Web Appendix

1. Data Extraction

Each retrieved citation was reviewed by two independently working reviewers. Most articles were excluded on the basis of information provided by the title or abstract. Citations that appeared to be appropriate or those that could not be excluded unequivocally from the title and abstract were identified, and the corresponding full text reports were reviewed by the two reviewers. Any disagreement between them was resolved by reviewer consensus. From the included articles, the following data were extracted: patient demographics, preexisting diagnosis, instability, treatment, follow-up, fusion rate, symptoms, and change in symptoms.

2. Study Quality

Articles selected for inclusion were classified by class of evidence. The method used for assessing the quality of evidence of individual studies as well as the overall quality of the body of evidence incorporates aspects of the rating scheme developed by the Oxford Centre for Evidence-based Medicine[1] and used with modification by *The Journal of Bone and Joint Surgery American Volume* (*J Bone Joint Surg Am*),[2] precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group,[3] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[4] Each individual study was rated by two different investigators against pre-set criteria that resulted in an evidence rating (Class of Evidence I, II, III, or IV). Disagreements were resolved through discussion.

Determination of Overall Strength of Evidence

After individual article evaluation, the overall body of evidence with respect to each outcome is determined based on precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group[1] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[5] Qualitative analysis is performed considering the following AHRQ required and additional domains.[4] ►**Table 7** provides an outline of the method used to determine the final strength of evidence (SoE).

- *Risk of bias* is evaluated during the individual study evaluation described above. After individual article review, the literature evidence was rated as "HIGH" initially if the majority of the articles are Level I or II. It is rated as "LOW" if the majority were level III or lower. This is the "baseline" SoE, ►**Table 7**. The consistency, directness, precision, and subgroup effects are considered for potential "downgrading" the strength of the body of evidence (one or two levels depending on the degree and number of domain violations).

Table 1 Outcomes reported in included studies for open door versus French door cervical laminaplasty

Investigator (y) Study design CoE	Diagnosis	Intervention[a]	Severity of myelopathy (JOA, JOA recovery rate, Nurick scores)[b]	Other patient-reported outcomes (pain, SF-36, etc)	Adverse events[c]	Post-op care (collar, orthosis, mobilization)
Okada (2009) RCT CoE: II	CSM, OPLL, or CDH	Expansive open door ($n = 20$)	JOA scores • Preop: NR • 26.9 months: 14.2 ± 1.6 JOA recovery rate (mean) • 26.9 months: 52.8% ± 28.1%	Axial pain (VAS, mm) • Preop: 14.3 ± 31.0 • 26.9 months: 39.8 ± 30.7 SF-36 • Preop: NR • 26.9 months: scores for subscales NR	• Bleeding from lateral gutter: 12% (2/17) • C7 radicular pain: 12% (2/17) • C5 palsy: 6% (1/17) • Restenosis at C3–C7: 6% (1/17) • Shoulder numbness or pain: 12% (2/17) • Infection at surgical site: 6% (1/17) • Transient hemiparesis of unknown cause: 6% (1/17)	NR
		French door ($n = 20$)	JOA scores • Preop: NR • 26.9 months: 13.2 ± 2.7 JOA recovery rate • 26.9 months: 42.0% ± 35.4%	Axial pain (VAS, mm) • Preop: 32.0 ± 33.5 • 26.9 months: 26.7 ± 30.4 SF-36 • Preop: NR • 26.9 months: scores for subscales NR	• CSF leakage: 6% (1/18) • Misrecognition of surgical level: 6% (1/18)	

Table 1 (*Continued*)

Investigator (y) Study design CoE	Diagnosis	Intervention[a]	Severity of myelopathy (JOA, JOA recovery rate, Nurick scores)[b]	Other patient-reported outcomes (pain, SF-36, etc)	Adverse events[c]	Post-op care (collar, orthosis, mobilization)
Yue (2000) Retrospective cohort CoE: III	Cervical spondylosis, OPLL, cervical prolapsed intervetebral discs, post-traumatic, or spinal stenosis post Cloward's procedure	Open door ($n = 12$)	NR	NR	• Overall complications, details NR: 67% (8/12) • Severe intraoperative blood loss, resulting in hypotension, C5 nerve root damage, and subendocardial acute myocardial infarction: 8% (1/12)	NR
		French door ($n = 25$	NR	NR	• Overall complications, details NR: 16% (4/25) • CSF leakage: 4% (1/25)	
Naito (1994) Retrospective cohort CoE: III	CSM, OPLL, traumatic lesions, tumor/ miscellaneous lesions	Open door ($n = 35$), including additional simultaneous staged anterior cervical fusion due to preexisting subluxation or instability ($n = 11$ for open door and Z-plasty groups)	JOA recovery rate • 60.1 months: - Excellent: 26% (5/19) - Good: 47% (9/19) - Fair: 16% (3/19) - Poor: 11% (2/19)	NR	• Intraoperative dural tear: 3% (1/35)	Bed rest for 1 week, then patient allowed to stand and walk wearing Philadelphia color or Somi brace for 12–16 weeks.
		French door with iliac bone graft spacer ($n = 29$), including additional simultaneous posterior fusion due to preexisting subluxation or instability ($n = 13$)	JOA recovery rate • 60.1 months: - Excellent: 31% (9/29) - Good: 48% (14/29) - Fair: 10% (3/29) - Poor: 10% (3/29)	NR	• Pseudarthrosis: 17% (5/29) (all patients had > 3 involved levels)	

CDH, cervical disc herniation; CoE, class of evidence; CSM, cervical spondylotic myelopathy; f/u, follow-up; NR, not reported; OPLL, ossification of posterior longitudinal ligament; RCT, randomized controlled trial.
JOA, Japanese Orthopedic Association score; evaluation of the neurological function of patients with cervical myelopathy; range of score 0–17 points, with a lower score indicating a poor outcome.
JOA recovery rate = 100% (postoperative JOA score – preoperative JOA score) / (17–preoperative JOA score).
Nurick score: evaluates severity of myelopathy; range of score 0–5, with higher scores indicating greater severity.
SF-36: short form 36 health survey questionnaire; measures physical functioning, limitations in usual role of activities resulting from physical health problems, bodily pain, general health perceptions, vitality, social functioning, limitations in usual role activities because of emotional problems, and mental health; range of score 0–100, with a lower score indicating a poor outcome.
VAS reported on 0–100 mm scale, with higher score indicating maximum pain.
[a]Study also included a third intervention group, Z-plasty ($n = 35$) (Naito, 1994).
[b]JOA recovery rates defined as: excellent (75–100%), good (50–74%), fair (20–49%), and poor ($< 20\%$) (Naito, 1994).
[c]Some patients experienced multiple adverse events (Okada, 2009); study reports reoperation (anterior cervical fusion) for 2 patients following pseudarthrosis, but it is unclear which type of cervical laminoplasty these patients received (Naito, 1994).

Table 2 Complications reported in included studies for mini-plates versus no plates used in cervical laminaplasty

Investigator (y) Study design	Diagnosis	Intervention	Complications[a]	Post-op care (collar, orthosis, mobilization)
Wang (2012) RCT CoE: II	CSM, OPLL, or CDH	Open door with titanium mini-plates ($n = 25$)	• Bilateral shoulder pain: 4% (1/25) • C5 radiculopathy: 4% (1/25) • Numbness at right shoulder: 4% (1/25) • CSF leakage: 4% (1/25) • No failed plates Axial pain (VAS, mm) • Preop: 34.4 ± 31.5 • 21.2 months: 27.2 ± 30.4	• Patients with plates: collar worn for 2 weeks, then gradual mobilization in flexion-extension, rotation, and side bending • Patients with no plates: collar worn for 6 weeks, then gradual mobilization
		Open door with sutures ($n = 24$)	• Bilateral shoulder pain:8% (2/24) • C5 radiculopathy: 13% (3/24) • C7 radiculopathy: 4% (1/24) • CSF leakage: 4% (1/24) • Restenosis at C3-C7, resulting in ACDF for 2 patients: 13% (3/24) Axial pain (VAS, mm) • Preop: 30.3 ± 32.0 • 21.2 months: 38.8 ± 30.2	
Jiang (2012) Retrospective cohort CoE: III	CSM with multilevel spinal stenosis or OPLL	Open door with titanium plates ($n = 32$)	• Axial pain: 38% (12/32) • Transient C5 palsy: 3% (1/32) • Superficial wound infection: 6% (2/ 32) • No failed plates or backed out/broken screws • No restenosis due to door reclosure	• Patients allowed to sit up or walk between 3–5 days postoperative • Cervical brace worn for 3 months
		Open door with sutures ($n = 17$)	• Axial pain: 35% (6/17) • Transient C5 palsy: 6% (1/17) • Superficial wound infection: 12% (2/ 17) • Cardiopulmonary event: 6% (1/17) • No restenosis due to door reclosure	

ACDF, anterior decompression and fusion; CDH, cervical disc herniation; CSF, cerebrospinal fluid; CSM, cervical spondylotic myelopathy; f/u, follow-up; MSCS, multisegmental cervical spondylosis; NR, not reported; n/a, not applicable; OPLL, ossification of posterior longitudinal ligament; RCT, randomized controlled trial.

VAS reported on 0–100 mm scale, with higher score indicating maximum pain.

[a]It is unclear whether patients experienced multiple complications (Jiang, 2012; Wang, 2012).

Table 3 Summary of inclusion and exclusion criteria

	Inclusion	Exclusion
Patient	Adults with cervical myelopathy, including: • CSM, or • OPLL	• Patients < 18 years of age • Tumor • Trauma • Infection • Deformity • Pathologic
Intervention	**Key question 1**: Cervical laminaplasty using open door technique **Key question 2**: Cervical laminaplasty using mini-plates	Fusion, anterior-posterior surgery, including corpectomy and laminotomy
Comparison	**Key question 1**: Cervical laminaplasty using French door technique **Key question 2**: Cervical laminaplasty without use of mini-plates	
Outcome	**Key Question 1:** • JOA, mJOA, JOA recovery rate • NDI • Nurick score • Axial pain • SF-36 • Complications **Key Question 2:** • Complications, including infection, neck pain, neurological complications, nonunions, and kyphosis; re-operations. **Key Question 3 (for all KQ1 and KQ2 studies)** • Postoperative care, including exercise, early cervical motion, and use of collar	• Radiographic outcomes (other than nonunion and kyphosis) • Motion/kinetics
Publication	• Peer-reviewed studies written in English • Comparative studies	• Abstracts, editorials, letters • Duplicate publications of the same study that do not report on different outcomes • Single reports from multicenter trials • White papers • Meeting abstracts, presentations, or proceedings • Narrative reviews • Articles identified as preliminary reports when results are published in later versions
Study design	• Comparative studies • At least 5 patients per treatment group	• Case series • Case reports • Comparative study with less than 5 patients per treatment group

Criteria Evaluated for "Downgrading"

- *Consistency* refers to the degree of similarity in the effect sizes of different studies within an evidence base. If effect sizes indicate the same direction of effect and if the range of effect sizes is narrow, an evidence base was judged to be consistent. If meta-analyses were conducted, we evaluated the consistency with an "eye ball test." This test consists of a visual appraisal of the forest plots by two independent reviewers. Single study evidence bases were judged "consistency unknown (single study)" and downgraded.
- *Directness is* concerned with whether the evidence being assessed reflected a single, direct link between the interventions of interest and the ultimate health outcome; that is, a determination of whether the most clinically relevant outcome was measured or if a surrogate outcome was assessed. Directness also applies to indirect comparisons of treatment when head to head comparisons of interest could not made within individual studies.
- *Precision* of evidence pertains to the degree of certainty surrounding an estimate of effect for a specific outcome. This is based on whether the estimate of effect reached statistical significance and/or the inspection of confidence intervals around effect estimates. When there are only two subgroups, the overlap of the confidence intervals of the summary estimates of the two groups is considered. No

Table 4 Critical appraisal for studies comparing open door with French door cervical laminoplasty

Methodological principle	Okada (2009)	Yue (2000)	Naito (1994)
Study design			
Randomized controlled trial	√		
Prospective cohort study			
Retrospective cohort study		√	√
Case-control			
Case-series			
Random sequence generation[a]			
Statement of concealed allocation[a]			
Intention to treat[a]			
Independent or blind assessment			
Co-interventions applied equally	√	√	√
Complete follow-up of ≥80%	√	√	√
Adequate sample size	√		
Controlling for possible confounding[b]			
Evidence level	**II**	**III**	**III**

[a]Applies only to randomized controlled trials.
[b]Groups must be comparable on baseline characteristics or evidence of control for confounding presented.
Blank cells indicate that the criterion was either not met or that it could not be determined.

Table 5 Critical appraisal for studies comparing the use of plates with no plates for cervical laminoplasty

Methodological principle	Wang (2012)	Jiang (2012)
Study design		
Randomized controlled trial	√	
Prospective cohort study		
Retrospective cohort study		√
Case-control		
Case-series		
Random sequence generation[a]		
Statement of concealed allocation[a]		
Intention to treat[a]		
Independent or blind assessment		
Co-interventions applied equally		√
Complete follow-up of ≥80%	√	√
Adequate sample size		
Controlling for possible confounding[b]	√	
Evidence level	**II**	**III**

[a]Applies only to randomized controlled trials.
[b]Groups must be comparable on baseline characteristics or evidence of control for confounding presented.
Blank cells indicate that the criterion was either not met or that it could not be determined.

Table 6 Definition of class of evidence for articles on therapy

		Studies of therapy	
Class	**Bias risk**	**Study design**	**Criteria**
I	**Low risk** Study adheres to commonly held tenets of high-quality design, execution, and avoidance of bias	Good quality RCT	• Random sequence generation • Allocation concealment • Intent-to-treat analysis • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size
II	**Moderately low risk** Study has potential for some bias; study does not meet all criteria for class I, but deficiencies not likely to invalidate results or introduce significant bias	Moderate- or poor-quality RCT	• Violation of one of the criteria for good-quality RCT
		Good-quality cohort	• Blind or independent assessment in a prospective study, or use of reliable data[a] in a retrospective study • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size • Controlling for possible confounding[b]
III	**Moderately high risk** Study has significant flaws in design and/or execution that increase potential for bias that may invalidate study results	Moderate- or poor-quality cohort	• Violation of any of the criteria for good-quality cohort
		Case-control	• Any case-control design
IV	**High risk** Study has significant potential for bias; lack of comparison group precludes direct assessment of important outcomes	Case series	• Any case series design

[a]Outcome assessment is independent of health-care personnel judgment. Reliable data are data such as mortality or re-operation.
[b]Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

overlap of the confidence intervals indicates statistical significance, but the confidence intervals can overlap to a small degree and the difference still is statistically significant.

- *Subgroup effects*. For evaluating subgroup effects (i.e., heterogeneity of treatment effects), we downgrade if the authors do not state *a priori* their plan to perform sub-group analyses and if there was no test for interaction.

Table 7 Methodology outline for determining overall strength of evidence (SoE)

All AHRQ "required" and "additional" domainsa are assessed. Only those that influence the baseline grade are listed in table.
Baseline strength: Risk of bias (including control of confounding) is accounted for in the individual article evaluations. HIGH = majority of articles Level I/II. LOW = majority of articles Level III/IV.
DOWNGRADE: Inconsistencyb of results (1 or 2); Indirectness of evidence (1 or 2); Imprecision of effect estimates (1 or 2); Sub-group analyses not stated apriori and no test for interaction (2)
UPGRADE: Large magnitude of effect (1 or 2); Dose response gradient (1)

Outcome	Strength of evidence	Conclusions and Comments	Baseline	DOWNGRADE	UPGRADE
Outcome	HIGH	Summary of findings	HIGH Level I/II studies	NO consistent, direct, and precise estimate	NO
Outcome	MODERATE	Summary of findings	LOW Level III studies	NO consistent, direct, and precise estimates	YES Large effect
Outcome	LOW	Summary of findings	HIGH Level I/II studies	YES (2) Inconsistent Indirect	NO

[a]Required domains: risk of bias, consistency, directness, precision. Plausible confounding that would decrease observed effect is accounted for in our baseline risk of bias assessment through individual article evaluation. Additional domains: dose-response, strength of association, publication bias.
[b]Single study = "consistency unknown".

Table 8 Evidence summary

Baseline quality: **HIGH = majority of articles Level I/II. LOW = majority of articles Levels III/IV** UPGRADE: **Large magnitude of effect (1 or 2 classes); dose response gradient (1 class)** DOWNGRADE: **Inconsistency of results (1 or 2 classes); indirectness of evidence (1 or 2 classes); imprecision of effect estimates (1 or 2 classes)**					
Outcomes	**Strength of evidence**	**Conclusions/comments**	**Baseline**	**UPGRADE (classes)**	**DOWNGRADE (classes)**
In adult patients with cervical myelopathy, what is the comparative effectiveness of open door versus French door cervical laminoplasty?					
Improvement in myelopathy	Low	Overall, data from one CoE II and two CoE III studies suggest that there is no difference between treatment groups in improvement in myelopathy. All three studies found no significant difference in improvement in myelopathy measured by JOA score and JOA recovery rate.	Low	No	No
Pain	Insufficient	There is insufficient strength of evidence on the comparative effectiveness of open versus French door laminoplasty regarding pain based on the results of one study. A CoE II RCT reported significant improvement in axial pain following French door laminoplasty compared with open door laminoplasty.	Low	No	(1) Inconsistency of results[a]
Healthcare-related quality of life	Insufficient	There is insufficient strength of evidence on the comparative effectiveness of open versus French door laminoplasty regarding healthcare-related quality of life based on the results of one study. A CoE II RCT reported significantly higher SF-36 scores in four subscales following French door laminoplasty compared with open door laminoplasty.	Low	No	(1) Inconsistency of results[a]
Complications	Low	Overall, data from one CoE II and three CoE III studies suggest that the incidence of complications appears to be higher in the open door laminoplasty group compared with the French door group. One CoE III study reported a higher overall incidence of complications in the open door group (67%) compared with the French door group (16%). Although complete reporting of complications was poor, incidence of pain, neurological complications, infection, bleeding, and restenosis appeared to be higher in the open door treatment group.	Low	No	No

In adult patients with cervical myelopathy, are postoperative complications, including pain and infection, different for the use of mini-plates versus the use of no plates following cervical laminoplasty?					
Complications	Low	Overall, data from one CoE II RCT and one CoE III retrospective cohort study suggest that the incidence of complications appears to be higher in the no plate treatment group compared with the mini-plate group. In both studies rates of reoperation, radiculopathy, and infection were higher in the no plate group. In one study patients in the no plate group experienced significantly greater pain as measured by the VAS score compared with the mini-plate group.	Low	No	No
Are results from cervical laminoplasty (open door compared with French door and the use of mini-plates compared with no plates) altered by early active postoperative cervical motion?					
Open door versus French door	Insufficient	No evidence available.	Insufficient		
Use of mini-plates versus no plates	Insufficient	There is insufficient strength of evidence on the effect of early cervical motion on postoperative axial pain. Although neither study conducted a formal analysis of this effect, evidence from one study suggests that earlier postoperative cervical motion might have an effect on pain. One RCT reported that mini-plate patients, who wore a collar for two weeks, experienced significantly less pain at follow-up than the no plate patients, who wore a collar for six weeks.	Low	No	(1) Inconsistency of results[a]

CoE, class of evidence; JOA, Japanese Orthopedic Association score; RCT, randomized controlled trial; VAS, visual analog scale.
[a]Consistency of results is unknown as it is based on a single study.

Table 9 Excluded articles and reason for exclusion

Author	Reason for exclusion
KQ1: open door versus French door laminaplasty	
Asgari S, Bassiouni H, Massoud N, Schlamann M, Stolke D, Sandalcioglu IE. Decompressive laminoplasty in multi-segmental cervical spondylotic myelopathy: bilateral cutting versus open-door technique. Acta neurochirurg-ica. Jul 2009;151(7):739–749; discussion 749	Patients received re-capping laminaplasty
Kaner T, Sasani M, Oktenoglu T, et al. Clinical outcomes following cervical laminoplasty for 19 patients with cer-vical spondylotic myelopathy. Turk Neurosurg 2009;19:121–6	Only one patient underwent French door laminoplasty
KQ2: plates versus no plates in cervical laminaplasty	
Agrawal D, Sharma BS, Gupta A, et al. Efficacy and results of expansive laminoplasty in patients with severe cervical myelopathy due to cervical canal stenosis. Neurol India 2004;52:54–8	Only two patients received non-plate cervical laminoplasty
Asgari S, Bassiouni H, Massoud N, et al. Decompressive laminoplasty in multisegmental cervical spondylotic myelopathy: bilateral cutting versus open-door technique. Acta Neurochir (Wien) 2009;151:739–49; discussion 49	All patients received mini-plates

Criteria used for "Upgrading"

- Finally, if the SoE is less than "HIGH," we "upgrade" the evidence if there is a dose–response association or a strong magnitude of effect.

Strength of Evidence for Existing Systematic Reviews

Level of evidence ratings for Cochrane reviews and other systematic reviews are assigned a baseline score of HIGH if RCTs were used, LOW if observational studies were used. The rating can be upgraded or downgraded based on adherence to the core criteria for methods, qualitative, and quantitative analyses for systematic reviews (there is a reference/evaluation table for this).

The following four possible levels and their definition are reported:

- **High**: High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- **Moderate**: Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- **Low**: Low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and likely to change the estimate.
- **Insufficient**: Evidence either is unavailable or does not permit a conclusion.

References

1 Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ 2004; 328(7454):1490

2 Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. J Bone Joint Surg Am 2003;85-A(1):1–3

3 Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64(4):401–406

4 Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(12)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2012. Available at: www.effectivehealthcare.ahrq.gov

5 West S, King V, Carey TS, et al. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290–97–0011). Rockville, MD: Agency for Healthcare Research and Quality; 2002