Eur Spine J (2000) 9 : 129–136

ORIGINAL ARTICLE

J. van Limbeek
W. C. H. Jacobs
P. G. Anderson
P. W. Pavlov

A systematic literature review to identify the best method for a single level anterior cervical interbody fusion

Received: 10 September 1999
Revised: 14 December 1999
Accepted: 28 December 1999

J. van Limbeek (✉)
Department of Rehabilitation,
Sint Maartenskliniek Nijmegen.
P.O. Box 9011,
6500 GM Nijmegen, The Netherlands
Tel.: +31-24-3659190,
Fax: +31-24-3659238

J. van Limbeek · W. C. H. Jacobs
P. G. Anderson Sint
Maartenskliniek Research,
Sint Maartenskliniek Nijmegen.
P.O.Box 9011,
6500 GM Nijmegen, The Netherlands

P. W. Pavlov
Department of Orthopedics,
Sint Maartenskliniek Nijmegen.
P.O.Box 9011,
6500 GM Nijmegen, The Netherlands

Abstract The goal of this systematic literature review was to determine, for patients with degenerative disc disease, which method of single-level anterior cervical interbody fusion using the anterior approach gives the best clinical and radiological outcome. The number of new techniques for obtaining a solid fusion has increased rapidly, but the rationale for choosing between different techniques is unclear. Randomised comparative studies on anterior cervical interbody fusions were identified in a sensitive Medline, Cochrane and Current Contents database search. Two independent reviewers evaluated the articles that met the selection criteria, using a checklist. The search yielded eight randomised, controlled trials for the systematic literature review. Three of these studies were judged to be of sufficient quality with regard to methodology and the information provided. In the three articles, five different treatment methods were investigated, four of which were interbody fusions. Fusion rates varied between 28% for an allograft method and 63% for a discectomy-alone method. In one study, kyphosis varied from 40% to 62% between treatments. Good clinical outcome (disability, pain and symptoms) ratings varied from 66% to 82%. A meta-analysis to determine the best method for an anterior interbody fusion could not be performed due to the heterogeneity of the methods reported and because no standard outcome parameter was used. From this systematic literature review, a gold standard for the treatment of degenerative disc disease could not be identified.

Key words Cervical vertebrae · Spine · Fusion · Spinal fusion · Literature review

Introduction

Degenerative disc disease in the cervical region causes pain and discomfort in both the neck and arm [14]. When conservative treatment gives no satisfactory results, the common surgical technique to treat cervical degenerative disc disease is discectomy and fusion of the two adjacent vertebral bodies. This can be achieved solely with bone grafts, usually obtained from the iliac crest, or by implantation of a fusion cage into the intervertebral disc space. The treatment of degenerative disc diseases can be divided into posterior procedures, anterior procedures or a combination of these. The popularity of the anterior approach for discectomy and fusion has increased because this approach avoids exposingthe spinal canal [10]. The majority of the interbody fusion techniques were designed to enhance ingrowth of bone into an implant to promote fusion

The choice of the treatment to be used should preferably be based on a knowledge of the published data [4] about which of the treatments available for the type of patient presented has obtained the best results [13, 26]. A systematic literature review is the report of an appraisal of

the available published articles conducted according to a stringent methodology that includes the identification, evaluation and synthesis of studies pertaining to a topic [7]. We conducted a search using databases in Medline, Current Contents and the Cochrane library but were unable to identify any previous systematic reviews on the anterior approach for cervical interbody fusion. The goal of this systematic literature review is to determine which treatment for a single-level cervical interbody fusion using the anterior approach gives the best clinical and radiological outcome for patients with degenerative disc disease.

Methods

Search strategy

In order to obtain all the relevant literature, we used a search in the most common databases of published literature:

- The Cochrane database of randomized controlled trials (issue 1998–4)
- Current Contents (1996–April 1999)
- Medline (through Pubmed; 1966–mid 1998)

The search strings are given in Table 1, along with the number of matches for each string. We made no restrictions on language or date. A final check that no relevant articles had been missed was carried out by screening the references from the articles that were selected on the basis of the abstract.

Table 1 Number of articles found with search strings and databases used (*TW* text word, *SH* subject heading, *PT* publication type)

Search strings	Database		
	Medline	Current Contents	Cochrane
Diagnosis			
Cervical (TW)	69244	12830	2027
Cervical vertebrae (SH)	13337	53	100
Degenerative (TW)	16692	3231	295
Union diagnoses	86333	15925	2310
Treatment			
Fusion (TW)	70145	24613	702
Spinal fusion (SH)	5929	151	69
Interbody (TW)	571	246	10
Spondylodesis (TW)	332	33	1
Union treatment	69793	24646	704
Study design			
Randomised controlled trial (PT)	115988	–	3999
Controlled clinical trial (PT)	48759	–	73
Clinical trial (PT)	253753	–	18697
Multicentre studies (PT)	28872	1386	1270
Multicase review (PT)	5839	60	–
Trial (TW)	94536	47521	58848
Random*[a] (TW)	197827	86437	108759
Controlled (TW)	140883	69591	53433
Prospective*[a] (TW)	146051	34787	28616
Union study design	585841	190083	141212
Intersection	187	61	20
Intersection databases			214

[a] Search word truncated with an asterisk

Selection criteria

To answer the question which surgical method should be regarded as the gold standard for treatment of degenerative disc disease using the anterior approach for cervical interbody fusion, we used the randomised, controlled trials (RCTs) obtained from the literature search. We excluded articles that used "semi"-random techniques such as alternate appointment or birthdate-dependent allocation. RCTs were included when the article met the following criteria:

1. The intervention evaluated in the trials had to be a single-level anterior cervical interbody fusion compared with another fusion technique or any treatment for cervical degenerative disc disease.
2. The subjects to be operated on had to be suffering from degenerative disc disease at a single level. Trials involving mainly patients with fractures, tumours or multilevel disorders were excluded.
3. The outcome measurements in the studies had to be clinical or radiological. We made no exclusions on the type of these outcome measures. The required minimum follow-up had to be 6 months.

When no consensus was reached between the two reviewers, a third reviewer was consulted in a renewed attempt to achieve consensus.

Methodological quality assessment

Articles that met all the above criteria were closely examined with the aid of a checklist. Articles were evaluated on formal methodological requirements and objectives. We used criteria for evaluating studies designed to be RCTs as used by Chalmers et al. [6] and the Cochrane Collaboration. We regrouped these criteria into internal validity, external validity, data presentation and statistical analysis (Table 2). In following this procedure we based our selection of studies on methodological considerations, rather than by applying a scoring system with an arbitrary cut-off point.

The internal validity was assessed by considering the following four criteria: (1) the randomisation procedure used, (2) the homogeneity of subgroups, (3) whether the intention to treat principle had been followed, and (4) the relevance of the effect measures used. Because blinding of surgeon and patient in orthopaedic surgery is difficult, we did not take this into account. Homogeneity of subgroups was rated "+" if the subgroups were homogeneous, had comparable prognostic baseline features and if a subgroup analysis had been carried out in cases where heterogeneity had been determined. Otherwise, homogeneity was rated "–". Internal validity was rated "–" if homogeneity of subgroups was "–". If homogeneity of subgroups was "+" and one of the other criteria had either a "–" or "+/–", then a score of "+/–" was given. In all other cases, the internal validity was rated "–". In the event that only certain data in a study met the internal validity criteria, only data from that section of the study were included in the systematic literature review.

If a study had a homogeneous population, homogeneous subgroups, or when a subgroup analysis was properly carried out in cases of heterogeneity, the study was included. We included studies with moderate external validity because the results of these studies can be of value provided critical comments are given concerning the generalisation of the results.

Table 2 Methodological aspects rated to assess validity of the individual studies

Aspect	Question
Internal validity	
Homogeneity	Are the treatment groups homogeneous?
Baseline features	Are the treatment groups comparable with regard to relevant prognostic baseline characteristics such as sex, age, disease, preoperative treatment?
Subgroup analysis	Is a subgroup analysis performed in case of heterogeneity?
Homogeneity subgroups	Composed on the basis of the above three questions: Are the subgroups homogeneous and comparable as to baseline characteristics, or is a subgroup analysis performed in cases of heterogeneity?
Randomisation procedure	Is the study randomised; is the randomisation procedure described and valid?
Intention to treat	Has the intention to treat principle been met?
Relevant measure of effect	Is the measure of effect relevant?
External validity	
Inclusion and exclusion criteria	Are the inclusion and exclusion criteria described and do they allow generalisation?
Description methods	Are the methods used described in sufficient detail?
Lost to follow-up	Is the percentage of patients lost to follow-up given and acceptable?
Data presentation	
Group sizes	Are the group sizes presented in sufficient detail?
Means/percentages	Are point estimates presented in sufficient detail?
Ranges/SDs	Are the ranges of the point estimates given?
Statistical analysis	Is the statistical analysis appropriate and described in sufficient detail?

The external validity was assessed by considering how complete the description of inclusion and exclusion criteria and the treatment methods used was and how large the percentage lost to follow-up was. "Lost to follow-up" was graded "+" if the percentage was below 10%, "+/–" if it was between 10% and 20%, and "–" if it was greater than 20%.

Data presentation was rated according to whether the sizes of the groups and/or subgroups were mentioned and according to whether means and standard deviations or proportions or other relevant point estimates and their precision were presented. In addition, the appropriateness of the statistical methods was rated.

Excluding articles on these formal grounds does not necessarily imply that the individual studies were of a poorer methodological quality, but that the quality could not be assessed from the article. Because this systematic review of published literature also aimed at re-analysis of published data, only studies which adequately presented their data were included.

Publications were managed with the aid of the Reference Manager software. In addition, relevant information pertaining to database source, reason for exclusion and reviewer consensus was recorded.

Results

The literature search yielded 214 references. Two-thirds (154) of these could be excluded on the basis of the title, from which it was clear that the article was not related to anterior cervical interbody fusion. More than half of the remaining 60 references could be excluded on the basis of the abstract. Articles were excluded on the basis of the abstract and title when it was evident that the subject of investigation was not anterior cervical interbody fusion. Of the remaining 28 references, 18 had an experimental design that could not be classified as an RCT [1, 3, 5, 9, 11, 12, 15–17, 21, 22, 24, 25, 28, 29, 31, 33, 34], and one was an abstract of Madawi et al. [19].

Nine RCTs thus remained for inclusion in the third step; these were subjected to a thorough investigation with the aid of the checklist. All articles were in the English language. One article failed to meet the requirements for a valid randomisation procedure [19]. In that study, randomisation turned out to be surgeon dependent, which makes subgroup results unreliable [36]. A brief summary of the eight RCTs found in the search, including critical comments, is given below. Five of these eight RCTs failed to meet the strict methodological requirements [8, 23, 27, 30, 39]. For completeness, these five studies have been included in Table 3, which gives an overview of the checklist score for each criterion on the checklist for the eight RCTs, but they were not analysed further. The remaining three studies were of sufficient methodological quality to be included in the review [20, 32, 35].

Table 4 gives an overview of the outcome measurements used in the three selected studies. Figure 1 shows the outcome measurements for the techniques used. These outcome measurements were standardised to a 100% scale, in which 100% indicates that all patients had a good outcome and 0% indicates that no patient had a good outcome. The standard deviation, based on the binomial distribution, was calculated for the outcome measurements as:

$$SD = \sqrt{(p^*(1-p)/n}$$

The outcome evaluation can be divided in three groups: fusion and kyphosis, both determined by radiographs, and clinical outcome, determined by questionnaires. Unfortunately, the outcome scales in the clinical outcome group are not the same for each study (see Table 4). Because of these differences in the definitions, we were unable to calculate pooled results and perform a meta-analysis. However, since the definitions of the outcome levels are comparable, the clinical outcome results have been grouped together for clarity.

The standardised clinical and radiological results are shown in Fig. 1. The confidence intervals (CI) given in the following summaries of the articles are based on the formula given above.

Table 3 Rating of methodological aspects (– poor, +/– questionable, + good, *na* not applicable)

Aspect (no. of items)	Study							
	van den Bent [35]	Martins [20]	Savolainen et al. [32]	Persson et al. [27]	Rosenorn et al. [30]	McGuire and St John[23]	Dowd and Wirth [8]	Zoega et al. [39]
Internal validity								
Homogeneity	+	+[a]	+	+	–	–	–	+/–
Baseline features	+	+	+	+	–	–	+	–
Subgroup analysis	Na	+/–	na	na	+/–	–	+	–
Homogeneity subgroups (3)	+	+	+	+	–	–	–	–
Randomisation procedure	+	+	+/–	+	+/–	+/–	+	+
Intention to treat	+	+[b]	+[c]	–[d]	+	+	–	–
Relevance measure of effect	+	+	+	+	+	+/–	+	+/–
Total (4)	+	+	+	+/–	–	–	–	–
External validity								
Inclusion and exclusion criteria	+	+	+	+	+	+	+	+/–
Description methods	+	+	+	+	+	+	+	+
Lost to follow-up	+	+	+/–	+	+	+	–	–
Total (3)	+	+	+/–	+	+	+	–	–
Data presentation								
Group sizes	+/–	+	+	–	–	–	+	+/–
Means/proportions	+	+	+	+/–	–	+/–	+	+/–
Ranges/SDs	+/–	na	na	+/–	na	–	na	+/–
Total (3)	+/–	+	+	–	–	–	+	+/–
Statistical analysis	+	+	+	+/–	+	+	+/–	+/–

[a] We used the follow-up results at 6 months, for which the data are presented for levels
[b] Outcome at 1 year could not be used because of the losts to follow-up
[c] Radiological outcome at 4 years could not be used because of the numbers lost to follow-up
[d] There is no valid follow-up of at least 6 months

Table 4 Outcome measurements per study

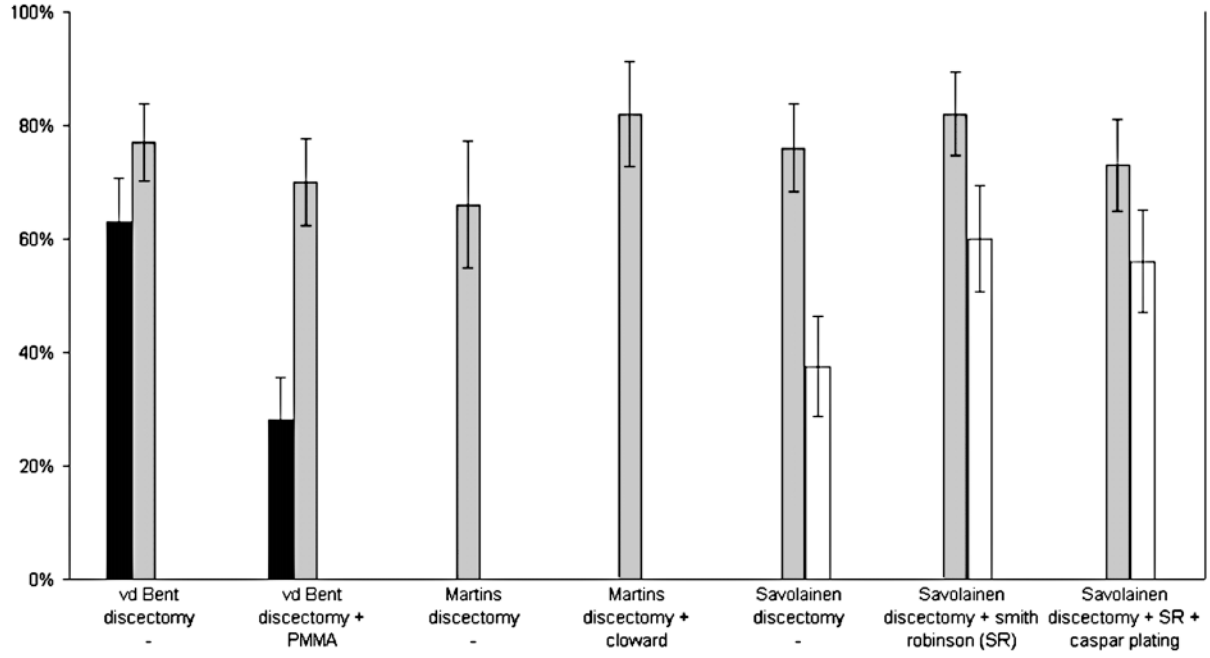
Study	Radiological measurements	Clinical outcome[a]
van den Bent et al. [35]	Plain X-ray: fusion	Clinical outcome (2): good (good or excellent on both scores); otherwise poor Odom (4): Excellent (no complaints, no impairment); good (discomfort but no limitations); satisfactory (improvement but limitations); poor (no improvement) Improvement (5): Excellent; good; moderate; hardly any; none Arm and neck pain (2): present (absent or troublesome on both scores); else not present Observer: Arm and neck pain (3): absent now and then; bothersome but not severe; disabling Patient: Arm and neck pain (3): absent now and then; bothersome but not severe; disabling
Martins [20]	Plain X-ray: alignment	Clinical outcome (4): Excellent (preoperative symptoms relieved and abnormal signs unchanged or improved); good (minimal persistence of preoperative symptoms and abnormal signs unchanged or improved); fair (definite relief of some symptoms, others unchanged or slightly improved); poor (signs and symptoms unchanged)
Savolainen et al. [32]	Plain X-ray: fusion Kyphosis	Clinical outcome (3): good (no symptoms); fair (some benefit but still complaints); poor (no benefits or worse than preoperatively)

[a] All ordinal. Number of items in parentheses

Short summaries of the studies

The eight studies that made up our final selected study material (Table 3) are briefly outlined here.

Van den Bent et al. [35] compared anterior cervical discectomy in which polymethylmethacrylate (PMMA) was used to promote fusion with discectomy alone in order to evaluate whether using PMMA improves the results of anterior discectomy. Outcome measurements were bony union on plain X-rays and a clinical outcome that was constructed on the basis of the pain scores obtained from independent reports made by the patient and an observer.



	vd Bent discectomy -	vd Bent discectomy + PMMA	Martins discectomy -	Martins discectomy + cloward	Savolainen discectomy -	Savolainen discectomy + smith robinson (SR)	Savolainen discectomy + SR + caspar plating
■ Fusion (X-ray)	63%	28%					
SD Fusion	8%	8%					
▣ Clinical outcome	77%	70%	66%	82%	76%	82%	73%
SD Clinical outcome	7%	8%	11%	9%	8%	7%	8%
□ No kyphosis (X-ray)					38%	60%	56%
SD No kyphosis					9%	9%	9%
Graft type	1	2	1	3	1	3	3
n	39	35	18	17	30	28	30

Fig. 1 Standardised outcome measures for the three studies that met the selection criteria. The treatments used are from left to right: van den Bent et al. [35]: discectomy alone and discectomy with fusion with PMMA; Martins [20]: discectomy alone and a Cloward procedure with iliac crest bone and Savolainen et al. [32]: discectomy alone, Smith Robinson with iliac crest bone and Smith Robinson with iliac crest bone with Caspar plating. The clinical outcome measures are grouped together, but they reflect different measurements. The *error bars* (see text for definition of SD) are included to indicate the expected ranges. The graft types are coded as: no graft *(1)*, allograft *(2)* or autograft *(3)*

Group sizes were 39 for discectomy alone and 42 for discectomy with PMMA. Follow-up was 2 years. One patient was lost to follow-up and marked as having a poor result. Two deaths unrelated to the discectomy procedures were recorded.

Fusion was observed in 63% of the discectomy-alone group and in 28% of the PMMA group. Clinical outcome was excellent or good in 77% (95%CI 61–89) for discectomy alone and 70% (95%CI 53–83) for discectomy with PMMA.

Some patients had two levels fused; they were equally divided among the groups. Pain was only investigated during the first 6 weeks. Standard deviations were not given for clinical outcome and pain scores. Group sizes could be reconstructed with the aid of lost-to-follow-up information. This study was considered to have sufficient internal and external validity. The data presentation could have been more informative and should have mentioned standard deviations.

Martins [20] compared discectomy alone to the Cloward fusion procedure using an iliac crest graft. Outcome measurements were based on radiographs and a four-point scale to measure clinical outcome that summarised the findings from an interview and physical examination. Follow-up was 6 months for all patients and 1 year for half of the patients. Group sizes were 25 for the fusion procedure and 25 for the discectomy-alone group, with 17 one-level procedures in the fusion group and 18 in the discectomy-alone group.

We found a percentage of two-level cases of 29%. Furthermore, only the clinical outcome was given for the one-level cases at 6 months. The Cloward fusion group had 6 excellent, 8 good, 2 fair and 1 poor result and the discectomy-alone group had 4 excellent, 8 good, 5 fair and 1 poor result. The authors did not carry out a statistical analysis on these results. In the cross-tab analysis performed by the authors of the present systematic literature review no difference was shown between the two treatments ($P = 0.65$).

Because the 1 year follow-up was incomplete (> 50% lost), these results are not suitable for the analysis. Since

the cervical alignment results are not given separately for one- and two-level procedures, these are not included in the analysis. Although this study has sufficient internal and external validity to be included in the analysis, only the results up to 6 months can be used, because these are given per number of treated levels.

Savolainen et al. [32] compared discectomy alone to two different fusion techniques: Smith-Robinson or Caspar plating. Outcome measurements were based on plain X-ray to determine fusion and kyphosis together with a clinical outcome questionnaire to evaluate benefit and complaints on a three-point scale. The total group size was 91 patients: 31 patients in the discectomy-alone group and 30 patients in each of the fusion groups. The follow-up was at 6 months and at 4 years. In respect of the radiological results, 20 patients were missing at 4 years, while for the clinical outcome data only three patients were lost for that follow-up.

Clinical outcome after 4 years showed 76% of the patients having a good result in the discectomy-alone group: 82% in the Smith-Robinson group, and 73% in the Caspar plating group. Prolonged severe iliac crest pain was observed in a total of five patients in the fusion groups.

The randomisation procedure was judged to be a block randomisation, although this was not mentioned in the text. The groups were homogeneous because the inclusion and exclusion criteria were well defined; the description of the group characteristics confirms this. The intention-to-treat principle was not met. A large number of patients were lost to follow-up for the radiological results at 4 years, and therefore these results were not used in the analysis. The clinical outcome results for 4 years pertain to 88 cases and are used in the analysis. This study had sufficient internal and external validity for the first 6 months for the radiological results, and for 4 years for the clinical results, to be used in the analysis.

Persson et al. [27] compared surgery, physiotherapy and a cervical collar for the treatment of cervical radicular pain. There was extensive cross over among therapies, and thus only the assessment after 3 months met the internal validity criteria. As the follow-up for an RCT had to have been 6 months or more to be included in the systematic literature review, this study could not be included in the analysis.

Rosenorn et al. [30] compared discectomy alone to discectomy with fusion according to a Cloward procedure using freeze-dried bone grafts. The adequacy of internal validity could not be assessed from the text due to the limited information reported. For instance, the only baseline characteristics given were the patient's sex and the number of fused levels. In addition, the randomisation procedure was not specified. Therefore, this study could not be included in the analysis.

McGuire and St John [23] compared cervical interbody fusion using autologous bone graft from the adjacent vertebral body to grafts from the iliac crest. In the data presentation it is not clear to which follow-up the results and analyses refer. Results for disc height measurements are only reported for the 1-month follow-up and lack standard deviations. In addition, there was no assessment of homogeneity or comparability for baseline features, and the randomisation procedure was not described. This study could not be used for the analysis because the uninformative data did not permit us to assess homogeneity or the precise outcome results. The internal validity could therefore not be guaranteed.

Dowd and Wirth [8] compared discectomy alone to interbody fusion using iliac crest bone. The number of two-level procedures was substantial and the results were not specified for the number of levels fused. Also, the number lost to follow-up is about 25%, which could introduce considerable bias in the results. As homogeneity of the study population could not be guaranteed, this study could not be included in the analysis.

Zoega et al. [39] compared spinal fusion with plating to fusion without plating. The homogeneity could not be assessed and no subgroup analysis was done to account for possible heterogeneity. Ranges were given for pain scores, but standard deviations were not mentioned for either pain or radiological measurement. This study was not used for the analysis because of the uninformative nature of both the data pertaining to homogeneity and the presentation of results. The internal and external validity of this study could not be assessed.

Discussion

In this systematic literature review on the effectiveness of anterior cervical interbody fusions, we found eight RCTs. Three of them were of adequate methodological quality. The three articles investigated reported on five different treatment techniques for degenerative disc disease of the cervical spine. Four of these were interbody fusions. Because both the number of procedures to treat degenerative disc disease and the number of possible outcome measurements are extensive and the number of RCTs is limited, we were unable to evaluate the effectiveness of these techniques in a meta-analysis.

Methodological aspects

The wide variety of outcome measures precludes the possibility of carrying out a meta-analysis of the pooled results. None of the studies reported adequate results at the handicap level. Persson et al. [27] reported having measured the Sickness Impact Profile (SIP) and Mood Adjective Check List (MACL), but only a comparison with a reference group was reported and no results for the subgroups was given. Other clinical outcome measurements in the eight RCTs were pain scores and disability scores,

which were either used individually or as a combined score. Anatomical measurements are fusion rates, kyphosis, and rotations and translations of the fused motion segment. Zoega et al. [39] reported rotations and translations in all directions, but the question remains as to the clinical relevance of these results and how they should be interpreted.

Although variety in outcome measurements interferes with the desire to combine different studies, it is obvious that different dimensions of outcome must be evaluated. For example, both radiological findings of fusion and patients' subjective experience of amelioration of pain, disability and handicap are important factors to consider. However, from the studies investigated, we could infer that the radiological evaluation of fusion did not correlate well with the clinical findings on pain and disability. Any outcome parameter to be used in evaluation studies for a new fusion technique must relate to the aim of the device. On the basis of this systematic literature review, it seems necessary to include both anatomical and pain/disability questionnaires since it is not obvious that a good radiological outcome will guarantee a good clinical outcome. A disability or handicap questionnaire could even be considered, as was used in the study by Persson et al. [27].

Another methodological problem is that the inclusion and exclusion criteria varied substantially from study to study. Some studies did not give clear inclusion and exclusion criteria [20, 23, 39], some studies included trauma in the study population [30], the duration of the complaints varied [8], and multi-level fusions were frequently included [8, 20, 23, 30, 35]. This raises the question whether the same base population was being treated. The data presentation is not adequate in all the studies. One study [23] gave no indication of confidence interval, range or standard error for the non-proportional results. The intention-to-treat principle was usually not met. Also, subgroup analyses were never carried out to compensate for possible heterogeneity of the base population. To carry out a meta-analysis, it is necessary to be able to extract means and standard deviations and to identify the results for homogeneous subgroups.

Best evidence synthesis

Although no meta-analysis could be performed, we have tried to synthesise some conclusions from the studies with the best methodological quality. We could draw the following conclusions, which are independent of those made by the respective authors.

The study with the most informative and the best methodological score on internal validity is the study by van den Bent et al. [35]. They reported a high rate of non-union in cases with PMMA fusion after discectomy. This was probably the consequence of the fibrous cement–bone interface that forms after implantation, and the fact that bone cannot invade the allograft. From the study by Martins[20] we cannot conclude that there is a difference between discectomy alone and addition of an autograft in a Cloward procedure. From the study by Savolainen et al. [32], it seems that additional plating does not improve outcome.

Limitations of the study

A weak step in our research might be the sensitivity our search. We did not use EMBASE, as is advised for systematic reviews in the medical field [2, 37]. EMBASE, however, primarily has a focus on pharmacological publications [13]; therefore, we do not expect to find additional studies for our review. Furthermore, the present search was aimed solely at identifying published literature. Therefore, in this study publication bias cannot be entirely ruled out.

For the selection of articles, we only blinded our reviewers to authors, institute and journal during the selection process based on the title and abstract. For the last step – reading of the actual articles – blinding is a labour-intensive procedure, the need for which has not yet been ascertained. We applied no blinding for the assessment of methodological quality; the need for this has not yet been established either [18, 38].

Conclusion

From this systematic literature review, no gold standard for the treatment of degenerative disc disease could be identified. The wide variety of treatments, the small group sizes and the variety in outcome measurements make it difficult to draw detailed conclusions about which technique is best for anterior interbody fusions. The need for more and better designed randomised, controlled trials in the field of the evaluation of degenerative cervical spine treatments is apparent.

References

1. An HS, Simpson JM, Glover JM, Stephany J (1995) Comparison between allograft plus demineralized bone matrix versus autograft in anterior cervical fusion. A prospective multicenter study. Spine 20:2211–2216
2. Assendelft WJ, Scholten RJ, van Eijk JT, Bouter LM (1999) De praktijk van de systematische reviews. III. Methodologische beoordeling van onderzoeken. Ned Tijdschr Geneeskd 143: 714–719
3. Bishop RC, Moore KA, Hadley MN (1996) Anterior cervical interbody fusion using autogeneic and allogeneic bone graft substrate: a prospective comparative analysis. J Neurosurg 85: 206–210

4. Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C (1999) Traditional reviews, meta-analyses and pooled analyses in epidemiology. Int J Epidemiol 28:1–9
5. Brooke NS, Rorke AW, King AT, Gullan RW (1997) Preliminary experience of carbon fibre cage prostheses for treatment of cervical spine disorders. Br J Neurosurg 11:221–227
6. Chalmers TC, Smith HJ, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A (1981) A method for assessing the quality of a randomized control trial. Control Clin Trials 2:31–49
7. Cook DJ, Sackett DL, Spitzer WO (1995) Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. J Clin Epidemiol 48:167–171
8. Dowd GC, Wirth FP (1999) Anterior cervical discectomy: is fusion necessary? J Neurosurg 90:8–12
9. Faraj AA, Webb JK (1997) Early complications of spinal pedicle screw. Eur Spine J 6:324–326
10. Fraser RD (1995) Interbody, posterior, and combined lumbar fusions. Spine 20:167S–177S
11. Geisler FH, Caspar W, Pitzen T, Johnson TA (1998) Reoperation in patients after anterior cervical plate stabilization in degenerative disease. Spine 23:911–920
12. Goffin J, Plets C, Van den Bergh R (1989) Anterior cervical fusion and osteosynthetic stabilization according to Caspar: a prospective study of 41 patients with fractures and/or dislocations of the cervical spine. Neurosurgery 25:865–871
13. Greenhalgh T (1999) How to read a paper. The basics of evidence based medicine. BMJ Publishing Group, London
14. Grob D (1998) Surgery in the degenerative cervical spine. Spine 23:2674–2683
15. Heidecke V, Rainov NG, Burkert W (1998) Anterior cervical fusion with the Orion locking plate system. Spine 23:1796–1802
16. Herkowitz HN, Kurz LT, Overholt DP (1990) Surgical management of cervical soft disc herniation. A comparison between the anterior and posterior approach. Spine 15:1026–1030
17. Hubach PC (1994) A prospective study of anterior cervical spondylodesis in intervertebral disc disorders. Eur Spine J 3:209–213
18. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 17:1–12
19. Madawi AA, Powell M, Crockard HA (1996) Biocompatible osteoconductive polymer versus iliac graft. A prospective comparative study for the evaluation of fusion pattern after anterior cervical discectomy. Spine 21:2123–2129
20. Martins AN (1976) Anterior cervical discectomy with and without interbody bone graft. J Neurosurg 44:290–295
21. Matge G (1998) Anterior interbody fusion with the BAK-cage in cervical spondylosis. Acta Neurochir (Wien) 140:1–8
22. Mayer T, McMahon MJ, Gatchel RJ, Sparks B, Wright A, Pegues P (1998) Socioeconomic outcomes of combined spine surgery and functional restoration in workers' compensation spinal disorders with matched controls. Spine 23:598–605
23. McGuire RA, St John K (1994) Comparison of anterior cervical fusions using autogenous bone graft obtained from the cervical vertebrae to the modified Smith-Robinson technique. J Spinal Disord 7:499–503
24. Mei F, Ren X, Wang W (1997) The biomechanical effect and clinical application of a Ni-Ti shape memory expansion clamp. Spine 22:2083–2088
25. Michelsen C, Jackson R, Lowe T, Farcy JP, Deinlein D (1998) A multicenter prospective study of the CD Spinal System in patients with degenerative disc disease. J Spinal Disord 11:465–470
26. Offringa M, de Craen AJ (1999) De praktijk van systematische reviews. I. Inleiding. Ned Tijdschr Geneeskd 143:653–656
27. Persson LC, Carlsson CA, Carlsson JY (1997) Long-lasting cervical radicular pain managed with surgery, physiotherapy, or a cervical collar. A prospective, randomized study. Spine 22:751–758
28. Rawlinson JN (1994) Morbidity after anterior cervical decompression and fusion. The influence of the donor site on recovery, and the results of a trial of surgibone compared to autologous bone. Acta Neurochir (Wien) 131:106–118
29. Razack N, Greenberg J, Green BA (1998) Surgery for cervical myelopathy in geriatric patients. Spinal Cord 36:629–632
30. Rosenorn J, Hansen EB, Rosenorn MA (1983) Anterior cervical discectomy with and without fusion. A prospective study. J Neurosurg 59:252–255
31. Russegger L, Monstadt H, Wenz F (1997) First experiences with a distractible titanium implant in ventral cervical disc surgery: report on 30 consecutive cases. Eur Spine J 6:70–73
32. Savolainen S, Rinne J, Hernesniemi J (1998) A prospective randomized study of anterior single-level cervical disc operations with long-term follow-up: surgical fusion is unnecessary. Neurosurgery 43:51–55
33. Schnee CL, Freese A, Weil RJ, Marcotte PJ (1997) Analysis of harvest morbidity and radiographic outcome using autograft for anterior cervical fusion. Spine 22:2222–2227
34. Silvers HR, Lewis PJ, Suddaby LS, Asch HL, Clabeaux DE, Blumenson LE (1996) Day surgery for cervical microdiscectomy: is it safe and effective? J Spinal Disord 9:287–293
35. van den Bent MJ, Oosting J, Wouda EJ, van Acker EH, Ansink BJ, Braakman R (1996) Anterior cervical discectomy with or without fusion with acrylate. A randomized trial. Spine 21:834–839
36. van Houwelingen JC (1998) Roaming through methodology. III. Randomization at the level of the physicians. (In Dutch.) Ned Tijdschr Geneeskd 142:1662–1665
37. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM (1997) Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. Spine 22:2323–2330
38. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Knipschild PG (1998) Balneotherapy and quality assessment: interobserver reliability of the Maastricht criteria list and the need for blinded quality assessment. J Clin Epidemiol 51:335–341
39. Zoega B, Kärrholm J, Lind B (1998) One-level cervical spine fusion. A randomized study, with or without plate fixation, using radiostereometry in 27 patients. Acta Orthop Scand 69:363–368