Evidence-based Practice Center Systematic Review Protocol

Project Title: *Diagnosis and Treatment of Tethered Spinal Cord*

I. Background and Objectives for the Systematic Review

Tethered spinal cord is most commonly caused by spinal dysraphism, including myelomeningocele, lipomyelomeningocele, diastematomyelia, dermal sinus tract, and thickened/fatty filum terminale. Tethered cord syndrome is a clinical disorder associated with excessive spinal cord tension that leads to motor and sensory deficits involving the cauda equina and spinal cord. Many patients initially present in childhood, adolescence, or early adulthood due to the congenital nature of spinal dysraphism disorders. Nevertheless, patients with tethered cord syndrome can present in adulthood and later in life when there is an occult tethered cord with delayed presentation or when patients develop recurrent tethered cord syndrome after prior surgical treatments.

The condition is believed to be caused by diverse etiologies resulting in the distal spinal cord and nerve tension. The main proposed pathophysiology is the ischemic hypothesis, in which the chronic tension on the spinal cord and nerves leads to impaired local blood flow, local spinal cord ischemic injury, and neuronal damage.[1-6] Much of the ischemia hypothesis is supported by animal models, often insufficient to mimic human conditions. In vivo, the degree of tension of the conus medullaris and filum in tethered spinal cord patients has never been measured. In addition, no alteration of blood flow has ever been measured or observed in patients with tethered spinal cord as compared to normal spinal cord blood flow. Finally, there is no human histological evidence of chronic ischemia resulting from tethered spinal cord. Thus, some experts believe that the ischemic hypothesis is theoretical and unproven as the pathophysiology of tethered spinal cord.

Diagnosis of Tethered Spinal Cord

Clinical assessments and imaging are the primary modalities for diagnosing tethered cord syndrome. Patients often present with pain, motor or sensory dysfunction, or bladder and bowel functional disturbances with symptomatic tethered spinal cord. Classically, those symptoms worsen with flexion of the spine in patients with symptomatic tethered spinal cord. Additionally, patients with spina bifida occulta related spinal dysraphisms may have cutaneous stigmata that includes tufts of hair, nevi, lipoma, dermal sinuses, or hemangiomas. On imaging, patients with classical tethered spinal cord have a low-lying conus medullaris that is generally associated with thickened filum terminale, spinal lipoma with extension through the dura and into the subcutaneous fat, or adhesion of the neural placode to the dura or surrounding soft tissues. When patients become symptomatic with motor or sensory deficits, clinicians generally believe that the neurological injury is progressive and likely irreversible. Therefore, it is recommended that patients with symptomatic tethered spinal cord should proceed with treatment early in their clinical course before significant and irreversible neurological insults occur. Direct surgical detethering of the spinal cord is the standard treatment for patients with

symptomatic tethered spinal cord. On the other hand, spinal column shortening to decrease spinal cord tension has been described and utilized for treating tethered spinal cord in cases where direct detethering is not feasible or deemed to be at high risk for surgically-related neurological injuries.

Treatment of Tethered Spinal Cord

Clinicians generally believe that tethered spinal cord results in a progressive and stepwise neurological decline in patients. While this belief is widely accepted, most evidence supporting that belief is based on low-level retrospective clinical studies. Most research has supported the natural history of neurological decline. However, there are a few studies in which patients treated conservatively were followed for their natural history. In those limited studies, most patients did not develop clinical or neurological decline to impair their functions.[7, 8] Thus, the notion that patients with tethered spinal cord will surely progress and deteriorate with motor or sensory loss is not supported by high-quality prospective randomized controlled studies and may not be true for all patients with tethered spinal cord.

While patients with symptomatic tethered spinal cord may benefit from surgical treatments, the neurological risks, and peri-operative morbidity associated with surgery for tethered spinal cord are not insignificant. Surgical treatment may prevent additional motor or sensory loss associated with tethered cord syndrome. Still, direct detethering of the spinal nerves and cord also carries considerable risks to patients. Despite the common belief that patients will have progressive and irreversible motor or sensory loss with symptomatic tethered spinal cord, the time course and severity of progression for such neurological injuries for any particular patient is mainly unknown. Some patients progress rapidly with significant neurological injuries, while others may have a more insidious course and gradual stepwise neurological decline. On the other hand, it is well known that neurological injury and other complications associated with surgical treatment can be quite high.[9] Therefore, the potential benefits of any surgical treatment for tethered spinal cord, particularly prophylactic surgery for asymptomatic or marginally symptomatic patients, should be carefully weighed against the possible complication and adverse effects of surgery.

Purpose of the Review

The review will summarize the evidence regarding diagnosis, prophylactic treatment, symptomatic treatment, and repeat surgery of tethered spinal cord. With funding from the Patient Centered Outcomes Research Institute (PCORI), the Agency for Healthcare Research and Quality (AHRQ), commissioned this work to synthesize the findings on the diagnosis and treatment of tethered spinal cord. The systematic review will support the Congress for Neurological Surgeons (CNS) clinical practice guidelines.

II. Key Questions

The key questions proposed for the systematic review, addressing diagnosis (Key Question 1), prophylactic treatment (Key Question 2), symptomatic treatment (Key Question 3), and repeat surgery (Key Question 4) of tethered spinal cord, were refined following input from Key Informants and stakeholder input through public posting.

Input on the project was received through public posting of the review questions on the AHRQ website in April 2023. The posting aimed to elicit responses from stakeholders to ensure that the review is addressing the right questions, and all aspects have been considered. One of the received comments addressed the diagnosis of tethered spinal cord and noted that babies are screened with ultrasound and MRI, that myelograms are rarely used anymore, that MRI is the gold standard for diagnosis, and that a neuroradiologist should be consulted by the neurosurgeon. The second set of comments addressed that a concern of prophylactic un-tethering is the number of times the same procedure needs to be repeated, concerns regarding the standards to justify or benchmark the necessity of surgery, information given to parents, whether the information considers long-term effects, and that detethering can be associated with unintended consequences and adverse events.

We also sought input from three key informants. Key Informants included a patient with tethered spinal cord, a patient advocate from the Spina Bifida Association, and a content expert developing the planned CNS guideline. The key informants showed strong support for the importance and relevance of three of the key questions but suggested broadening key question two further to prophylactic surgery (rather than exclusively asymptomatic treatment). They emphasized the need for clinical guidance supporting patients with evidence-based information. suggested relevant references and provided important input on terminology relevant to the literature searches.

Discussions with the technical expert panel (TEP) resulted in adding a further stratification by symptom type for KQ3a to distinguish between pain and functional outcomes to the key questions. Discussions with the TEP also informed pre-planned subgroup analyses described in the data synthesis section.

Following the described input, the final key questions are as follows:

Key Question 1: What is the accuracy of radiographic and other diagnostic criteria in diagnosing tethered spinal cord?

Key Question 2: What are the benefits and harms of prophylactic surgery for asymptomatic tethered spinal cord patients?

Key Question 3: What are the effectiveness, comparative effectiveness and harms of surgical and non-surgical treatments for symptomatic tethered spinal cord?

  a. Stratified by symptom type, intensity, and patient age?
 b. Are effects modified by use of special surgical equipment or techniques?

Key Question 4: Among individuals who experience retethering after spinal detethering surgery, what are the benefits, harms and long-term outcomes of another surgery compared with no treatment?

  a. Are individual factors with which a patient presents (such as primary symptoms, symptom intensity, age, etc.) associated with better or worse outcomes after repeat surgery?

III. Logic Model

The analytic framework depicts the patient population, the interventions, and the outcomes that will be addressed in the evidence synthesis.

Figure 1: Analytic Framework
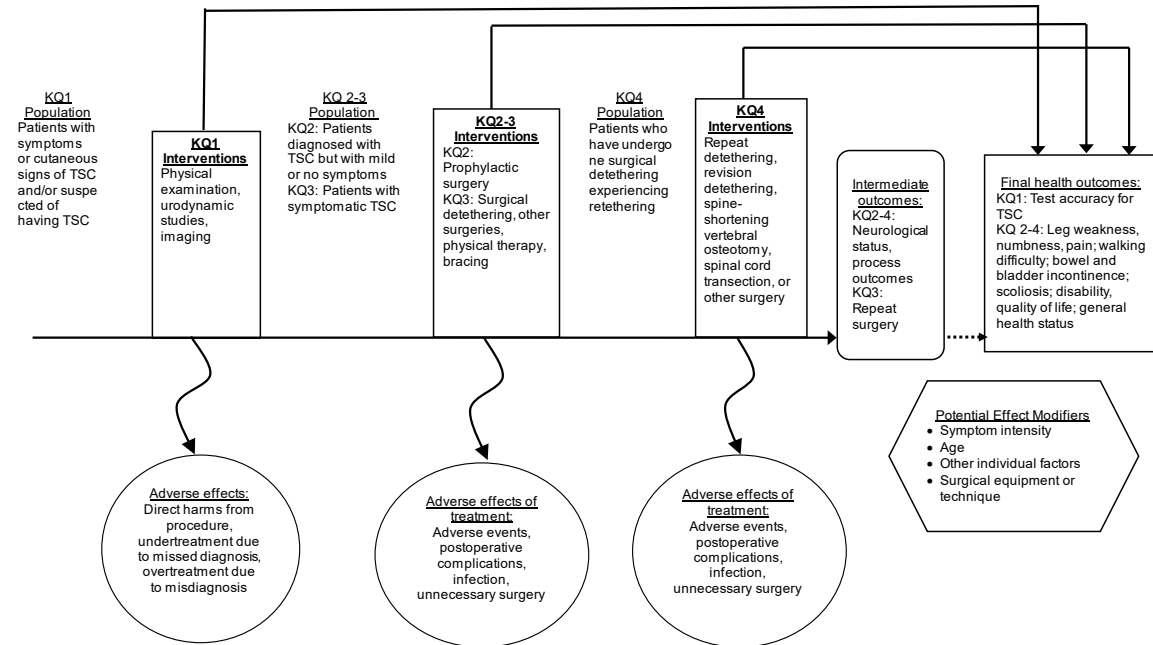
IV. Methods

Criteria for Inclusion/Exclusion of Studies in the Review

The eligibility criteria are shown in the table.

Table 1. Eligibility Criteria

Element	Inclusion Criteria	Exclusion Criteria
Population	KQ1: Pediatric or adult patients assessed for tethered spinal cord KQ2: Pediatric or adult patients with tethered spinal cord and no symptoms or marginally symptomatic without functional deficits KQ3: Pediatric or adult patients with symptomatic tethered spinal cord KQ4: Pediatric or adult patients who experience retethering after spinal detethering surgery	Tethering of the spine as an adverse event associated with an intervention (not patients being treated for tethered spinal cord)
Interventions	KQ1: Screening and diagnostic approaches, tools, and criteria such as physical examination, urodynamic studies, (MRI), myelogram, computed tomography (CT) scan, computed axial tomography (CAT) scan, or ultrasound KQ2: Prophylactic or early surgery KQ3: Surgical or non-surgical treatment or management interventions such as surgical detethering, or other surgery (e.g., spine-shortening vertebral osteotomy,	Interventions and approaches not addressing tethered spinal cord

4

Element	Inclusion Criteria	Exclusion Criteria
	spinal cord transection), physical therapy, bladder therapy for bladder function, or bracing KQ4: Surgical interventions such as repeat detethering, revision detethering, spine-shortening vertebral osteotomy, vertebral column shortening, spinal cord transection, or other surgery	
Comparators	KQ1: Confirmation of diagnosis by a neurosurgeon or neurologist KQ2-4: No surgery, sham surgery, no treatment, or alternative treatments for effectiveness outcomes; no comparator is required for studies reporting adverse events of interest (eligible adverse events will be determined with the help of the TEP)	KQ 1: no comparator For KQ 2-4, Studies without comparator except for studies for an adverse event of interest
Outcomes	KQ1: Diagnostic performance (e.g., diagnostic accuracy measured as concordance with neurosurgeon or neurologist diagnosis); adverse events of the diagnostic procedure; and clinical impact of a correct or incorrect diagnosis such as (e.g., overtreatment due to misdiagnosis, delayed treatment, or undertreatment due to missed diagnosis) KQ2-4: Patient health and other patient effects such as leg weakness, leg numbness, leg pain, other pain, gait, walking difficulty, bowel incontinence, bladder incontinence, scoliosis, disability, adverse events, postoperative complications, infection, 30-day complication rate, morbidity, quality of life, or general health status, as well as process measures such as repeat surgery	Provider satisfaction and frequency of procedures
Timing	No restrictions regarding the timing or duration of the intervention or the follow up	N/A
Setting	Settings compatible with US healthcare settings, no restrictions regarding the clinical setting	Very low resource countries or conflict zones
Study Design	KQ1: Diagnostic accuracy and diagnostic impact analyses KQ2-4: Randomized controlled trials (RCTs), clinical trials without randomization, cohort studies comparing two cohorts, controlled post-only studies, and case-control studies. Experimental single arm trials and observational case series, with or without structured pre- and post-intervention data, need to report on neurological status or bladder or bowel function to be eligible.	Secondary data, but systematic reviews will be retained for reference-mining
Other limiters	Data published in journal manuscript and trial records	Data reported in abbreviated format (e.g., conference abstracts)

Note: KQ key question, TEP technical expert panel

Relevant systematic reviews and meta-analyses will be retained as background or for reference-mining but will not be included as evidence. Publications reporting on the same participants will be consolidated into one study record. Uncontrolled studies exclusively published in non-English language publications will be excluded; controlled studies exclusively reported in non-English language publications will be assessed for applicability to the US healthcare system and the aim to support a U.S. clinical practice guideline.

Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

For primary research studies we will search PubMed (biomedical literature), EMBASE (pharmacology emphasis), CINAHL (allied nursing), Web of Science (technical innovation), and SCOPUS (general research). We will also search US and international research registries (clinicaltrials.gov, ICTRP) to capture all relevant data regardless of the publication status; increasingly, these registries include data and often provide a complete record of adverse events, making them an important evidence review tool. We will also use existing reviews for reference-mining where available. We will search the same databases used for primary research plus the Cochrane Database of Systematic Reviews and PROSPERO to systematically identify existing research syntheses. We also systematically searched for existing clinical practice guidelines, using the ECRI repository, G-I-N, MagicApp, and ClinicalKey to inform this protocol. The systematic review will include a collection of guidelines to provide further context.

We will use detailed pre-established criteria to determine eligibility for inclusion and exclusion of publications in accordance with the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews. To reduce reviewer errors and bias, all citations will be reviewed by a human reviewer and a machine learning algorithm. Citations deemed potentially relevant will be obtained as full text. Each full-text article will be independently reviewed for eligibility by two literature reviewers, including any articles suggested by peer reviewers or that arise from the public posting process, submissions through the Supplementary Evidence And Data for Systematic reviews (SEADS) portal, or responses to a Federal Register notice. Any disagreements will be resolved by consensus. We will maintain a record of studies excluded at the full-text level with reasons for exclusion.

While the draft report is posted for public comment, we will update the search and include any eligible studies identified either during that search or through peer or public reviews in the final report.

Data Abstraction and Data Management

The review team will create data abstraction forms for the key questions in DistillerSR, an online program for systematic reviews. We will abstract detailed information from controlled studies to answer the key questions. Forms will include detailed guidance to support reviewers to aid both reproducibility and standardization of data collection. Based on their clinical and methodological expertise, researchers will be assigned to abstract data from each of the eligible articles. One researcher will abstract the data, and a second reviewer will check for accuracy and completeness. Disagreements will be resolved by consensus.

We will document the diagnostic approaches and their diagnostic performance in detail for all suggested indicators to address KQ1 (diagnosis). We will abstract all reported diagnostic performance data, including false negatives and false positives, sensitivity, specificity, accuracy, area under the curve, negative and positive predictive value. In addition, we will accept other measures of concordance, including rater agreement. We will also abstract adverse events associated with the diagnostic procedure. Where reported, we will abstract information on the consequences of misdiagnosis or false positives to provide an accurate picture of the diagnostic approaches to tethered spinal cord.

We will design the data abstraction forms for this project to collect the data required to evaluate the study, as well as demographic and other data needed for determining outcomes. Abstraction categories will be informed by the National Institute of Neurological Disorders and Stroke (NINDS) spinal cord injury (SCI) common data elements.[10] Given the controversy regarding the best approach in the presence of no symptoms, we will abstract cases of prophylactic treatment (KQ2) in detail, reporting patient characteristics and context that may provide further detail on why the approach was chosen and what the observed results were.

For KQ3 (effects, comparative effectiveness and comparative safety) we will document the included patients and treatment approach in addition to the study design, analysis, and conceptual framework for measuring effects. We will pay particular attention to describing the details of the treatment (e.g., approach, surgical equipment, technique), patient characteristics (e.g., symptom intensity, age), and study design (e.g., statistical power, comparator) that may be related to outcomes. We will differentiate short-term and long-term outcomes for all studies. Studies that reported on outcomes after skeletal maturity in children and five years of follow up in adults will be considered long-term. In addition, we will carefully describe comparators, as treatment standards may have changed during the period covered by the review.

For KQ4 (repeat surgery), we will document the sequence of events in terms of timing of the surgeries and duration of follow up to clearly document the existing research evidence. Throughout KQ2-4, we will capture the treatment approach in detail so that the reader can evaluate the study results in context. This will include more information on the patients (e.g., clinical presentation) and interventions (e.g., surgical approach and experience of surgeon) than typically provided in systematic reviews, because of the complexity of condition as well as the treatment.

Data necessary for assessing quality and applicability as described in the EPC Methods Guide will also be abstracted. Forms will be pilot-tested with a sample of included articles to ensure that all relevant data elements are captured and that ambiguity is avoided. Final abstracted data will be uploaded to SRDR+.

Assessment of Methodological Risk of Bias of Individual Studies

Many different study designs are eligible for the review, hence the critical appraisal for individual studies needs to be conducted thoughtfully. We believe it is important that studies can still be compared across study designs, and we will apply a set of evaluation criteria that focuses on the underlying risk of biases, rather than applying dozens of different study design-specific tools.

For the diagnostic studies, we will apply criteria consistent with QUADAS-2. The instrument evaluates four domains: *patient selection*, *index test* characteristics, *reference standard* quality, as well as *flow and timing*[11]:

- Patient selection: The domain *patient selection* addresses whether the selection of patients could have introduced bias, taking into account whether the study enrolled a consecutive or random sample, whether the data are not based on a retrospective case-control design, and whether the study avoided inappropriate or problematic exclusions from the patient pool.

- Index test: The *index test* domain evaluates whether the conduct or interpretation of the test could have introduced bias, taking into account whether the results of the test were interpreted without knowledge of the results of the reference standard and whether any thresholds or cut-offs were pre-specified (e.g., instead of determined in the study to maximize diagnostic performance).
- Reference standard: The domain *reference standard* evaluates whether the reference standard, its conduct, or its interpretation may have introduced bias, taking into account the quality of the reference standard in correctly classifying the condition (e.g., a gold standard may not exist) and whether the reference standard test results were interpreted without knowledge of the results or index test.
- Flow and timing: The last domain, *flow and timing*, evaluates whether the conduct of the study may have introduced bias. The assessment takes into account whether the interval between the test and the reference standard was appropriate, whether all patients received the reference standard and whether they received the same reference standard, and whether all patients were included in the analysis. For each domain, we assessed the potential risk of bias in the study in order to identify high risk of bias and low risk of bias studies. Consistent with QUADAS-2,[11] the critical appraisal will evaluate for each study and appraisal domain whether there are concerns regarding the applicability of the study results to the review question. This encompassed whether the patients included in the studies do not match the review question; whether the test, its conduct, or interpretation differ from the review question; or whether the target condition as defined by the reference standard does not fully match the review question.

Throughout, the critical appraisal will be focused on how study design features may have affected the reported results. For all intervention studies we will also use a bias-focused approach, i.e., determining whether reported effects are distorted from the true value. The critical appraisal for all treatment studies will be based on the RoB 2 guidance for common sources of bias in intervention studies adapted for the eligible study designs.[11, 12] Because of the large proportion of observational studies in this topic area, assessing confounding variables will be of particular importance.

The risk of bias assessment will address selection, detection, performance, attrition, reporting, and study-specific sources of bias:

- Selection bias: For *selection bias*, we will assess the randomization sequence and allocation concealment in RCTs as well as baseline differences and potential confounders in all studies.
- Performance bias: *Performance bias* will evaluate whether patient- or caregiver knowledge of the intervention allocation or circumstances such as the trial context may have affected the outcome, and whether any deviations from intended interventions were balanced between groups.
- Attrition bias: *Attrition bias* will consider the number of dropouts, any imbalances across study arms, and whether missing values may have affected the reported outcomes.
- Detection bias: *Detection bias* will assess whether outcome assessors were aware of the intervention allocation, whether this knowledge could have influenced the

outcome measurement, and whether the outcome ascertainment could differ between arms.

- Reporting bias: *Reporting bias* assessment will include an evaluation of whether a pre-specified analysis plan exists (e.g., a published protocol), whether the numerical results likely have been selected on the basis of the results, and whether key outcomes were not reported (e.g., an obvious effectiveness indicator is missing) or inadequately reported (e.g., anecdotal adverse event reporting).
- Study-specific sources of bias: In addition to the types of bias listed above, we will assess *other potential sources of bias* such as early termination of studies, inadequate reporting of intervention details, and lack of intention-to-treat analyses.

The overall goal of the appraisal will be to identify high risk of bias studies for sensitivity analysis (e.g., to determine whether effects are primarily based on low-quality studies) as well as low-risk studies that can strengthen evidence statements through confirmation of results in strong studies. We will incorporate the risk of bias results into the strength of evidence assessment and downgrade our confidence in evidence summaries in the presence of study limitations.

Data Synthesis

We will answer each key question with the available evidence, highlighting findings from controlled studies. We will order our findings by diagnostic and treatment strategy and then by outcome.

We will determine the feasibility of a quantitative synthesis (i.e., meta-analysis) for each intervention and outcome. Feasibility depends on the volume of relevant literature, conceptual homogeneity of the studies, and completeness of the reporting of results. When a meta-analysis is appropriate, we will use random-effects models corrected for small numbers of studies where necessary to synthesize the available evidence quantitatively.[13] We will present summary estimates and 95 percent confidence intervals. We will test for heterogeneity using graphical displays and the I-squared statistic; we will highlight I-squared values that exceed 70 percent. We will explore potential sources of heterogeneity while recognizing that the ability of statistical methods to detect heterogeneity may be limited.[14]

We anticipate that the included studies and reported effects will be heterogeneous. We hypothesize that the methodological rigor of individual studies, intervention characteristics, and patients' underlying clinical presentation are potentially associated with the intervention effects. We will stratify key question 1 (diagnosis) by studies evaluating first-time diagnosis versus studies evaluating retethering. We will stratify key question 3 (treatment) by symptom type, intensity, and patient age. Furthermore, we will differentiate patients with suspected (occult tethered cord syndrome) versus confirmed tethered spinal cord syndrome We will perform meta-regression analyses for study type, intervention characteristics (diagnostic approach, surgical equipment/technique), and patient presentation characteristics, for each key question.

Regardless of the suitability for statistical pooling, all studies will be summarized in a narrative synthesis. The synthesis will be guided by the key questions, evaluated diagnostic and therapeutic interventions, and key outcomes.

Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes

The strength of evidence assessment will clearly document uncertainty, outline the reasons for insufficient evidence where appropriate, and communicate our confidence in the findings.

The strength of evidence for each body of evidence (based on the Key Question, diagnostic and treatment approach, comparator, and outcome) will be initially assessed by one researcher with experience in determining strength of evidence for each primary clinical outcome by following the principles for adapting GRADE (Grading of Recommendations Assessment, Development and Evaluation), outlined in the AHRQ methods guide.[15] The initial assessment will be discussed in the team.

We prioritized outcomes with the help of the Technical Expert Panel (TEP) in combination with team expertise. We considered outcomes most clinically relevant and important to patients and clinicians to guide clinical practice. The outcomes that will be considered for summary of findings statements are as follows:

- KQ1 (diagnosis) outcomes: any diagnostic accuracy measure most commonly reported, overtreatment or undertreatment due to misdiagnosis, clinical impact of correct or incorrect diagnosis, specificity, sensitivity, accuracy, concordance with neurosurgeon's or neurologist's diagnosis, and inter-rater reliability.
- KQ2 (prophylactic surgery) outcomes: bladder or bowel function, ambulation, quality of life, standardized symptom scores, pain, post-operative complications, number of patients with adverse events, need for repeat surgery.
- KQ3 (interventions for symptomatic patients) outcomes: neurological status, ambulation, standardized symptoms scores, pain, post-operative complications, 30-day complication rate, number of patients with adverse events, need for repeat surgery.
- KQ4 (repeat surgery) outcomes: neurological status, ambulation, bladder or bowel function, quality of life, pain, post-operative complications, 30-day complication rate, number of patients with adverse events, need for repeat surgery.

In determining the strength of a body of evidence, the following domains will be evaluated:

- Study limitations: The extent to which studies reporting on a particular outcome are likely to be protected from bias. The aggregate risk of bias across individual studies reporting an outcome is considered; graded as low, medium, or high level of study limitations.
- Consistency: The extent to which studies report the same direction or magnitude of effect for a particular outcome; graded as consistent, inconsistent, or unknown (in the case of a single study).
- Directness: Describes whether the intervention (test, treatment, or strategy) and the comparator were directly compared (i.e., in head-to-head trials) or indirectly (e.g., through meta-regressions across studies). In addition, indirectness reflects

whether the outcome is directly or indirectly related to health outcomes of interest. The domain is graded as direct or indirect.

- Precision: Describes the level of certainty of the estimate of effect for a particular outcome, where a precise estimate is one that allows a clinically useful conclusion. Graded as precise or imprecise. When quantitative synthesis is not possible, sample size and assessment of variance within individual studies will be considered.
- Reporting bias: Occurs when publication or reporting of findings is based on their direction or magnitude of effect. Publication bias, selective outcome reporting, and selective analysis reporting are types of reporting bias. Reporting bias is difficult to assess as systematic identification of unpublished evidence is challenging. If sufficient numbers of RCTs are available, we reviewed Begg and Egger tests and used trim and fill methods to assess the robustness of effect estimates.

Bodies of evidence consisting of RCTs are initially considered as high strength, while bodies of comparative observational studies begin as low-strength evidence. The strength of the evidence may be downgraded based on the limitations described above. There are also situations where observational evidence may be upgraded (e.g., large magnitude of effect, presence of dose-response relationship or existence of plausible unmeasured confounders) as described in the AHRQ Methods guides.[15]

A final strength of evidence grade will be assigned by evaluating and weighing the combined results of the above domains. To ensure consistency and validity of the evaluation, the grades will be reviewed by the entire team of investigators. The strength of evidence will be assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale:

- High: We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).
- Moderate: We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low: We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient: We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available, or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Summary tables will include ratings for individual strength of evidence domains (i.e., risk of bias, consistency, precision, directness) based on the totality of underlying evidence (i.e., the existing evidence included in the prior report in combination with newly identified studies). We will summarize updated evidence and describe what it adds to the previous review and highlight changes to the key findings.

Assessing Applicability

Applicability will be assessed in accordance with the AHRQ's Methods Guide. Factors that may affect applicability, which we have identified *a priori*, include patient, intervention, setting, and study design features.

We will address whether outcomes are different across studies that recruit different populations (e.g., age groups, clinical presentations, exclusions for comorbidities) or use different methods to implement the interventions of interest. We will use these data to evaluate the applicability to clinical practice, paying special attention to the following: study eligibility criteria; demographic features of the enrolled population in comparison to the target population; characteristics of the intervention used in comparison with care models currently in use; the possibility of diagnostic tool or treatment intervention learning curves; and clinical relevance and timing of the outcome measures.

We will use this information to assess the situations in which the evidence is most relevant and to evaluate applicability to real-world clinical practice in typical U.S. settings, summarizing applicability assessments qualitatively.

Use of Artificial Intelligence and/or Machine Learning

All citations retrieved by the literature searches will be screened by at least one human literature reviewer and a DistillerSR software machine learning algorithm trained by the human reviewers to ensure that no relevant citation will be missed. Any citations identified as potentially relevant by the algorithm that have not been selected for full text publication review will be rescreened for relevance by an independent literature reviewer.

V. References

1. Fuse T, Patrickson JW, Yamada S. Axonal transport of horseradish peroxidase in the experimental tethered spinal cord. Pediatr Neurosci. 1989;15(6):296-301. doi: 10.1159/000120486. PMID: 2484977.

2. Koçak A, Kiliç A, Nurlu G, et al. A new model for tethered cord syndrome: a biochemical, electrophysiological, and electron microscopic study. Pediatr Neurosurg. 1997 Mar;26(3):120-6. doi: 10.1159/000121176. PMID: 9419028.

3. Yamada S, Iacono RP, Andrade T, et al. Pathophysiology of tethered cord syndrome. Neurosurg Clin N Am. 1995 Apr;6(2):311-23. PMID: 7620356.

4. Yamada S, Zinke DE, Sanders D. Pathophysiology of "tethered cord syndrome". J Neurosurg. 1981 Apr;54(4):494-503. doi: 10.3171/jns.1981.54.4.0494. PMID: 6259301.

5. Yamada S, Won DJ, Yamada SM. Pathophysiology of tethered cord syndrome: correlation with symptomatology. Neurosurg Focus. 2004 Feb 15;16(2):E6. doi: 10.3171/foc.2004.16.2.7. PMID: 15209489.

6. Kang JK, Kim MC, Kim DS, et al. Effects of tethering on regional spinal cord blood flow and sensory-evoked potentials in growing cats. Childs Nerv Syst. 1987;3(1):35-9. doi: 10.1007/BF00707191. PMID: 3594467.

7. Kulkarni AV, Pierre-Kahn A, Zerah M. Spontaneous regression of congenital spinal lipomas of the conus medullaris. Report of two cases. J Neurosurg. 2004 Nov;101(2 Suppl):226-7. doi: 10.3171/ped.2004.101.2.0226. PMID: 15835113.

8. Tuuha SE, Aziz D, Drake J, et al. Is surgery necessary for asymptomatic tethered cord in anorectal malformation patients? J Pediatr Surg. 2004 May;39(5):773-7. doi: 10.1016/j.jpedsurg.2004.01.023. PMID: 15137017.

9. Pierre-Kahn A, Zerah M, Renier D, et al. Congenital lumbosacral lipomas. Childs Nerv Syst. 1997 Jun;13(6):298-334; discussion 5. doi: 10.1007/s003810050090. PMID: 9272285.

10. NINDS. NINDS Common Data Elements: Spinal Cord Injury. n.d. https://commondataelements.ninds.nih.gov/Spinal%20Cord%20Injury. Accessed on July 19, 2023.

11. University of Bristol. QUADAS-2. https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/. Accessed on February 10, 2021.

12. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019 Aug 28;366:l4898. doi: 10.1136/bmj.l4898. PMID: 31462531.

13. Rover C, Knapp G, Friede T. Hartung-Knapp-Sidik-Jonkman approach and its modification for random-effects meta-analysis with few studies. BMC Med Res Methodol. 2015 Nov 14;15:99. doi: 10.1186/s12874-015-0091-1. PMID: 26573817.

14. Hempel S, Miles JN, Booth MJ, et al. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. Syst Rev. 2013 Nov 28;2:107. doi: 10.1186/2046-4053-2-107. PMID: 24286208.

15. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Content last reviewed March 202`. Effective Health Care Program Agency for Healthcare Research and Quality. Rockville, MD: 2021. https://effectivehealthcare.ahrq.gov/products/cer-methods-guide

VI. Definition of Terms

None

VII. Summary of Protocol Amendments

None

VIII. Review of Key Questions

The Agency for Healthcare Research and Quality (AHRQ) posted the Key Questions on the AHRQ Effective Health Care Website for public comment. The Evidence-based Practice Center (EPC) refined and finalized them after reviewing of the public comments and seeking input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the Key Questions are specific and relevant.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into the decisional dilemmas and help keep the focus on Key Questions that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for the systematic review or when identifying high-priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report. They do not review the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The AHRQ Task Order Officer (TOO) and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and suggest approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The AHRQ TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers.

The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published 3 months after the publication of the evidence report.

Potential peer reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Invited peer reviewers with any financial conflict of interest greater than $5,000 will be disqualified from peer review. Peer reviewers who disclose potential business or professional conflicts of interest can submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Direct financial conflicts of interest that cumulatively total greater than $1,000 will usually disqualify an EPC core team investigator.

XIII. Role of the Funder

This project was commissioned and funded by the Patient-Centered Outcomes Research Institute (PCORI) and executed under Contract No. 75Q80120D00009 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The AHRQ Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by PCORI, the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

XIV. Registration

This protocol will be registered in the international prospective register of systematic reviews (PROSPERO).

Appendix A. Search strategies

Search date: 7/5/2023

<u>PubMed</u>
No Limits
tethered cord syndrome[Title/Abstract] OR tethered spinal cord[Title/Abstract] OR tethered cord[Title/Abstract] OR spinal column shortening[Title/Abstract] OR (spinal column shortening[Title/Abstract] AND tethered[Title/Abstract]) OR low lying conus[Title/Abstract] OR fatty filum terminale [Title/Abstract] OR thickened filum terminale[Title/Abstract] OR tight filum terminale[Title/Abstract] OR low lying spinal cord[Title/Abstract] OR filum terminale syndrome[Title/Abstract] OR lipomyelomeningocele[Title/Abstract] OR tight filum syndrome[Title/Abstract]

<u>EMBASE</u>
('tethered cord syndrome' OR 'tethered spinal cord' OR 'tethered cord' OR 'spinal column shortening' OR ('spinal column shortening' AND tethered)) NOT [medline]/lim OR ("low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome" OR "Lipomyelomeningocele" OR "tight filum syndrome") NOT [medline]/lim

<u>CINAHL</u>
Limit: Academic Journals
TI ("tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR ("spinal column shortening" AND tethered)) OR AB ("tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR ("spinal column shortening" AND tethered)) OR SU ("tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR ("spinal column shortening" AND tethered)) OR TI ("low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome OR "Lipomyelomeningocele" OR "tight filum syndrome") OR AB ("low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome OR "Lipomyelomeningocele" OR "tight filum syndrome") OR SU ("low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome" OR "Lipomyelomeningocele" OR "tight filum syndrome")

<u>Web of Science</u>
Limit: Science Citation Index Expanded, Conference Proceedings Citation Index, Emerging Sources Citation Index
"tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR ("spinal column shortening" AND tethered) (Topic) OR "low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum

terminale" OR "low lying spinal cord" OR "filum terminale syndrome" OR "Lipomyelomeningocele" OR "tight filum syndrome") (Topic)

<u>Scopus</u>
TITLE-ABS-KEY ("tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR ("spinal column shortening" AND tethered)) OR TITLE-ABS-KEY ("low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome" OR "Lipomyelomeningocele" OR "tight filum syndrome")

<u>Clinicaltrial.gov</u>
No limits
"tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR "low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome OR Lipomyelomeningocele" OR "tight filum syndrome"
Recruitment: Completed studies
Study Results: All studies
Phase: Phase 2, Phase 3, Phase 4

<u>ICTRP</u>
No limits
"tethered cord syndrome" OR "tethered spinal cord" OR "tethered cord" OR "spinal column shortening" OR "low lying conus" OR "fatty filum terminale" OR "thickened filum terminale" OR "tight filum terminale" OR "low lying spinal cord" OR "filum terminale syndrome OR Lipomyelomeningocele" OR "tight filum syndrome"
Phase: Phase 2, Phase 3, Phase 4

<u>Cochrane Database of Systematic Reviews</u>
tethered cord syndrome OR tethered spinal cord OR tethered cord OR spinal column shortening OR (spinal column shortening AND tethered) OR low lying conus OR fatty filum terminale OR thickened filum terminale OR tight filum terminale OR low lying spinal cord OR filum terminale syndrome OR lipomyelomeningocele OR tight filum syndrome

<u>PROSPERO</u>
tethered cord syndrome OR tethered spinal cord OR tethered cord OR spinal column shortening OR (spinal column shortening AND tethered) OR low lying conus OR fatty filum terminale OR thickened filum terminale OR tight filum terminale OR low lying spinal cord OR filum terminale syndrome OR lipomyelomeningocele OR tight filum syndrome
Completed Studies