of daily life, pain, and anxiety/depression), ranging from -0.59 to 1. Higher scores indicate better health status. We used the validated Norwegian version [25], and syntax files obtained from the EQ-5D society using the UK time trade off tariff to calculate the utility index [26].

The Dysphagia Short Questionnaire consists of five items (ability to swallow, incorrect swallowing, globus sensation, involuntary weight loss, and pneumonia), with scores ranging from 0 to 18. Lower scores represent milder symptoms.

Statistical analyses

The trial was planned to have 80% power to detect a difference of 10/100 in NDI score, considered to be the minimal level required for clinically important change [27, 28]. On the basis of a significance level of 0.05 and a standard deviation of 18, 104 participants were required for the trial. Correcting for 40% lost to follow-up gave a total

of 146 participants. A P value of <0.05 was used as a level of significance. PASW (Predictive Analytics Software) Statistics 18 (IBM Corporation, Armonk, New York, USA) was used for all analysis.

Outcomes were analyzed according to the intention-to-treat principle. Continuous data are described as means and standard deviations (SD), or medians and interquartile ranges (IQR), as appropriate, and were statistically tested between the groups with independent t test or Mann–Whitney U test depending on assumptions on statistical distribution. Ninety-five percent confidence intervals (CI) are specified in Figure illustrations and Tables for the outcome measures. Categorical data are described as number of patients and percentages, tested with χ^2 test or Fischer’s exact test, as appropriate.

To assess change in outcome from baseline to each follow-up time-point, paired samples t tests were used for parametric data, and Wilcoxon signed rank tests for non-parametric data.

The repeated measurements after intervention were analyzed using linear mixed models with a random intercept adjusted for baseline score. Follow-up time-points, treatment modality and baseline score were included as fixed main effects together with interaction terms between follow-up time-points and treatment modality. The mean differences between treatment modalities with 95% CI at each follow-up time-point were estimated using linear combinations of estimators. The linear mixed model analysis was not described in the original study protocol, but applied due to a difference in NDI scores between the treatment modalities at baseline. A sensitivity analysis including seven patients who were randomized and excluded from the trial was performed based on intention-to-treat principle with extreme values (best possible score) for all outcome measures.

Possible effect or difference between the five neurosurgical departments was also evaluated, but neither the statistical assessment nor the trial design indicated that any multicenter effect should be taken into account in our final statistical analysis.

Ethical considerations

The trial was approved by the Regional Committee for Medical and Health Research Ethics in Central Norway, and the Data Protection Official for Research. All enrolled patients gave their written informed consent. Participating senior surgeons at each hospital performed all operations. The accuracy of the study to the protocol was vouched for by all authors, and it was a unanimous agreement to submit the final manuscript for publication.

Trial registration <http://www.clinicaltrials.gov> NCT00735176.19.

Results

Patient characteristics

During the study period, 3922 patients attended specialist outpatient clinics for cervical radiculopathy at the study sites, 143 of which were eligible for inclusion. Seven patients were excluded, leaving 136 patients for inclusion, 68 in each group (Fig. 1). Of these, 120 attended the 2-year follow-up and returned the questionnaires, giving a dropout rate of 11.8%. The groups were well matched with respect to demographic and clinical characteristics at baseline (Table 1).

Primary and secondary outcomes

Change in outcome from baseline to each follow-up time-point revealed no significant difference between the NDI scores of the two treatment modalities, $P = 0.25$ (Fig. 2; Table 2), and no differences in secondary outcome measures [Table 3; Fig. 3 (supplementary appendix)] after 2 years.

Both procedures demonstrated a statistically significant improvement in NDI score from baseline to 3 months, 6 months, 1 and 2 years, $P < 0.001$ (Fig. 2). For arthroplasty, the mean reduction in NDI score was from 45.7 (95% CI 42.9–48.6) at baseline to 25.0 (95% CI 20.1–29.9) after 2 years, $P < 0.001$, and for fusion from 51.2 (95% CI 48.0–54.4) at baseline to 21.2 (95% CI 16.7–25.6) after 2 years, $P < 0.001$. The improvement from baseline to each follow-up time was also statistically significant for all secondary outcomes, $P < 0.001$ (Fig. 3; supplementary appendix).

From three months, there was no significant change in NDI score for either arthroplasty or fusion ($P = 0.20$). The proportion of patients reaching the minimal clinically important change of ten or more improvement in NDI score from baseline, was 70.0% ($n = 42$) for arthroplasty, and 78.3% ($n = 47$) for fusion, $P = 0.30$.

Statistical analysis using linear mixed models for repeated measurements that correct for baseline differences, dropouts and missing data demonstrated a mean difference in NDI score of 5.9% in favor of fusion after 2 years, $P = 0.049$ (Table 2). Figure 4 in the supplementary appendix shows a plot for NDI where the two groups are compared based on estimated results from the statistical model. For the secondary outcomes, there was a mean difference in NRS arm pain of 1.0 in favor of fusion after 2 years, $P = 0.03$ (Table 3).

The surgical procedure was significantly longer with arthroplasty, $P < 0.001$. There were no major complications in either group, and no difference in dysphagia score

Fig. 1 Eligibility, randomisation, and follow-up of the patients. ^aOne patient withdrew consent before surgery, and one patient's MRI was too old, but he/she was unable to undertake a new preoperative MRI due to claustrophobia. One patient had a short neck, which made visualization of the relevant level C6/C7 with fluoroscopy impossible, and for another, the prostheses were not available in the operating theater at time of surgery due to a misunderstanding. In the last patient, the surgeons had to convert to fusion with anterior plating fixation due to instability. ^bShort necks made visualization of the relevant level C6/C7 with fluoroscopy impossible. ^cOne patient did not receive allocated intervention due to problems with positioning of the arthroplasty device, resulting in conversion to fusion. ^dFour patients in each group did not attend the follow-up and did not return the questionnaires. ^eEleven patients in the arthroplasty group and 14 patients in the fusion group did not return the questionnaires. ^fFive patients in the arthroplasty group and six patients in the fusion group did not attend the follow-up and did not return the questionnaires. ^gTwo patients did not attend the follow-up and did not return the questionnaires, one patient attended the follow-up without returning the questionnaires, and 1 patient had undergone brain surgery that resulted in postoperative problems with hand writing. ^hOne patients did not attend the follow-up and did not return the questionnaires

in the 2 years following treatment (Table S2, supplementary appendix).

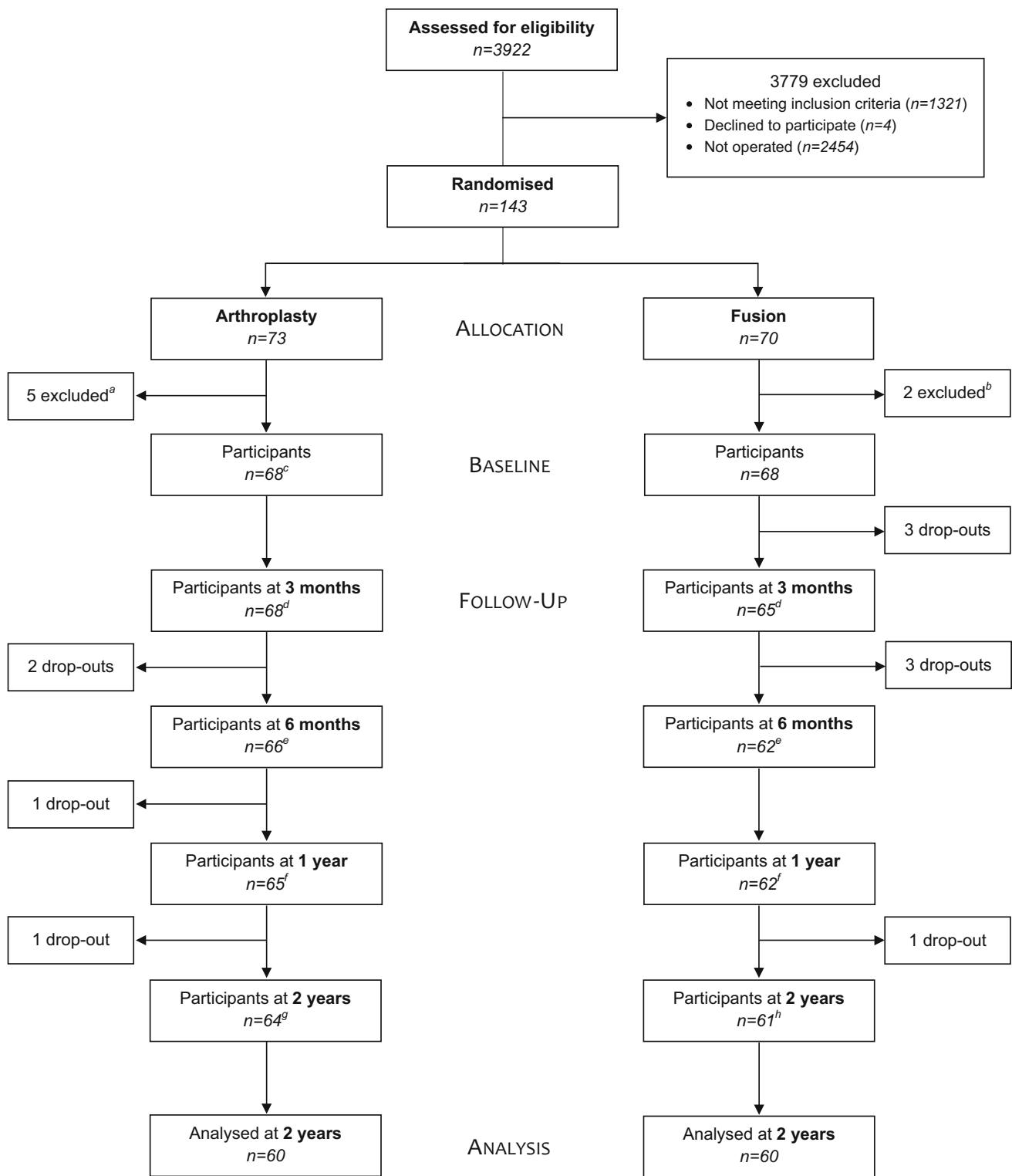
Two years after surgery, one patient in the fusion group and eight in the arthroplasty group had undergone index level reoperation, $P = 0.03$ (Table S2, supplementary appendix). One patient in the fusion group and five in the arthroplasty group were reoperated due to index level uncovertebral restenosis. Decompression was performed with a posterior foraminotomy, leaving the implants in place. Three patients in the arthroplasty group were reoperated because of migration and anterior displacement of the implant with suspected instability and secondary neck pain. Revision surgeries were performed with removal of the prostheses, cage implantation and/or anterior plating.

The median duration of sick leave after surgery was 10 weeks (IQR 6–27) with arthroplasty, and 12 weeks with fusion (IQR 6–30), $P = 0.17$. After 2 years, 59.7% in the arthroplasty group, and 71.7% in the fusion group had resumed work, $P = 0.16$.

A sensitivity analysis including the seven patients who were randomized but excluded from the trial was performed based on an intention-to-treat principle with extreme values (best possible scores) for all outcomes. Independent t tests revealed no differences between the groups up to 2 years.

Discussion

We found excellent clinical results for both treatment modalities at 3 months, which were sustained at 2 years. There was no significant difference between arthroplasty



and fusion at any of the follow-up times. However, statistical analyses using linear mixed models that adjust for baseline values, dropout and missing data showed a difference in self-rated neck disability and the numeric rating score for arm pain in favor of fusion after 2 years.

This is not consistent with most randomized controlled trials [6–12], the recent study on available registry data by Staub and colleagues [29], and three recent meta-analyses [15–17] reporting clinical outcome in favor of arthroplasty.

Table 1 Baseline characteristics of the study participants operated for single-level cervical radiculopathy with either arthroplasty or fusion

Characteristics	Arthroplasty (<i>n</i> = 68)	Fusion (<i>n</i> = 68)
Mean (SD) age (years)	44.7 (7.2)	43.4 (6.8)
Women	36 (52.9)	37 (54.4)
Mean (SD) height (cm)	174.1 (10.6)	172.7 (8.9)
Mean (SD) weight (kg)	79.1 (14.6)	76.8 (15.8)
Mean (SD) body mass index ^a	26.0 (3.7)	25.5 (3.6)
Level operated C5/C6	38 (55.9)	36 (52.9)
Median (IQR) duration of sick leave (weeks) ^b	21 (6–39)	24 (1–55)
Duration of arm pain ^c		
<3 months	3 (4.5)	6 (9.1)
3 months–1 year	35 (53.0)	30 (45.1)
1–2 years	14 (21.2)	20 (30.3)
>2 years	14 (21.2)	10 (15.1)
Duration of neck pain ^c		
No neck pain	3 (4.5)	2 (3.0)
<3 months	4 (6.1)	3 (4.5)
3 months–1 year	27 (40.9)	28 (41.8)
1–2 years	11 (16.7)	19 (28.4)
>2 years	21 (31.8)	15 (22.4)
Working until surgery	14 (20.6)	17 (25.0)
Higher education ^d	28 (41.2)	26 (38.2)
Smoking ^e	23/66 (34.9)	29/61 (47.5)
Married or cohabitating	59 (86.8)	50 (73.5)
Co-morbidity		
Heart disease	1 (1.5)	1 (1.5)
Hypertension	6 (8.8)	2 (2.9)
Diabetes	4 (6.1)	3 (4.4)
Mean (95% CI) NDI ^f	45.7 (42.9–48.6)	51.2 (48.0–54.4)

Figures are number (percentage) unless stated otherwise

^a The body mass index is weight in kilograms divided by the square of the height in meters

^b Prior to surgical treatment

^c Number of patients with arm and neck pain in the specified period before surgery. Missing data regarding arm pain in four patients (two in each group) and regarding neck pain in three patients (two in the arthroplasty group and one in the fusion group)

^d College or university

^e Missing data for nine patients (two in the arthroplasty group and seven in the fusion group)

^f *NDI* Neck Disability Index (0–100%, lower scores indicate less severe symptoms)

The between-group difference in NDI score of 5.9%, shown in the present study is small and the statistical significance is weak, and the results must, therefore, be interpreted with caution. One might argue that the difference should not be considered clinically important, but there is no clear consensus-based agreement on how large the between-group difference should be [30, 31]. There were 78.3% in the fusion group and 70.0% in the arthroplasty group reporting an NDI change of 10 or more from baseline to 2-year follow-up. Even though the difference was not statistically significant, the direction did not favor arthroplasty. There may be several reasons for the discrepancy compared with previous studies, such as different

implant design, different study methods, different fusion technique, different lengths of follow-up, and the impact of funding by arthroplasty manufacturers.

Different arthroplasty designs have revealed different biomechanical performances for the treatment of single-level cervical disc disease [32]. Arthroplasty devices are considered constrained in certain planes if they restrict motion to less than that seen physiologically. The usual designs are, however, “semiconstrained”, which allows for physiological movement, or “nonconstrained”, where there is no mechanical stop and extremes of motion are prevented by the perispinal soft tissue and inherent compression across the disc space [33]. The nonconstrained device

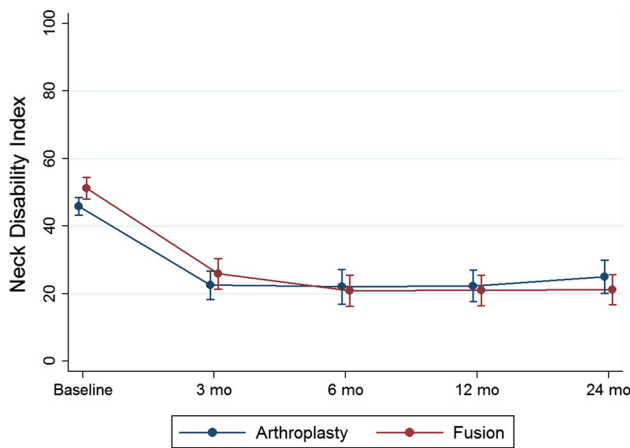


Fig. 2 Plot of the primary outcome measure Neck Disability Index from baseline to 2-year follow-up. The primary outcome measure was the Neck Disability Index (NDI). It is composed of 10 items, each scored from 0 to 5. It was calculated in percentage where a higher score indicates more severe symptoms. The figure shows the results of the observed improvement in NDI score from baseline to 3-month, 6-month, 1-year and 2-year follow-up for both treatment modalities (the intention-to-treat population) without adjustment for baseline differences, dropouts and missing data

used in the present trial is comparable in this respect with the Bryan device (Medtronic Spine and Biologics) [7, 8, 10] and the Porous Coated Motion (PCM) device (NuVasiveInc. San Diego, CA, USA) [11]. The Prestige ST (Medtronic Sofamore Danek) [6, 9] differs from the present study implant by its semiconstrained design, and by the implantation technique, where the device is fixed with screws to the vertebrae cranial and caudal to the disc space. In addition to different degree of constraint, implants may also differ in design of their articulating surfaces. The ball and socket design of the device used in the present trial has a different impact on range of motion (ROM) compared with the Bryan and PCM devices, and the adjacent level intradiscal pressure has been shown to differ according to implant design [32].

The study methods of the present trial also differ from the previously mentioned studies where only two describe blinding of the participating patients [7, 11]. However, Heller and colleagues [7] could not continue blinding of

patients after completion of the surgical procedure due to treatment with non-steroid anti-inflammatory medication (NSAID) in the arthroplasty group for two weeks after surgery. Phillips and colleagues [11] blinded patients only until after the surgical procedure was completed. Blinding of the surgical team until after decompression of the compressed nerve root has rarely been included in previous study designs, but was conducted in the study by Skeppholm and colleagues [14], consistent with the previous study methods. Strict study methods are probably important to avoid expectation bias in both patients and surgeons, and may have been a contributing factor to the discrepancy with previous trials.

Another aspect, which may influence the outcome, is the applied fusion technique. Stand-alone polyetheretherketone (PEEK) cage implant as used in the NORCAT differs from most other comparable trials, where allograft and anterior plating are most commonly used [6–11]. The reported fusion rates between the two techniques after 2 years are, however, similar at 97.5% [6], 94.3% [7], and 92.1% [11] for allograft with plating and 92% [34] for stand-alone PEEK cage. Nemoto and colleagues [34] recently assessed clinical outcome and complications regarding postoperative dysphagia between stand-alone cage implant versus cage and anterior plating in single-level cervical disc disease, and found no difference between the two surgical methods.

The length of follow-up may also have an impact on the clinical outcome, and longer observational period after surgery is often requested. Time is naturally highly relevant in relation to the impact of adjacent level disease [35]. However, the present study results demonstrate that there is little change in clinical outcome from 3 months up to 2 years after surgery. A longer follow-up has probably little effect on clinical outcome related to the completed surgery, as recently demonstrated by Gornet et al. [36].

Arthroplasty manufacturers are often represented as sponsors of large randomized, controlled trials, as was the case in the present study. Their role in relation to outcome is probably important to include in the overall discussion regarding outcome discrepancy between authors, and was recently discussed by Alvin and colleagues [37]. They

Table 2 Comparison of Neck Disability Index^a between arthroplasty and fusion group at each follow-up

	Arthroplasty	Fusion	<i>P</i>	Mean difference ^b	<i>P</i>
3 months	22.5 (18.3 to 26.7)	25.9 (21.4 to 30.4)	0.28	−0.5 (−6.4 to 5.4)	0.86
6 months	22.0 (16.9 to 27.2)	20.8 (16.2 to 25.5)	0.73	3.1 (−3.1 to 9.3)	0.32
1 year	22.3 (17.7 to 26.9)	20.9 (16.4 to 25.5)	0.68	2.9 (−3.0 to 8.9)	0.33
2 years	25.0 (20.1 to 29.9)	21.2 (16.7 to 25.6)	0.25	5.9 (0.02 to 11.9)	0.049

^a Mean Neck Disability Index (0 to 100%, lower scores indicate less severe symptoms), 95% CI

^b Mean difference in Neck Disability Index in percentage (95% CI) analyzed using linear mixed models for repeated measurements

Table 3 Comparison of secondary outcome measures between arthroplasty and fusion group at each follow-up

	Arthroplasty	Fusion	<i>P</i>	Mean difference ^a	<i>P</i>
Mean (95% CI) EQ-5D-3L					
Baseline	0.37 (0.29 to 0.45)	0.28 (0.20 to 0.35)			
3 months	0.73 (0.67 to 0.79)	0.67 (0.60 to 0.74)	0.19	0.04 (−0.05 to 0.13)	0.36
6 months	0.80 (0.74 to 0.85)	0.81 (0.75 to 0.86)	0.78	−0.03 (−0.13 to 0.07)	0.54
1 year	0.70 (0.63 to 0.78)	0.72 (0.65 to 0.79)	0.72	−0.03 (−0.12 to 0.06)	0.50
2 years	0.72 (0.64 to 0.80)	0.72 (0.67 to 0.79)	0.95	−0.02 (−0.11 to 0.07)	0.72
Mean (95% CI) SF-36					
Physical component					
Baseline	34.6 (32.9 to 36.2)	34.9 (32.9 to 36.8)			
3 months	44.6 (42.5 to 46.8)	43.9 (41.3 to 46.6)	0.68	1.1 (−2.2 to 4.3)	0.53
6 months	46.9 (44.0 to 49.7)	47.9 (45.3 to 50.4)	0.62	−0.5 (−4.0 to 3.1)	0.79
1 year	46.1 (43.7 to 48.4)	48.2 (45.5 to 50.9)	0.26	−1.4 (−4.8 to 1.9)	0.40
2 years	46.4 (43.7 to 49.3)	46.9 (44.5 to 49.1)	0.81	0.1 (−3.2 to 3.3)	0.97
Mental component					
Baseline	47.4 (44.5 to 50.1)	44.2 (41.6 to 46.5)			
3 months	52.0 (49.8 to 54.3)	50.7 (47.7 to 53.2)	0.45	0.3 (−3.0 to 3.7)	0.85
6 months	51.0 (48.2 to 53.7)	51.7 (49.2 to 54.8)	0.67	−1.8 (−5.5 to 1.9)	0.34
1 year	52.2 (49.5 to 54.8)	53.3 (51.0 to 55.6)	0.56	−1.9 (−5.4 to 1.5)	0.28
2 years	52.3 (49.3 to 54.9)	50.3 (47.2 to 53.3)	0.33	0.6 (−2.8 to 4.0)	0.72
Median (range) NRS 11					
Arm pain					
Baseline	6.0 (1.0 to 10.0)	6.5 (1.0 to 10.0)			
3 months	2.0 (0.0 to 8.0)	1.0 (0.0 to 10.0)	0.17	0.8 (−0.1 to 1.7)	0.10
6 months	2.0 (0.0 to 8.0)	1.0 (0.0 to 8.0)	0.16	1.1 (0.1 to 2.0)	0.03
1 year	2.0 (0.0 to 8.0)	1.0 (0.0 to 7.0)	0.11	0.9 (−0.03 to 1.8)	0.06
2 years	2.0 (0.0 to 10.0)	1.5 (0.0 to 8.0)	0.06	1.0 (0.1 to 1.9)	0.03
Neck pain					
Baseline	7.0 (0.0 to 10.0)	7.0 (1.0 to 10.0)			
3 months	3.5 (0.0 to 9.0)	3.0 (0.0 to 10.0)	0.89	0.1 (−0.8 to 0.9)	0.84
6 months	3.0 (0.0 to 9.0)	3.5 (0.0 to 8.0)	0.78	0.1 (−0.8 to 1.1)	0.79
1 year	3.0 (0.0 to 9.0)	3.0 (0.0 to 9.0)	0.64	0.2 (−0.6 to 1.2)	0.57
2 years	3.0 (0.0 to 10.0)	3.0 (0.0 to 10.0)	0.64	0.4 (−0.5 to 1.2)	0.44

^a Values are mean differences (95% CI) analyzed using linear mixed models for repeated measurements EQ-5D-3L, EuroQol-5 Dimension-3 Level score (−0.59 to 1, higher scores indicate better health status); SF-36, Short Form 36 (0 to 100, higher scores indicate better health status); NRS 11, Numeric Rate Scale 11 (0–10, 0 labeled as “no pain at all” and 10 as “worst imaginable pain”)

assessed whether trials funded by arthroplasty manufacturers had a greater likelihood of reporting results in favor of arthroplasty, and found lower complication rates when a conflict of interest was reported, but no impact on health-related quality of life outcomes.

Critical issues which may explain the discrepancy in clinical outcome between the present study and most previous comparable trials are difficult to point out. The truth, however, may be a combination of physiological and actual differences between the implants, as well as different study designs as discussed above.

The expected clinical outcome is important in the surgical decision-making for individual patients. In addition,

differences between surgical techniques are also key factors to consider. In the present trial, patients operated with arthroplasty had significantly longer duration of surgery, which corresponds to the results from a newly published meta-analysis [15]. Even though experienced spinal surgeons operated the patients, all surgeons were more familiar with the fusion procedure as it was the standard treatment in the departments involved. Thus, level of experience is one possible explanation for the difference in surgery duration. Other possible explanations are that implantation of the specific arthroplasty device is technically more demanding and time consuming. There were no severe complications in the present study, but the

reoperation rate differed from previous trials reporting more secondary surgeries with fusion [6, 8, 9]. The difference in index level reoperations could be explained by suboptimal implantation technique or incorrect size of the arthroplasty device. However, all patients who were reoperated had their primary surgery at a time-point when all surgeons had good experience with the particular arthroplasty device. In a recent study using the same implant [38], instability and accompanying neck pain after arthroplasty were found in 8% of patients, all of whom underwent revision surgery.

Corresponding with previous reports [6, 7], patients in the arthroplasty group returned to work two weeks earlier than patients in the fusion group, but there was no difference in employment status at 2-year follow-up. A previous study concluded that the duration of preoperative sick leave influenced return to work postoperatively [39]. In the present trial, preoperative sick leave was 3 weeks shorter in the arthroplasty group, but the difference was not significant.

Ament and colleagues recently assessed the cost-effectiveness of 2-level arthroplasty or fusion at 2- and 5-years follow-up. Arthroplasty was more expensive than fusion, but came out with higher total quality adjusted life years, suggesting it to be a highly cost-effective treatment option [40, 41]. Consistent with these results, Zou and colleagues recently presented a meta-analysis on clinical outcome after two-contiguous level cervical disc surgery and concluded that arthroplasty was equivalent, and in some aspects significantly superior to fusion regarding clinical outcome [42]. Considering the results of the present trial, the growing interest among physicians for arthroplasty as an alternative to fusion, and the high number of surgical procedures performed each year [43], future studies should focus on both clinical outcome as well as cost-effectiveness analyses.

The role of adjacent level disease was not addressed in the present study since clinical outcome was the only focus of this report. The impact of adjacent level disease will be presented in a forthcoming paper including the NORCAT 5-year follow-up data. Regarding maintenance of mobility, which is the main goal of choosing arthroplasty over fusion, the authors of the present study have recently shown that high-grade heterotopic ossification around the Discover arthroplasty device was found in 62% after 2 years [44].

Limitations

Our study may be criticized for a too short follow-up period. However, the present study shows that there is little change in clinical outcome from 3 months up to 2 years.

Similar results at even longer follow-up was recently presented by Staub and colleagues who reported quite stable postoperative course of patient-reported outcomes between 2 and 5 years both after arthroplasty and fusion based on registry data [29]. Their results also strengthen the external validity of randomized controlled trials comparing cervical arthroplasty and fusion, where a large number of patients often do not meet the inclusion criteria, as was the case in the present trial.

Even though no patients with severe spondylosis should have been included in the NORCAT, the degree of spondylosis using radiographic parameters for evaluation could have been emphasized specifically in the inclusion/exclusion criteria. Therefore, one cannot exclude the possibility that some patients not meeting the criteria for arthroplasty may have been included, which again could have biased the study in favor of the fusion group.

Conclusion

There was a high level of success for both treatment modalities at 2 years. Arthroplasty was not superior to fusion regarding clinical outcome. The rate of index level reoperations was higher and duration of the surgical procedure was longer with arthroplasty. More studies assessing clinical outcome and cost-effectiveness analyses are needed.

Acknowledgements We thank all patients for their benevolent participation. The authors would also like to thank Vigdis Skogli and Torun Gillebo at the Department of Neurosurgery, and Jorund Roos at the Department of Radiology, both Oslo University Hospital, Rikshospitalet, Oslo, Camilla Seif at the Department of Neurosurgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, and Monica Finnkirk at the Department of Neurosurgery, Haukeland University Hospital, Bergen for efforts to convene the patients, coordination of investigations and follow-up of the study.

Compliance with ethical standards

Disclosures Supported by the South-Eastern Norway Regional Health Authority (Grant Number 2012009). The study committee has received an unconditional Grant from DePuy Synthes Spine (325 Paramount Drive Raynham, MA 02767), a division of Johnson & Johnson AB, Norwegian branch office and OrtoMedic AS (Vollsveien 13E, 1366 Lysaker, Norway).

Conflict of interest None.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bogduk N (2003) The anatomy and pathophysiology of neck pain. *Phys Med Rehabil Clin N Am* 14:455–472 (v. **Review**)
- Carette S, Fehlings MG (2005) Clinical practice, Cervical radiculopathy. *N Engl J Med* 353:392–399
- Lee CH, Hyun SJ, Kim MJ, Yeom JS, Kim WH, Kim KJ, Jahng TA, Kim HJ, Yoon SH (2013) Comparative analysis of 3 different construct systems for single-level anterior cervical discectomy and fusion: stand-alone cage, iliac graft plus plate augmentation, and cage plus plating. *J Spinal Disord Tech* 26:112–118
- Nesterenko SO, Riley LH 3rd, Skolasky RL (2012) Anterior cervical discectomy and fusion versus cervical disc arthroplasty: current state and trends in treatment for cervical disc pathology. *Spine* 37:1470–1474
- Hillibrand AS, Robbins M (2004) Adjacent segment degeneration and adjacent segment disease: the consequences of spinal fusion. *Spine J* 4:190–194
- Mummaneni PV, Burkus JK, Haid RW, Traynelis VC, Zdeblick TA (2007) Clinical and radiographic analysis of cervical disc arthroplasty compared with allograft fusion: a randomized controlled clinical trial. *J Neurosurg Spine* 6:198–209
- Heller JG, Sasso RC, Papadopoulos SM, Anderson PA, Fessler RG, Hacker RJ, Coric D, Cauthen JC, Riew DK (2009) Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. *Spine* 34:101–107
- Garrido BJ, Taha TA, Sasso RC (2010) Clinical outcomes of Bryan cervical disc arthroplasty a prospective, randomized, controlled, single site trial with 48-month follow-up. *J Spinal Disord Tech* 23:367–371
- Burkus JK, Haid RW, Traynelis VC, Mummaneni PV (2010) Long-term clinical and radiographic outcomes of cervical disc replacement with the Prestige disc: results from a prospective randomized controlled clinical trial. *J Neurosurg Spine* 13:308–318
- Sasso RC, Anderson PA, Riew KD, Heller JG (2011) Results of cervical arthroplasty compared with anterior discectomy and fusion: 4-year clinical outcomes in a prospective, randomized controlled trial. *J Bone Joint Surg Am* 93:1684–1692
- Phillips FM, Lee JY, Geisler FH, Cappuccino A, Chaput CD, DeVine JG, Reah C, Gilder KM, Howell KM, McAfee PC (2013) A prospective, randomized, controlled clinical investigation comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion. 2-year results from the US FDA IDE clinical trial. *Spine* 38:E907–E918
- Rozankovic M, Marasanov SM, Vukic M (2016) Cervical disc replacement with discover versus fusion in a single level cervical disc disease: a prospective single center randomized trial with a minimum 2-year follow-up. *Clin Spine Surg*. 2016 Jun 9 [**Epub ahead of print**]
- Coric D, Kim PK, Clemente JD, Boltz MO, Nussbaum M, James S (2013) Prospective randomized study of cervical arthroplasty and anterior cervical discectomy and fusion with long-term follow-up: results in 74 patients from a single site. *J Neurosurg Spine* 18:36–42
- Skeppholm M, Lindgren L, Henriques T, Vavruch L, Lofgren H, Olerud C (2015) The Discover artificial disc replacement versus fusion in cervical radiculopathy—a randomized controlled outcome trial with 2-year follow-up. *Spine J* 15:1284–1294
- Luo J, Huang S, Gong M, Dai X, Gao M, Yu T, Zhou Z, Zou X (2015) Comparison of artificial cervical arthroplasty versus anterior cervical discectomy and fusion for one-level cervical degenerative disc disease: a meta-analysis of randomized controlled trials. *Eur J Orthop Surg Traumatol* 25(Suppl 1):S115–S125
- Hu Y, Lv G, Ren S, Johansen D (2016) Mid- to Long-Term Outcomes of Cervical Disc Arthroplasty versus Anterior Cervical Discectomy and Fusion for Treatment of Symptomatic Cervical Disc Disease: a Systematic Review and Meta-Analysis of Eight Prospective Randomized Controlled Trials. *PLoS One* 11:e0149312
- Aragonés M, Hevia E, Barrios C (2015) Polyurethane on titanium unconstrained disc arthroplasty versus anterior discectomy and fusion for the treatment of cervical disc disease; a review of level I-II randomized clinical trials including clinical outcomes. *Eur Spine J* 24:2735–2745
- Vernon H, Mior S (1991) The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 14:409–415
- Johansen JB, Andelic N, Bakke E, Holter EB, Mengshoel AM, Roe C (2013) Measurement properties of the norwegian version of the neck disability index in chronic neck pain. *Spine* 38:851–856
- Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA (1978) Studies with pain rating scales. *Ann Rheum Dis* 37:378–381
- Ware JE Jr, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30:473–483
- Brooks RG, Jendteg S, Lindgren B, Persson U, Bjork S (1991) EuroQol: health-related quality of life measurement. Results of the Swedish questionnaire exercise. *Health policy (Amsterdam, Netherlands)* 18:37–48
- Skeppholm M, Ingebro C, Engstrom T, Olerud C (2012) The Dysphagia Short Questionnaire: an instrument for evaluation of dysphagia: a validation study with 12 months' follow-up after anterior cervical spine surgery. *Spine* 37:996–1002
- Loge JH, Kaasa S, Hjermstad MJ, Kvien TK (1998) Translation and performance of the Norwegian SF-36 Health Survey in patients with rheumatoid arthritis. I. Data quality, scaling assumptions, reliability, and construct validity. *J Clin Epidemiol* 51:1069–1076
- Stavem K (1999) Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease. *Qual Life Res* 8:45–54
- EuroQol Group EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy (Amsterdam, Netherlands)* 16:199–208
- MacDermid JC, Walton DM, Avery S, Blanchard A, Etruw E, McAlpine C, Goldsmith CH (2009) Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther* 39:400–417
- Pool JJ, Ostelo RW, Hoving JL, Bouter LM, de Vet HC. (2007) Minimal clinically important change of the Neck Disability Index and the Numerical Rating Scale for patients with neck pain. *Spine (Phila Pa 1976)* 32:3047–3051
- Staub LP, Ryser C, Röder C, Mannion AF, Jarvik JG, Aebi M, Aghayev F (2016) Total disc arthroplasty versus anterior cervical interbody fusion: use of the Spine Tango registry to supplement the evidence from randomized control trials. *Spine J* 16:136–145
- Revicki D, Hays RD, Cella D, Sloan J (2008) Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 61:102–109
- Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS (1998) Interpreting treatment effects in randomised trials. *BMJ* 316:690–693
- Mo Z, Zhao Y, Du C, Sun Y, Zhang M, Fan Y (2015) Does location of rotation center in artificial disc affect cervical biomechanics? *Spine* 40:E469–E475

33. Puttlitz CM, DiAngelo DJ (2005) Cervical spine arthroplasty biomechanics. *Neurosurg Clin North Am* 16:589–594
34. Nemoto O, Kitada A, Naitou S, Tachibana A, Ito Y, Fujikawa A (2015) Stand-alone anchored cage versus cage with plating for single-level anterior cervical discectomy and fusion; a prospective, randomized, controlled study with a 2-year follow-up. *Eur J Orthop Surg Traumatol* 25:127–134
35. Helgeson MD, Bevevino AJ, Hilibrand AS (2013) Update on the evidence for adjacent degeneration and disease. *Spine J* 13:342–351
36. Gornet MF, Burkus JK, Shaffrey ME, Nian H, Harrell FE Jr (2016) Cervical disc arthroplasty with prestige LP disc versus anterior cervical discectomy and fusion. Seven-year outcomes. *Int J Spine Surg* 22(10):24
37. Alvin MD, Abbott EE, Lubelski D, Kuhns B, Nowacki AS, Steinmetz MP, Benzel EC, Mroz TE (2014) Cervical arthroplasty: a critical review of the literature. *Spine J* 14:2231–2245
38. Skeppholm M, Svedmark P, Noz ME, Maguire GQ Jr, Olivecrona H, Olerud C (2015) Evaluation of mobility and stability in the Discover artificial disc: an in vivo motion study using high-accuracy 3D CT data. *J Neurosurg Spine* 23:383–389
39. Bhandari M, Louw D, Reddy K (1999) Predictors of return to work after anterior cervical discectomy. *J Spinal Disord* 12:94–98
40. Ament JD, Yang Z, Nunley P, Stone MB, Kim KD (2014) Cost-effectiveness of cervical total disc replacement vs fusion for the treatment of 2-level symptomatic degenerative disc disease. *JAMA Surg* 149:1231–1239
41. Ament JD, Yang Z, Nunley P, Stone MB, Lee D, Kim KD (2016) Cost utility analysis of the cervical artificial disc vs fusion for the treatment of 2-level symptomatic degenerative disc disease: 5-year follow-up. *Neurosurgery* 79:135–145
42. Cou S, Gao J, Xu B, Lu X, Han Y, Meng H (2016) Anterior cervical discectomy and fusion (ACDF) versus cervical disc arthroplasty (CDA) for two contiguous levels cervical disc degenerative disease: a meta-analysis of randomized controlled trials. *Eur Spine J* 17 [**Epub ahead of print**]
43. Patil PG, Turner DA, Pietrobon R (2005) National trends in surgical procedures for degenerative cervical spine disease: 1990–2000. *Neurosurgery* 57:753–758 (**discussion 753–758**)
44. Sundseth J, Jacobsen EA, Kolstad F, Sletteberg RO, Nygaard OP, Johnsen LG, Pripp AH, Andresen H, Fredriksli OA, Myrseth E, Zwart JA (2016) Heterotopic ossification and clinical outcome in nonconstrained cervical arthroplasty 2 years after surgery: the Norwegian Cervical Arthroplasty Trial (NORCAT). *Eur Spine J* 25:2271–2278