The Spine Journal 19 (2019) 965–975

Clinical Study

The NECK trial: Effectiveness of anterior cervical discectomy with or without interbody fusion and arthroplasty in the treatment of cervical disc herniation; a double-blinded randomized controlled trial

Carmen L.A. Vleggeert-Lankamp, MD, MSc, PhD[a,*], Tessa M.H. Janssen, BSc[a], Erik van Zwet, PhD[b], Caroline M.W. Goedmakers[a], Lisette Bosscher, MD[d], Wilco Peul, MD, PhD[a,c], Mark P. Arts, MD, PhD[c]

[a] *Department of Neurosurgery, Leiden University Medical Center, Leiden, the Netherlands*
[b] *Department of Biostatistics, Leiden University Medical Center, Leiden, the Netherlands*
[c] *Department of Neurosurgery, Haaglanden Medical Center, the Hague, the Netherlands*
[d] *Medical Center Alkmaar, Alkmaar, the Netherlands*

Received 27 June 2018; revised 17 December 2018; accepted 18 December 2018

Abstract

BACKGROUND: Motion preserving anterior cervical disc arthroplasty (ACDA) in patients with cervical radiculopathy was introduced to prevent symptomatic adjacent disc degeneration as compared with anterior cervical discectomy and fusion (ACDF). Prior reports suggest that ACDF is not more effective than anterior cervical discectomy (ACD) alone for the treatment of cervical radiculopathy.
PURPOSE: To evaluate whether patients with cervical radiculopathy due to a herniated disc benefit more from undergoing ACDA, ACDF, or ACD in terms of clinical outcome measured by the neck disability index (NDI).
STUDY DESIGN: Double-blinded randomized controlled trial.
METHODS: One hundred-nine patients with one level herniated disc were randomized to one of the following treatments: ACDA, ACDF with intervertebral cage, ACD without fusion. Clinical and radiological outcome was measured by NDI, Visual Analogue Scale (VAS) neck pain, VAS arm pain, SF36, EQ-5D, patients' self-reported perceived recovery, radiographic cervical curvature, and adjacent segment degeneration parameters at baseline and until two years after surgery. BBraun Medical paid €298.837 to cover the costs for research nurses.
RESULTS: The NDI declined from 41 to 47 points at baseline to 19 ± 15 in the ACD group, 19 ± 18 in the ACDF group, and 20 ± 22 in the ACDA group after surgery (p=.929). VAS arm and neck pain declined to half its baseline value and decreased below the critical value of 40 mm. Quality of life, measured by the EQ-5D, increased in all three groups. Adjacent segment degeneration parameters were comparable in all three groups as well. No statistical differences were demonstrated between the treatment groups.

FDA device/drug status: Not applicable.
Grant: BBraun Medical (F, paid directly to institution).

Author disclosures: ***CLAV-L:*** Grant: BBraun Medical (F, paid directly to institution). Speaking/Teaching Arrangements: CSRS (A), CSRS AP (A). Research Support (Investigator Salary, Staff/Materials): Medtronic, Paradigm, Covidien (paid direclty to institution), Dutch National Organization (paid directly to institution). Fellowship Support: Chinese Student Council (student support/sponsorship). ***TMHJ:*** Grant: BBraun Medical (F, paid directly to institution). ***EVZ:*** Nothing to disclose. ***CMWG:*** Grant: BBraun Medical (F, paid directly to institution). ***LB:*** Grant: BBraun Medical (F, paid directly to institution). ***WP:*** Grant: BBraun Medical (F, paid directly to institution), Hersenstichting Brain Foundation Netherlands (G, paid directly to institution), Center TBI (F, paid directly to institution). ***MPA:*** Royalties: EIT (amount not disclosed). Stock Ownership: Galapagos (amount not disclosed), Nuvasive (amount not disclosed). Consulting: Zimmer-Biomet (amount not disclosed, paid directly to institution), EIT (amount not disclosed, paid directly to institution), Silony (amount not disclosed, paid directly to institution), Intrinsics (amount not disclosed, paid directly to institution). Research Suppport: Zimmer-Biomet (amount not disclosed, paid directly to institution) EIT (amount not disclosed, paid directly to institution), Instrinsics (amount not disclosed, paid directly to institution).

* Corresponding author. Department of Neurosurgery, Leiden University Medical Center, Albinusdreef 2, 2300 RC Leiden, the Netherlands.
Tel.: +31-71-5262109.
E-mail address: cvleggeert@lumc.nl (C.L.A. Vleggeert-Lankamp).

https://doi.org/10.1016/j.spinee.2018.12.013

CONCLUSIONS: The hypothesis that ACDA would lead to superior clinical outcome in comparison to ACDF or ACD could not be confirmed during a 2-year follow-up time period. Single level ACD without implanting an intervertebral device may be a reasonable alternative to ACDF or ACDA.

Keywords: Cervical discectomy; Cervical radiculopathy; Prosthesis; Neck disability index; Radicular pain; Disc herniation

Introduction

Anterior cervical discectomy (ACD) is an accepted surgical treatment for patients with radicular pain caused by cervical disc herniation. In 1958, Cloward and Smith and Robinson first described the first anterior cervical discectomy and fusion (ACDF) using autologous iliac crest interbody graft to maintain disc height [1]. However, arthrodesis of a motion segment can lead to increased mechanical load at the adjacent levels, hypothetically leading to adjacent segment degeneration (ASD). In some cases, this can be symptomatic and cause neck disability and pain. The rationale for anterior cervical discectomy with arthroplasty (ACDA) is to maintain motion while maintaining disc space height.

It should be understood that solid proof that ACDA prevents ASD is lacking. We recently published a systematic review on the radiological follow-up of patients after implanting a prosthesis following anterior discectomy [2]. It revealed low-level evidence that ASD increased more (10%−20%) in the fusion group compared with the prosthesis group after long-term follow-up. However, this was only studied in patients suffering from both radiculopathy and myelopathy, which are prone to have a more degenerative spine than those with soft disc herniations. Further study of ASD and its clinical implications after ACDA surgery is therefore desirable.

Most outcome studies regarding ACDA are industry-sponsored and primarily report on the safety of the device and the relief of pain after surgery [3−6]. Only a few exclusively evaluated patients suffering from cervical radiculopathy due to a herniated disc, excluding patients with myelopathy [7−16]. A systematic review of these studies revealed that only minimal clinical differences between the groups exist with clinical debatable relevance [2]. Issues with the vast majority of these studies are that the support for the authors' conclusions is disappointing, since the risk of bias is high, and the results are contradictory and not presented precisely enough. Thus, in our view, ACDA cannot be unequivocally presented as better or a suitable alternative.

In this study, the Netherlands Cervical Kinetics (NECK) trial, we compare single level ACDA not only to ACDF but also to ACD. Single level ACDA was randomly and blindly compared with ACD and ACDF (using a PEEK cage and no plate). It was our hypothesis that better clinical outcomes as measured by the neck disability index (NDI) would be found with ACDA at 2 years follow-up period.

Materials and methods

A prospective, randomized double-blinded multicenter trial was conducted among patients with cervical radiculopathy due to single-level disc herniation (Netherlands Cervical Kinematics, or NECK, trial). Patients were randomly assigned using a computer into three groups: ACDA, ACDF, and ACD. Both patients and research nurses evaluating clinical outcome were blinded to the allocated treatment. The protocol was approved by the Central Medical Ethics Committee Leiden ("Commissie Medische Ethiek Leiden University Medical Center," decision letter P08.011) and the board of directors of the Rijnland hospital Leiderdorp, Diaconessenhuis Leiden, Haaglanden Medical Center and Antoniushove the Hague, including an approval for randomization after anesthetic induction, in agreement with the Central Ethics Committee Leiden. The protocol was also approved by the "Medical Ethics Committee Noord-Holland" for the Medical Center Alkmaar (M08−038). Written informed consent was obtained from all patients. The design and study protocol were previously published [17]. Dutch Trial Register Number: NTR1289.

Eligibility and randomization

Patients (aged 18−65 years old) with radicular signs and symptoms in one or both arms (pain, paresthesia, or paresis in a specific nerve root distribution) for at least 8 weeks and for whom conservative therapy failed were eligible for inclusion. All patients were diagnosed with cervical radiculopathy by a neurologist in one of the participating hospitals. If MRI demonstrated a single-level cervical disc herniation with or without an accompanying osteophyte at one level (C3-C4 to C7-Th1) in accordance with clinical signs and symptoms, patients could be included as surgical candidates for the study by the consulting neurosurgeon. At the time of enrollment, an independent research nurse verified the persistence of the symptoms. Patients with previous cervical surgery (either anterior or posterior), absence of motion, increased anteroposterior translation, very narrow (<3 mm) intervertebral space, severe segmental kyphosis (>3°) at the index level on static or dynamic X-rays, neck pain only and symptoms and signs of myelopathy were excluded. Furthermore, patients with metabolic and bone diseases (osteoporosis, severe osteopenia), neoplasm or trauma of the cervical spine, spinal anomaly (Klippel Feil, Bechterew, OPLL), severe mental or psychiatric disorders were excluded.

A randomized design with variable block sizes was used (computer-assisted) in a 1:1:1 ratio, with allocations stratified by center. Allocations were stored in prepared opaque, coded, and sealed envelopes. The key was only accessible to the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. All patients gave informed consent.

After induction of anesthesia, the prepared envelope was opened and patients were randomly allocated to one of the treatment arms. Patients, the nursing department and research nurses maintained the blind allocation of treatment group during the follow-up period of 2 years.

Disc prosthesis

The investigational device used in the ACDA group was the activC flat artificial cervical disc (Aesculap AG, Tuttlingen, Germany). The activC device is composed of two flat Cobalt-Chrome-Molybden alloy metal endplate components with spikes on the superior endplate and an inferior endplate and a keel for primary stability. The inferior prosthesis plate has an integrated polyethylene inlay. The implants are available in six different sizes (XS, S, M, L, XL, and XXL) and in three heights (5, 6, and 7 mm).

Interventions

All patients were in the supine position with their neck slightly extended under general anesthesia. The affected cervical disc level was identified using fluoroscopy. A small transverse incision was made either on the right or left side depending on the surgeon's preference. Medial to the carotid sheath, the prevertebral space was opened and the anterior cervical spine was exposed. Caspar spreader and two distraction pins were placed in the effected segment. Care was taken to not damage the adjacent level discs. A standard anterior discectomy using loupe magnification or microscope (depending on the surgeon's preference) was performed in all cases. The posterior longitudinal ligament was opened and the nerve root and dura were decompressed adequately. If required, a vacuum drain was applied and the wound was closed in layers.

For patients randomized to the ACD group no intervertebral device was placed the procedure was a discectomy without fusion. In patients randomized to the ACDF group, an interbody PEEK cage either filled with synthetic bone substitute or autologous bone (chips locally harvested) was placed within the intervertebral space under fluoroscopic guidance. The brand of cage depended on the surgeons' preference and daily practice. For patients randomized to the ACDA group, special attention was given to the placing of the distraction pins in the adjacent levels. With anteroposterior fluoroscopy the midvertebral body position was ensured. After decompression of dura and nerve roots, the implant size and height were determined and the endplates were prepared for proper fitting of the activC flat prosthesis, including preparing a sleeve for the keel to fit in. The device was inserted under slight distraction and fluoroscopic guidance. Postoperatively, all patients were encouraged to mobilize as soon as possible. No collars were prescribed.

In all participating centers, one senior surgeon with advanced training in cervical spine surgery was trained to implant the prosthesis. After implanting 10 prostheses, the surgeon was allowed to implant prostheses for the RCT. Four of the participating hospitals referred their patients for surgery to the main referral hospital (LUMC), where patients were operated on by one trained surgeon dedicated to this trial. Overall, three surgeons in three hospitals were responsible for the implantation of the prostheses (16, 29, and 64 interventions per center).

Outcomes

The primary outcome measure was the NDI. The NDI is a 10-item questionnaire on three different aspects; pain intensity, daily work-related activities, and nonwork-related activities. Each item is scored from 0 to 5and the total score ranges from 0 (best score) to 50 (worst score).This 50-point score was converted to a 100-point scale (50 points = 100 points). The NDI is a modification of the Oswestry Low Back Pain Index and has been shown to be reliable and valid for patients with cervical pathology [18–20].

Secondary outcome measures were the Visual Analogue Scale for arm pain (VAS arm) and for neck pain (VAS neck), the EuroQol quality of life questionnaire, including a Visual Analogue Scale for health (VAS health), the Likert patient perceived recovery scale, and the Medical Outcome Study 36-item short-form Generated Health Survey (SF 36). The VAS pain measures the experienced pain intensity during the week before visiting the research nurse. Pain was assessed on a horizontal 100 mm scale varying from 0 mm (no pain) to 100 mm (worst pain imaginable). Patients do not see the results of earlier assessments and score the pain experienced at the visit. Reliability, validity, and responsiveness of VAS have been shown previously [21].

The EuroQol (EQ-5D) measures five dimensions (mobility, self-care, daily activities, pain/discomfort, anxiety/depression), on a three-point scale (no, some, or extreme problems).

Whereas the EQ-5D provides society's valuation for the patients' health, the patients themselves will also provide their own valuations for their health on a visual analogue scale, ranging from 0,0 (as bad as death) to 1,0 (optimal health).

The patients were asked to judge their postoperative recovery ("perceived recovery") on a scale varying from "complete recovery" to "worse than ever" in 7 steps (7-point Likert scale). This outcome scale has been used in previous studies and is regarded valid and responsive to change [22]. "Complete recovery" and "almost complete recovery" are defined as a good result, which was used to dichotomize the data. A Likert perceived recovery evaluation was performed for recovery of global health and recovery of arm pain separately.

The SF36 is a generic health status questionnaire that can easily be filled out at home. The questionnaire consists of 36 items on physical and social status of the patient divided into subscales. The questions are scored on a scale of 0 (worst health) to 100 (ideal health). This questionnaire has been used frequently and is validated in surgical studies on spinal column pathology [23–25].The PCS and MCS

are derived from the SF-36 and are summary scores for respectively the Physical Quality of Life and the Mental Quality of Life. The PCS and MCS range from 0 to 100 with higher scores representing better self-reported health.

All outcome scores were assessed at baseline, at 2, 4, 8, 12, 26, 52, and 104 weeks, except for the Likert perceived recovery score which was not evaluated at baseline. Follow-up will be extended to 10 years.

Radiological outcome measures

Fluoroscopy of the cervical spine was obtained at baseline and at the 1- and 2-year follow-up point. ASD was evaluated using the Goffin criteria, scoring both the upper adjacent and lower adjacent level as normal, mild, moderate, or severe ASD, focusing on decrease in disc height and presence of osteophytes [26]. Data were assessed at baseline and at 2 years after randomization. Outcome data were dichotomized as "no ASD" and "ASD" (mild, moderate, and severe) and compared between the three randomization arms.

To evaluate the shape of the cervical spine and to judge lordosis or kyphosis, a line was drawn from the posterior inferior part of C2 to the posterior superior part of C7. If the posterior lining of the elements of C3–C6 was anterior of this line, we judged the cervical spine to be lordotic, if this line crossed the C2–C7 line, we judged the cervical spine to be kyphotic. If the posterior line coincided with the line from C2 to C7, we judged the cervical spine to be straight.

All patients underwent MRI at baseline and after 1 year. MR images were performed at each study center using a standardized protocol tailored to a 1.5 or 3 Tesla scanner. Standard sagittal T_1 and T_2 and T_2 axial images were obtained, using 3 mm contiguous slices in all directions and an in-plane resolution of 1 mm^2 or less. Disc degeneration at the superior and inferior level of the index level on MRIs were independently evaluated using the Pfirrman scale [27] by one senior neurosurgeon dedicated to spine surgery and a junior medical doctor educated for this purpose. The reviewers were not provided with any clinical information of the included patients. Before the evaluation of MRIs, the reviewers met in person to evaluate and refine the definitions.

Sample size

The original aim of the NECK trial was to evaluate the incidence of symptomatic ASD with clinical outcome (NDI) as a secondary outcome parameter. To that end, a sample size calculation was made. An annual incidence of ASD of 2% after arthroplasty and 7% after interbody fusion was assumed for the power calculations [28]. In order to obtain 90% power, 0.05 significance level, and accrual duration of 3 years, corresponding to a Hazard Ratio of 0.28, we would need 675 patients in total. Assuming 10% lost to follow-up, a total of 750 patients was calculated. With this goal, the study was started. However, after a few years it became obvious that it would take more than 15 years to include this number of patients. Therefore, we changed our protocol and made new calculations using the NDI as a primary outcome measure.

It was assumed that, in general, 60% of the patients will have an excellent outcome with no complaints [29].We defined a 20-point lower NDI score (on a 100 point scale) as a clinically relevant benefit to justify ACDA. This value was decided based upon the assumption that superiority would be convincing enough to change the surgical guidelines [30–32].

The sample size calculation was based on the hypothesis that the average NDI at 2 years after ACDA is superior to the average NDI among anterior discectomy with or without interbody fusion. We required a power of 80%, a significance level of 0.05, and we assumed a standard deviation of 9 (based on independent data). This led to the need for 111 patients for a two-group comparison. Since two 2-group comparisons are separately performed, we calculated a need of 166 patients in total. Data were only accessible to the researchers after the follow-up period of 2 years. The East software (version 6) was used for the design of this study.

Statistical analysis

Groups were compared based on an intention-to-treat analysis. Differences between groups at baseline were tested. We tested the effect of treatment at 2 years. In case of significance, we performed post hoc tests between the three treatment groups. Differences between groups at all follow-up (2, 4, 8, 12, 26, 52 weeks) points were analyzed with repeated measurement analysis. To account for the correlation between repeated measurements of the same individual, generalized estimating equations were used.

At the moment of randomization, the study was stratified by the administrative center for the purpose of analyzing possible heterogeneity among centers and attempting a clinical interpretation of such heterogeneity. Those centers that were referring patients to the same hospital and same surgeon for treatment were combined. Hence, for the purpose of the analysis of heterogeneity, a center means the actual location where the treatment (according to random allocation) took place.

For analysis of the ASD data, an ANOVA was used to compare the three intervention groups. If no data were present in one of the groups, a Fisher exact test was performed. Data collection and checking for quality were performed with the ProMISe data management system of the Department of Medical Statistics and BioInformatics of the Leiden University Medical Center. IBM SPSS software, version 22.0, was used for all statistical analysis.

Results

Between October 2010 and July 2014, 156 consecutive patients with cervical radiculopathy due to a herniated disc were eligible for inclusion. Forty-four patients declined from participation and 112 patients signed informed consent

and were enrolled in the NECK Trial (Fig. 1). Three patients were excluded because baseline data were missing.

The intended number of 166 patients was not achieved. The rate of inclusion was low, and it was calculated that a further 2-year inclusion time would be needed to reach the total number of 166 patients. An interim study analysis demonstrated that the clinical evaluation parameters were comparable in the three treatment arms. No trend toward better results in one of the treatment groups could be observed (across all the outcome parameters). Even the calculation of the maximum possible differences revealed minor dissimilarities between the treatment strategies. It was decided that continuing the study would not lead to other conclusions since standard deviations were low and groups were comparable. We decided that outcome data were more relevant if it would be available within due time.

The included 109 patients were randomly assigned to ACD (38 patients), ACDF (36 patients), or ACDA (35 patients). No cross over occurred. There were no significant differences in baseline characteristics between the three treatment arms (Table 1). At the 2-year follow-up point, data were available of 98 out of 109 patients (90% compliance rate; Fig. 1). Data on operating time, blood loss, and hospital stay are summarized in Table 2.

Clinical outcome measures

All outcome measures improved after surgery, regardless of the treatment strategy (Table 3). The NDI decreased significantly from41 to 47 points at baseline to 19 ± 15 in the ACD group, 19 ± 18 in the ACDF group, and 20 ± 22 in the ACDA group after surgery to 19–20 points after 1 year in all three groups ($p<.001$). There was no significant difference in the absolute value of NDI, nor in the decline of NDI among the three groups. After 2 years, the NDI outcomes were comparable to the 1-year follow-up evaluation data and likewise no difference among the three groups existed. Overall, there was a clinically relevant decrease of 60% in NDI from baseline to 2 years after surgery, irrespective of the surgical strategy (Fig. 2). Moreover, the estimated marginal means were calculated from the mixed models analysis and demonstrated small standard errors (Table 4).

The calculated sample size of 166 patients (based on the NDI 20-point difference on a 100-point scale) was not achieved. However, the differences between the absolute NDI values between the groups were small and more importantly, the differences in effect between treatments were low (Table 5) varying from −4.5 to 3.6 on a 100-point NDI scale. Additionally, the maximum possible differences between the individual treatment strategies were estimated based on the distribution in outcomes using a linear mixed model. The model revealed that the maximum possible differences in outcome of NDI between the individual groups ranged from −10.5 to 9.8 on a 100-point scale (Table 5), never exceeding the minimal clinically important difference of 20 points on the 100-point NDI scale.

The VAS arm pain improved significantly from 57 to 64 mm before surgery to 16–24 after 1 year in all treatment

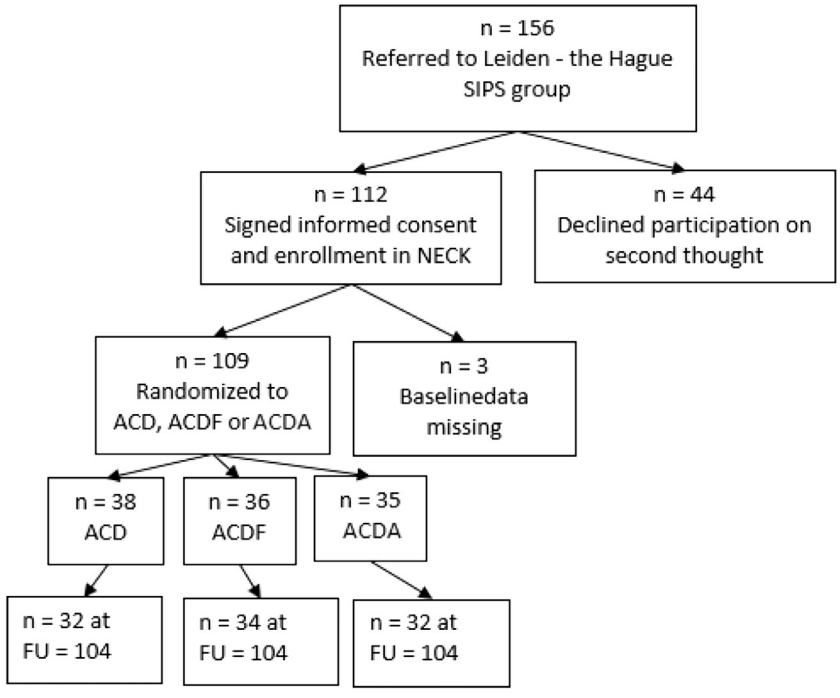

Fig. 1. Overview of patients' enrollment.

One hundred fifty-six were willing to consider participating in the study, 44 patients declined participation after the initial approach, and 112 patients signed informed consent. For 3 patients, baseline data was missing, so these were excluded from evaluation. The remaining 109 patients were randomly assigned to ACD (38 patients), ACDF (36 patients), or ACDA (35 patients).

Table 1
Patient demographics

	ACD	ACDF	ACDA	p value
Age (y; mean ± SD)	46.4 ± 7.3	47.5 ± 8.0	46.5 ± 8.7	.806
Body mass index (mean ± SD)	25.5 ± 3.5	27.6 ± 5.5	26.9 ± 3.6	.118
Gender				
Man	20	14	17	
Woman	18	22	18	.480
Smoking				
None	21	20	21	
Occasionally	1	3	1	
Regularly	16	13	13	.748
Level HNP				
C5C6	19	19	19	
C6C7	19	16	16	
C7T1	0	1	0	.932
Duration of complaints (weeks, mean ± SD)	38.4 ± 53.5	56.9 ± 91.2	46.6 ± 68.2	.548
Baseline NDI	45.2 ± 15.6	41.0 ± 13.2	46.5 ± 17.2	.294
Baseline VAS arm (1)	64.2 ± 21.6	56.6 ± 20.3	60.1 ± 21.8	.331
Baseline VAS neck (2)	56.5 ± 31.3	52.8 ± 25.8	49.6 ± 27.2	.586

Baseline characteristics of all patients in the ACD, ACDF, and ACDA group. No statistical significant differences were present between the three groups, with the exception of the body mass index. This was significantly smaller in the ACD group. Running BMI as a covariate in the GEE analysis did not result in a significant influence of BMI on NDI (p=.148).

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty; BMI, body mass index; NDI, neck disability index; GEE, generalized estimating equations; VAS, visual analogue scale.

Table 2
Outcome measures with concern to the surgical data for ACD, ACDF, and ACDA group

	ACDA	ACDF	ACD	
Surgery time (min)	73 ± 23	48 ± 18	43 ± 19	p < .001
Blood loss (ml)	113 ± 169	95 ± 132	68 ± 120	p = .396
Hospital stay (d)	1.5 ± 0.6	1.1 ± 0.4	1.2 ± 0.4	p = .079

No statistical significant differences were present between the three groups.

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty.

arms and remained at this value at the 2-year follow-up time point. There was no significant difference in improvement of VAS arm pain among the three surgical interventions (Fig. 3).

The VAS neck pain demonstrated a similar pattern as the VAS arm pain, with comparable reduction of neck pain in all treatment arms after 1 and 2 years without significant differences among the groups (Fig. 4). Quality of life, measured by the EQ-5D, increased in all treatment groups and the VAS health increased by 40% after surgery. No statistically significant differences were demonstrated between the treatment groups.

The Likert perceived recovery scale, both the global scale and the scale with specific attention for arm pain, demonstrated that two-thirds of the patients were satisfied with the intervention, both after 1 and after 2 years, irrespective of the treatment (Table 3). SF-36 baseline scores started around 40 for PCS and 55–60 (out of 100)for MCS at baseline and increased to the more favorable scores of 70–80 during the follow-up period (Table 3). No significant differences between the groups in improvement of PCS and MCS were seen.

Table 3
Clinical outcome at baseline, after 1 year and after 2 years of follow-up

	Baseline	1-year FU	2-year FU
NDI			
ACD	45 ± 16	21 ± 16	19 ± 15
ACDF	41 ± 13	18 ± 17	19 ± 18
ACDA	47 ± 17	18 ± 18	20 ± 22
p value	.294	.711	.929
VAS arm pain			
ACD	64 ± 22	24 ± 31	18 ± 25
ACDF	57 ± 20	18 ± 26	15 ± 23
ACDA	60 ± 24	16 ± 19	17 ± 30
p value	.331	.442	.880
VAS neck pain			
ACD	56 ± 31	24 ± 27	21 ± 23
ACDF	53 ± 26	29 ± 28	23 ± 27
ACDA	50 ± 27	17 ± 19	23 ± 32
p value	.586	.140	.934
EQ-5D			
ACD	0.54 ± 0.20	0.85 ± 0.16	0.83 ± 0.17
ACDF	0.70 ± 0.18	0.83 ± 0.15	0.83 ± 0.18
ACDA	0.59 ± 0.20	0.84 ± 0.16	0.82 ± 0.23
p value	.003	.901	.985
VAS health			
ACD	48 ± 26	71 ± 23	69 ± 24
ACDF	53 ± 23	76 ± 22	74 ± 24
ACDA	45 ± 22	72 ± 21	74 ± 25
p value	.336	.697	.663
Likert; global health (%)			
ACD		52.9	62.5
ACDF		76.5	67.6
ACDA		61.8	65.6
p value		.125	.907
Likert; arm pain (%)			
ACD		58.8	68.8
ACDF		76.5	73.5
ACDA		67.7	65.6
p value		.298	.781
Physical component score (PCS)			
ACD	41.2 ± 14	69.7 ± 24	68.3 ± 24
ACDF	44.7 ± 15	71.8 ± 24	75.9 ± 23
ACDA	41.3 ± 14	70.1 ± 21	72.2 ± 27
p value	.754	.529	.622
Mental component score (MCS)			
ACD	57.9 ± 21	75.6 ± 21	71.2 ± 23
ACDF	61.7 ± 22	76.8 ± 23	81.6 ± 19
ACDA	54.9 ± 25	75.7 ± 20	74.3 ± 25
p value	.689	.643	.554

p values for the between-treatment comparisons are given for each time point, calculated using the ANOVA test for continuous data and the chi-squared test for binary data in SPSS. Only in the EQ-5D statistically significant differences at baseline existed between the three treatment arms: the patients in the ACDF group had a significantly higher baseline score.

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty.

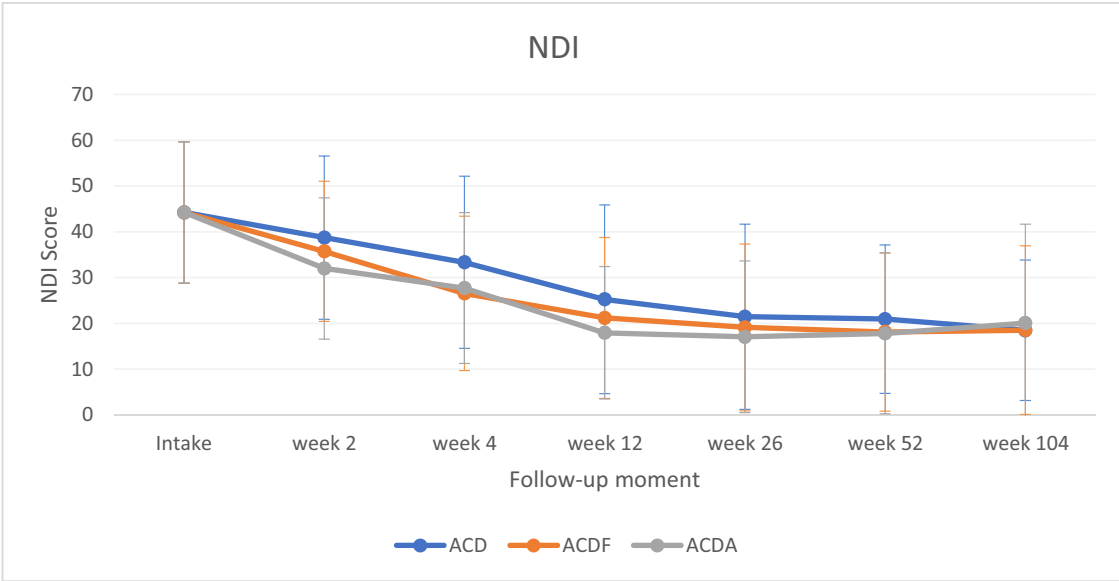

Fig. 2. Neck disability index values during follow-up.

NDI: value for intake for all three groups was set at the mean value of NDI at intake (covariate in GEE analysis) because there were no significant differences between the groups at baseline. There were no significant differences between the three groups. All three groups reach comparable values after two years.

Table 4
Estimated Marginal mean values of the NDI for all three treatment groups at baseline, after 1 year and 2 years

	Baseline		1-year FU		2-year FU	
NDI	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
ACD	42.7	1.109	20.3	1.984	18.8	2.176
ACDF	40.7	0.911	21.6	2.184	23.3	2.189
ACDA	42.1	1.336	18.5	2.800	19.7	2.371

Scores are based on the linear mixed model, in which was corrected for NDI at baseline.

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty; NDI, neck disability index.

Radiographic data

At baseline in 89 patients, a neutral lateral standing X-ray was available. Both in the superior and inferior level of the index level, osteophyte presence and formation were more prevailing than decrease in disc height (Table 6), both at baseline and after 2 years. At the superior level, ASD was present in 20%–26% of patients at baseline and in 38%–45% of patients at the 2-year follow-up point. No statistical difference was present among the three treatment arms. At the inferior level, ASD was present in 16%–20% of patients at baseline and in 29%–40% of patients after 2 years. No statistical difference was present among the three treatment arms (Table 6).

Table 5
Treatment effects of ACD, ACDF, and ACDA at 2 years follow-up

	Treatment effect	Minimal treatment effect (95% CI)	Maximal treatment effect (95% CI)	
ACD vs. ACDF	−4.493	−10.530	1.544	p = .145
ACD vs. ACDA	−0.882	−7.070	5.306	p = .780
ACDF vs. ACDA	3.611	−2.556	9.778	p = .251

Corrected for preoperative NDI score. The treatment effects are the differences in estimated marginal means between groups computed with a linear mixed model.

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty; CI, confidence interval.

For analysis of the cervical curvatures oblique view X-ray was available of an additional four patients that clearly demonstrated the cervical spine curvature. For 11 patients, a baseline X-ray was missing, but an X-ray made within 3 months after surgery was available. Therefore, for 104 patients, baseline information on the shape of the cervical spinal column in neutral position was available. Fifty-five patients demonstrated a lordotic spine, 41 patients had a straight spine, and 7 a kyphotic spine. After a follow-up period of 2 years most patients that had a lordotic or straight spine at baseline remained lordotic or straight; only one patient developed a kyphotic spine (Table 1). In the patient group with a kyphotic cervical curvature at baseline, only one patient remained kyphotic; the other six patients recovered to a straight or lordotic spine during 2-year follow-up. The number of patients with a kyphotic spine was too small to correlate this to clinical data (Table 7).

MRI analysis of ASD at baseline demonstrated that the majority of patients had severe ASD at baseline in the superior level and mild ASD at the inferior level. No statistical difference was present between the three treatment arms. At 1-year follow-up, again, ASD was severe in the superior level and mild in the inferior level in the majority of patients,

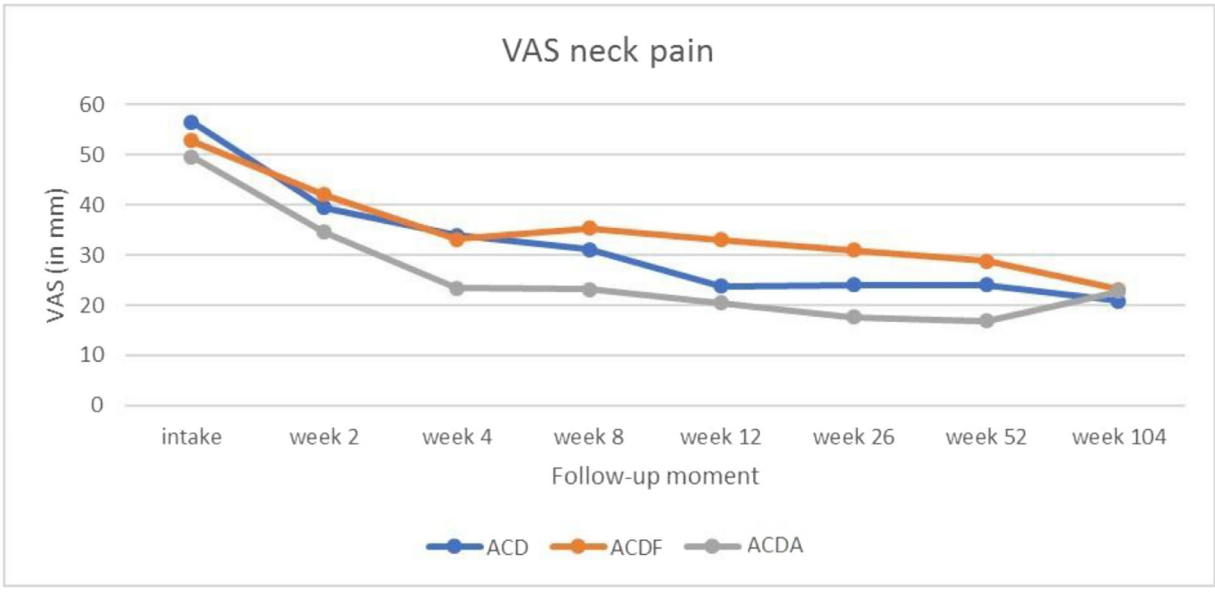

Fig. 3. VAS arm pain values during follow-up.

VAS arm pain: patients in the three groups demonstrate a decline in VAS arm pain shortly after surgery. There were no significant differences between the three groups. In contrast to the curves demonstrated for VAS neck pain, the values for VAS arm pain in the prosthesis group do not demonstrate a tendency to increase after 1-year follow-up.

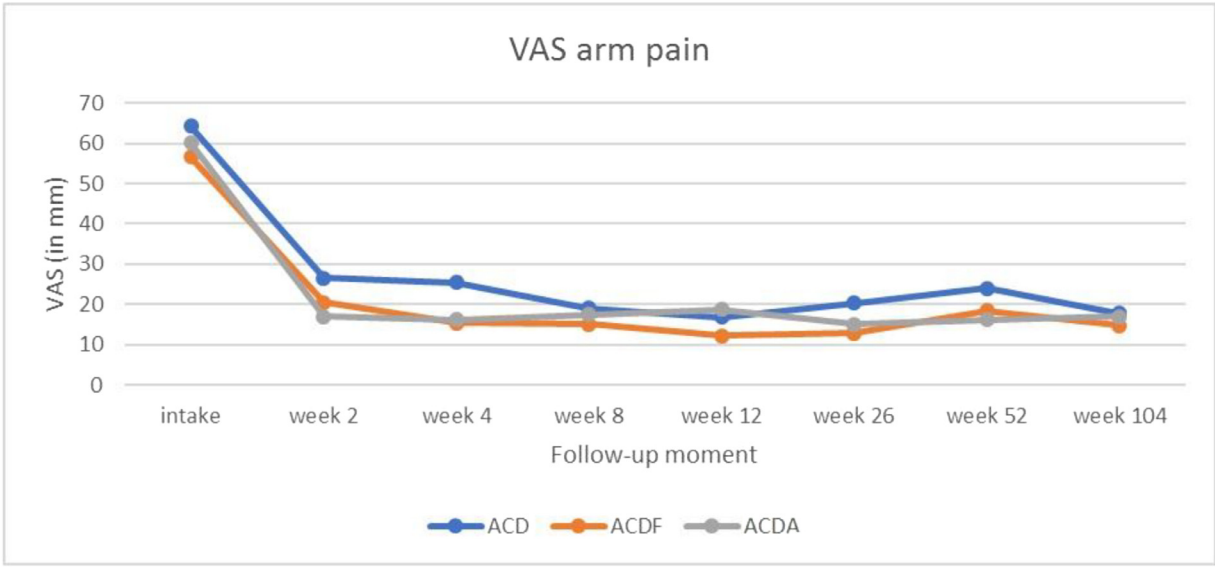

Fig. 4. VAS neck pain values during follow-up.

VAS neck pain: patients in all three treatment arms drop in VAS neck pain shortly after decompressive surgery. There were no significant differences between the three groups. In the prosthesis group, there is a tendency to an increase in neck pain after one year.

without significant differences between the groups. Before the evaluation of MRIs, the reviewers met in person to evaluate and refine the definitions. Outcome data were dichotomized as “mild ASD” (grade 1 and 2) and “severe ASD” (grade 3, 4, and 5) and compared among the three randomization arms.

Complications

Eight out of 109 patients were reoperated during the follow-up period. Four of these were operated for cervical radiculopathy due to nerve compression at an adjacent disc level with new complaints of radiculopathy indicating adjacent level disease (3 ACDF, 1ACDA). Two patients were operated on the adjacent level within 6 months of the initial surgery. They had a herniated disc at the adjacent level upon randomization, which was not associated with radiculopathy at the time of randomization. However, after surgery, complaints of the adjacent level started (1 ACD, 1 ACDA). One patient was reoperated at the index level and the adjacent level. The patient was initially

Table 6
Adjacent segment degeneration assessed with Goffin's method [26]

	ACD		ACDF		ACDA		p value (ANOVA)	
Superior level	Disc height	Osteophyte	Disc height	Osteophyte	Disc height	Osteophyte	Disc height	Osteophyte
Baseline								
ASD	4 (5.6%)	19 (26.8%)	8 (10%)	21 (26.3%)	5 (6.6%)	15 (20%)	.557	.562
Non-ASD	67 (94.4%)	52 (73.2%)	72 (90%)	59 (73.8%)	71 (934%)	60 (80%)		
2-year follow-up								
ASD	6 (10.7%)	25 (44.6%)	7 (14%)	20 (40%)	5 (8.3%)	23 (38.3%)	.635	.777
Non-ASD	50 (89.3%)	31 (55.4%)	43 (86%)	30 (60%)	55 (91.7%)	37 (61.7%)		
Inferior level								
Baseline								
ASD	5 (9.1%)	11 (20%)	5 (7.9%)	11 (17.5%)	5 (7.8%)	10 (15.6%)	1*	.822
Non-ASD	50 (90.9%)	44 (80%)	58 (92.1%)	52 (82.5%)	59 (92.2%)	54 (84.4%)		
2-year follow-up								
ASD	6 (14.6%)	15 (36.6%)	3 (7.5%)	16 (40%)	6 (11.5%)	15 (28.8%)	.618*	.510
Non-ASD	35 (85.4%)	26 (63.4%)	37 (92.5%)	24 (60%)	46 (88.5%)	37 (71.2%)		

"Mild ASD" corresponds to grade 1 and 2, and 'severe ASD' corresponds to the presence of grade 3, 4, or 5.

ACD, anterior cervical discectomy; ACDF, anterior cervical discectomy and fusion; ACDA, anterior cervical discectomy with arthroplasty; ASD, adjacent segment degeneration.

* Fisher's exact test

operated at C5–C6 (ACD) for C6 radicular symptoms and developed additional complaints in the second and third digits in the weeks after surgery. Together with persisting neck pain and slight radicular problems in the C6 area, this convinced the surgeon to implant a cage at the index level (C5–6) to correct kyphosis at the index level and to perform an additional anterior discectomy fusion at C6–7 (the adjacent level). However, complaints persisted up till 104 weeks follow-up. The other patient had persisting complaints of neck pain after having received a prosthesis. The pain was related to neck movements and subsequently the level was immobilized by adding a plate to the target level; complaints, however, persisted. Four patients experienced a subcutaneous hematoma (3.3%), one patient developed a subcutaneous infection (0.8%), nine patients experienced temporary hoarseness (7.5%), and 13 patients had temporary dysphagia (10.8%). Damage of the dura did not occur and no implant was dislocated.

Table 7
Cervical curvature at baseline and evolution during follow-up

Baseline	Follow-up	Number of patients
Lordosis	Lordosis	44
	Straight	9
	No FU	1
Straight	Straight	25
	Lordosis	8
	Kyphosis	1
	No FU	7
Kyphosis	Kyphosis	1
	Straight	5
	Lordosis	1

FU, follow-up.

Clinically relevant adjacent level disease

Since clinical data and radiological data were comparable in the three treatment arms, correlating the data was not deemed relevant. As summarized above, four patients of 109 were reoperated for adjacent level disease. Three of those patients were in the ACDF group, and one patient was in the ACDA group. None of the patients in the ACD group developed adjacent level disease that was surgically treated.

Discussion

In contrast to what is generally believed, this trial demonstrated no significant differences in clinical or radiological outcomes among single level ACDA, ACDF, and ACD. Therefore, superior outcome of the prosthesis could not be confirmed. All patients improved significantly regardless the treatment strategy.

It was demonstrated that ACD patients performed comparably well to patients that underwent ACDF or ACDA. The usual justification for an intervertebral device is that the original height of the removed disc should be restored in order to keep the neuroforamen at its original height and to prevent kyphosis. However, the results demonstrated in this study contradict that theory; only two patients developed kyphosis. Moreover, outcomes were comparable in the three groups and possible radiological differences were apparently not clinically relevant. This conclusion is in agreement with the results presented in the Cochrane review on comparison of interbody fusion techniques in the treatment of cervical radiculopathy, published in 2011 [33]. Likewise, a 10-year follow-up study of 102 patients being subjected to discectomy alone (ACD) demonstrated satisfactory results [34]. Although the size of the ACD group in the current RCT is limited, the NDI treatment effect difference between the

ACD and ACDA group was only 0.8 out of 100 (Table 3), with a maximum possible difference between ACD and the other two treatment strategies (Table 4) between −10.5 and 1.5 and −2.6 and 9.8(on a 100 point scale). This difference is hardly clinically relevant.

Specific subgroups might benefit more from receiving a prosthesis and therefore maintaining disc height. Young patients receiving a prosthesis may fare better, as they often have a more hydrated disc. Additionally, pseudarthrosis after ACD in these patients may eventually lead to kyphosis of the index level or to a significant decrease of the height of the neuroforamen. Currently, the group of patients examined is too small to do a subgroup analysis. Moreover, patients with a small disc height at baseline were not included and therefore conclusions cannot be generalized. The phenomenon should be explored in a future study.

The strong points of our study are that we maintained the double-blinded allocation of patients for 2 years, that we implanted only one type of prosthesis, that we included only patients suffering from cervical radiculopathy excluding patients suffering from myelopathy or patients with a disc height less than 3 mm, and that it is a multicenter study.

Adding our results to the literature available on the comparison of ACDF and ACDA is relevant. Several RCTs comparing ACDA and ACDF have been published, and even a few meta analyses have been performed. The meta-analyses of Zhang, Gao, and Wu demonstrated significant differences between treatment arms, but differences were hardly clinically relevant [4–6]. In addition, they included mainly studies that did not exclude myelopathy patients. These patients are prone to have more degenerative cervical spines and perform different on outcome scales. This justifies the performance of an RCT on this subject in a group of patients exclusively suffering from radicular symptoms due to a herniated disc of sufficient height.

At baseline, the patients in this RCT experienced a VAS neck pain of around 50 mm and an NDI of over 50 points, which indicates that patients experience incapacitating neck pain and disability in combination with a radiculopathy. It is therefore defendable that research is performed on preventing or diminishing neck pain and disability in this population, besides treating the arm pain. Preoperative neck pain may be due to muscle tension that is induced by the body to prevent specific neck movements irritating the compressed cervical spinal root. Degeneration of the cervical spine may also induce neck pain and disability. By decompressing the nerve, muscle tension may be lowered, but degeneration of the cervical spine cannot be removed by a surgical intervention. It may however be advantageous to prevent accelerated degeneration at the adjacent levels that is referred to as ASD. Our X-ray and MRI results demonstrate that within 2 years after surgery, ASD does not significantly differ among the three treatment arms. However, power calculation was not based on ASD differences, and only the clinical implications of possible ASD prevalence can be evaluated with sufficient statistical power. The absence of differences in VAS neck pain or NDI between the patients that received a prosthesis and those in which fusion of the segment was induced illustrates that the cervical spine apparently has compensation mechanisms to overcome the drawback of segmental fusion. This may be due to slight compensatory changes in the sagittal balance of the cervical column in order to rebalance the mechanical load.

A limitation of this study is the failure to include the 166 patients that were calculated in the power analysis. The small treatment effects however justify the publication of our results. Another limitation of this study is that it describes a follow-up period of 2 years. It is possible that adjacent level disease will occur after this period and that it has subsequent clinical effects. We do however consider a follow-up period of 2 years an appropriate length of time to compare the initial clinical outcome of the three surgical methods.

In our view, the main reason not to implant a prosthesis is a financial one, since the prosthesis is more expensive than the tested alternatives. Also, the surgical intervention to introduce a prosthesis is somewhat longer and more challenging, though it was demonstrated that this does not lead to more complications during or after surgery. Our results suggest that implanting an intervertebral device does not add to improvement of clinical outcome. To support discontinuation of the use of cages in single level cervical anterior discectomy, larger studies with subgroup analyses are needed. Furthermore, data of longer follow-up periods will be critically evaluated in order to verify if conclusions remain unchanged on the long term.

Conclusions

Notwithstanding the drawbacks of this three arm RCT, it seems that there is no strong evidence in favor of one of the three treatment strategies based on the 2-year evaluation of results. They all give comparable clinical results and all three options are acceptable.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.spinee.2018.12.013.

References

[1] Cloward RB. The anterior approach for removal of ruptured cervical disks. J Neurosurg 1958;15:602–17.
[2] Yang X, Janssen T, Arts MP, Peul WC, Vleggeert-Lankamp CLA. Radiological follow-up after implanting cervical disc prosthesis in anterior discectomy: a systematic review. Spine J 2018;18:1678–93 Epub 2018 May 8. https://doi.org/10.1016/j.spinee.2018.04.021.
[3] Bartels R, Donk R, Verbeek A. No justification for cervical disk prostheses in clinical practice. Neurosurgery 2010;66:1153–60.
[4] Gao F, Mao T, Sun W, Guo W, Wang Y, Li Z, et al. An updated meta-analysis comparing artificial cervical disc arthroplasty (CDA) versus anterior cervical discectomy and fusion (ACDF) for the treatment of cervical degenerative disc disease (CDDD). Spine 2015;40:1816–23.

[5] Wu A, Xu H, Mullinix K, Jin H, Huang Z, Lv Q, et al. Minimum 4-year outcomes of cervical total disc arthroplasty versus fusion. Medicine 2015;94:e665.
[6] Zhang Y, Liang C, Tao Y, Zhou X, Li H, Li F, et al. Cervical total disc replacement is superior to anterior cervical decompression and fusion: a meta-analysis of prospective randomized controlled trials. PLOS One 2015;10:e0117826.
[7] Hou Y, Nie L, Pan X, Si M, Han Y, Li J, et al. Effectiveness and safety of mobi-C for treatment of single-level cervical disc spondylosis: a randomised control trial with a minimum of five years of follow-up. Bone Joint J 2016;98-B:829–33.
[8] Janssen M, Zigler J, Spivak J, Delamarter R, Darden B, Kopjar B. ProDisc-C total disc replacement versus anterior cervical discectomy and fusion for single-level symptomatic cervical disc disease. J Bone Joint Surg 2015;97:1738–47.
[9] Loumeau T, Darden B, Kesman T, Odum S, Van Doren B, Laxer E, et al. A RCT comparing 7-year clinical outcomes of one level symptomatic cervical disc disease (SCDD) following ProDisc-C total disc arthroplasty (TDA) versus anterior cervical discectomy and fusion (ACDF). Eur Spine J 2016;25:2263–70.
[10] 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 2009;9:275–86.
[11] Nabhan A, Ahlhelm F, Shariat K, Pitzen T, Steimer O, Steudel W, et al. The ProDisc-C prothesis. Spine 2007;32:1935–41.
[12] Nabhan A, Steudel W, Nabhan A, Pape D, Ishak B. Segmental kinematics and adjacent level degeneration following disc replacement versus fusion: RCT with three years of follow-up. J Long-Term Eff Med Implant 2007;17:229–36.
[13] Park J, Roh K, Cho J, Ra Y, Rhim S, Noh S. Comparative analysis of cervical arthroplasty using mobi-C® and anterior cervical discectomy and fusion using the solis® -cage. J Korean Neurosurg Soc 2008; 44:217.
[14] Sala V, Lisi C, Di Natali G, Zanellato S, Dall'Angelo A, Tinelli C, et al. Functional and quality of life evaluation after single level cervical discectomy and fusion or cervical disc replacement. Giornaleitaliano di medicina del lavoro 2015;37:239–44.
[15] Zhang H, Shao Y, Chen Y, Hou Y, Cheng L, Si M, et al. A prospective, randomised, controlled multiCenter study comparing cervical disc replacement with anterior cervical decompression and fusion. Int Orthop 2014;38:2533–41.
[16] Zigler J, Delamarter R, Murrey D, Spivak J, Janssen M. ProDisc-C and anterior cervical discectomy and fusion as surgical treatment for single-level cervical symptomatic degenerative disc disease. Spine 2013;38:203–9.
[17] Arts M, Brand R, van den Akker E, Koes B, Peul W. The NEtherlands Cervical Kinematics (NECK) trial. Cost-effectiveness of anterior cervical discectomy with or without interbody fusion and arthroplasty in the treatment of cervical disc herniation; a double-blind randomised multicenter study. BMC Musculoskelet Disord 2010;11:122–30.
[18] Vernon H, Mior S. The neck disability index: a study of reliability and validity. J Manip Physiol Ther 1991;14:409–15.
[19] Pietrobon R, Coeytaux RR, Carey TS, Richardson WJ, DeVellis RF. Standard scales for measurement of functional outcome for cervical pain or dysfunction: a systematic review. Spine 2002;27:515–22.
[20] Vos CJ, Verhagen AP, Koes BW. Reliability and responsiveness of the Dutch version of the neck disability index in patients with acute neck pain in general practice. Eur Spine J 2006;15:1729–36.
[21] Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983;16:87–101.
[22] Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders. Spine 2000;25:3100–3.
[23] Ware Jr. JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
[24] Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992;305:160–4.
[25] Stansfeld SA, Roberts R, Foot SP. Assessing the validity of the SF-36 general health survey. Qual Life Res 1997;6:217–24.
[26] Goffin J, Geusens E, Vantomme N, Quintens E, Waerzeggers Y, Depreitere B, et al. Long-term follow-up after interbody fusion of the cervical spine. J Spinal Disord Tech 2004;17:79–85.
[27] Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 2001;26:1873–8.
[28] Robertson J, Papadopoulos S, Traynelis V. Assessment of adjacent-segment disease in patients treated with cervical fusion or arthroplasty: a prospective 2-year study. J Neurosurg 2005;3:417–23.
[29] Pimenta L, McAfee PC, Cappuccino A, Bellera FP, Link HD. Clinical experience with the new artificial cervical PCM (Cervitech) disc. Spine J 2004;4:315S–21S.
[30] Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopaedic conditions. J Orthop Surg Res 2015;10:15.
[31] Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. Am J Phys Med Rehabil 2010;89:831–9.
[32] Bartels RHMA, Donk RD, Verhagen WIM, Hosman AJF, Verbeek ALM. Reporting the results of meta-analyses: a plea for incorporating clinical relevance referring to an example. Spine J 2017;11:1625–32.
[33] Jacobs W, Willems P, Kruyt M, van Limbeek J, Anderson P, Pavlov P, et al. Systematic review of anterior interbody fusion techniques for single- and double-level cervical degenerative disc disease. Spine 2011;36:E950–60.
[34] Nandoe Tewarie RD, Bartels RH, Peul WC. Long-term outcome after anterior cervical discectomy without fusion. Eur Spine J 2007;16:1411–6.