



# Ceiling effects in the EOSQ-24 may limit its ability to assess treatment outcomes in early-onset scoliosis

Antti Juhani Saarinen<sup>1</sup>

Received: 29 September 2025 / Accepted: 22 January 2026  
© The Author(s) 2026

## Abstract

**Background** The 24-item Early-Onset Scoliosis Questionnaire (EOSQ-24) is a condition-specific instrument to assess health-related quality of life in children with early-onset scoliosis (EOS). Previous studies have raised concerns regarding ceiling effects, i.e. clustering of scores at the upper limit, which may reduce the ability of the EOSQ-24 to detect clinical improvements.

**Methods** A structured review of published clinical studies reporting EOSQ-24 scores was performed. Data were standardized to a 0–100 scale. For each domain weighted mean scores, standard deviations (SD), and ceiling effect percentages were calculated. A ceiling effect  $\geq 15\%$  was considered significant.

**Results** Data from 16 studies were included. Significant ceiling effects were observed in the following domains: pulmonary function (28%), physical function (25%), overall satisfaction (22%), transfer (20%), financial burden (22%), daily living (18%), fatigue/energy level (18%), and emotion (16%). Lower ceiling effects were noted in general health, pain/discomfort, parental burden, and child/parent satisfaction domains.

**Conclusions** Eight of EOSQ-24 domains exhibit marked ceiling effects, potentially limiting responsiveness to treatment effects in patients with high baseline function. These findings support consideration of revised scoring strategies or adjunctive measures in EOS outcome assessment. The presence of marked ceiling effect should be noted as a limitation in studies. Objective pulmonary function testing should be implemented in future studies.

**Keywords** Patient reported outcome measure · Methodology · Quality of life · Early onset scoliosis

## Introduction

The 24-item Early-Onset Scoliosis Questionnaire (EOSQ-24) is a condition-specific instrument to assess health-related quality of life in children with early-onset scoliosis (EOS) [1]. The EOSQ-24 can be completed by either the patient or their caregiver. The EOSQ-24 consists of 24 items across ten domains: General Health, Pain/Discomfort, Pulmonary Function, Transfer, Physical Function, Daily Living, Fatigue, Emotion, Child Satisfaction, and Parent Satisfaction. EOS can be divided into syndromic, neuromuscular, congenital, and idiopathic scoliosis [2]. Etiology is significantly associated with the EOSQ-24 scores in patients with EOS, with

children with syndromic or neuromuscular scoliosis reporting lower scores than patients with congenital or idiopathic scoliosis [3].

A ceiling effect occurs when a substantial proportion of respondents attain the highest possible score on a scale, which compresses variability at the top end and reduces sensitivity to detect further improvement [4]. In patient reported outcome measure (PROM) methodology, thresholds of 15% of respondents at the maximum are commonly used to flag a problematic ceiling effect and potential limitations in content validity and responsiveness [4].

While the EOSQ-24 has demonstrated good psychometric properties, concerns have been raised regarding ceiling effects, i.e. clustering of scores at the upper limit, which reduce sensitivity to detect change [5–10]. Ceiling effects arise when item difficulty is poorly matched to respondent ability, resulting in score clustering at the upper bound. Several EOSQ-24 domains assess overt functional limitation or severe symptom states (e.g., respiratory compromise, need

✉ Antti Juhani Saarinen  
anjusaa@utu.fi

<sup>1</sup> Department of Orthopedics and Traumatology, Turku University Hospital and University of Turku, Luolavuorentie 2, 20700 Turku, Finland

for assistance with transfers), which may be absent in many treated or higher-functioning patients. When combined with a limited five-level response structure and proxy-reported assessments, this design may insufficiently discriminate among patients with mild residual impairment, predisposing several domains to ceiling effects.

The aim of this study was to systematically review clinical studies using the EOSQ-24 and quantify domain-specific ceiling effects across published EOS populations.

## Methods

A systematic review was performed for clinical studies reporting EOSQ-24 outcomes in early-onset scoliosis. PubMed/MEDLINE, Embase, and Scopus databases were searched for “early-onset scoliosis questionnaire” OR “EOSQ-24”. Preliminary screening was performed using titles and abstracts. The inclusion criteria were clinical research setting using EOSQ-24 as an outcome. EOSQ-24 validation studies were excluded, as these patient samples may not reflect relevant clinical settings. Full texts were screened for final inclusion. If stated, the primary outcome timepoint EOSQ-24 was collected. If not specified, the most recent follow-up was considered the primary outcome for this review. When multiple groups were reported within a study, each group was treated as a separate data point. Studies reporting only subsets of EOSQ-24 items were also considered eligible. The original EOSQ-24 included two items in the Satisfaction domain: Child Satisfaction and Parent Satisfaction. However, some studies reported only a single overall satisfaction score; in such cases, the data was collected in Overall Satisfaction and were not included in Parent or Child Satisfaction domains. Each EOSQ-24 item is scored on a five-point Likert scale. Domain scores are calculated as the mean of item responses within that domain and expressed on a scale with a maximum score of 100. When EOSQ-24 domain scores were reported using score ranges other than the conventional 0–100 scale, they were linearly re-expressed on the 0–100 scale to allow comparison across studies. Weighted means for each domain were calculated using sample sizes.

For each EOSQ-24 domain, reported mean, standard deviation (SD), sample size, and scale range were extracted. When only median and interquartile range (IQR) were available, mean and SD were approximated using the method by Wan et al., where the mean is set equal to the median and SD estimated as  $IQR/1.35$  [11]. When normal distribution is assumed, the proportion of the sample expected to reach or exceed the maximum possible score (100 points) can be estimated as  $1 - \Phi((100 - \text{mean})/SD)$ , where  $\Phi$  denotes the cumulative distribution function of the standard normal distribution [12]. Ceiling proportions were then summarized

across domains. Ceiling effect of at least 15% was considered as significant. Because EOSQ-24 domain scores are ordinal, bounded, and frequently right-skewed, the assumption of normality may underestimate the true proportion of observations at the maximum score. Consequently, calculated ceiling proportions should be interpreted as conservative estimates rather than exact values.

## Results

In all, 116 references were found (Fig. 1). After the preliminary screening, 60 studies were included for the full text review. Studies that lacked SD or IQR were excluded. Two studies reported adjusted EOSQ-24 data without further explanation and were excluded from the analysis. In total, 16 studies comprising 32 groups and 357 domains met the inclusion criteria (Supplementary data) [3, 13–25]. Five studies reported medians and IQRs.

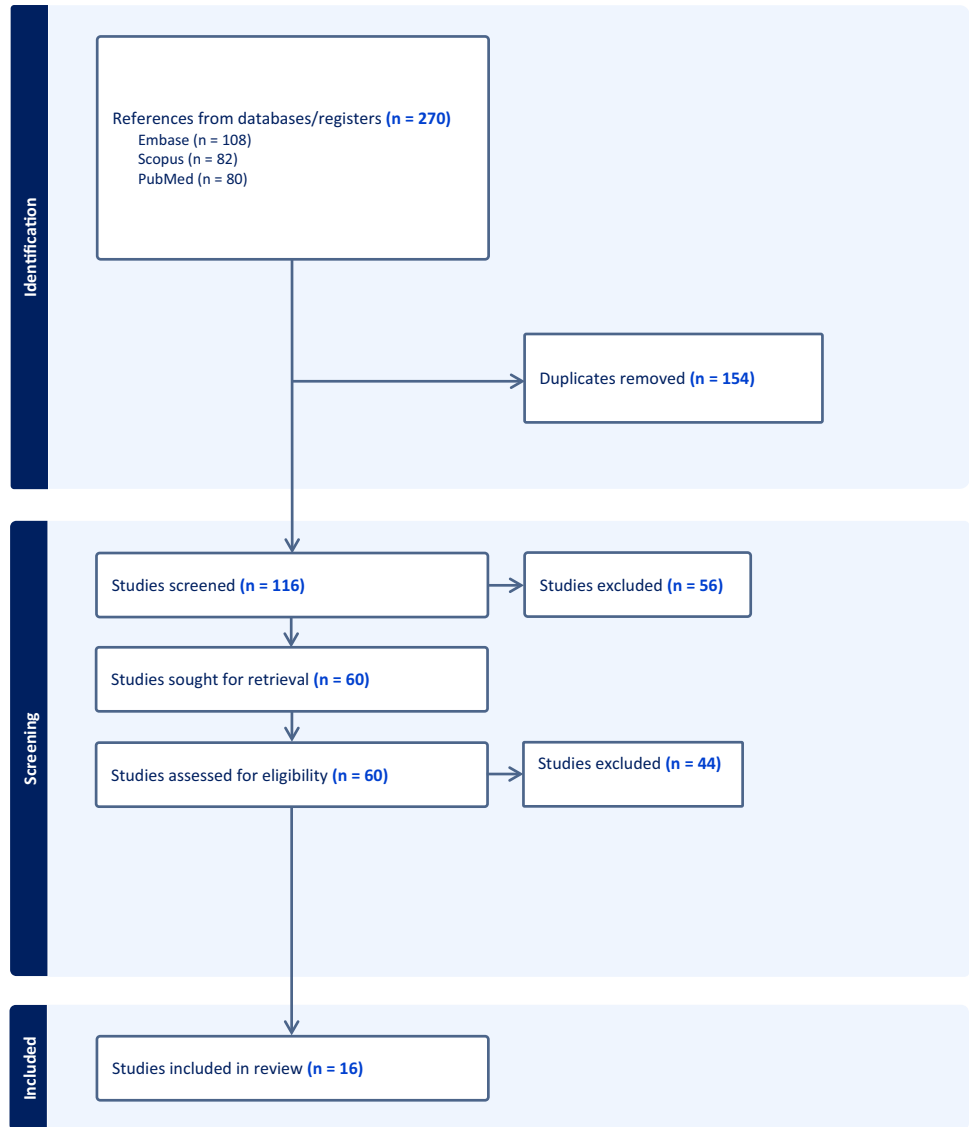
Every included study had at least one EOSQ-24 domain with a significant ceiling effect present. Pulmonary Function, Transfer, and Financial Burden items had the largest incidence of ceiling effect, and General Health, Parental Burden, and Child Satisfaction the lowest (Table 1).

When weighted means were calculated for all studies, a mean ceiling effect of at least 15% was observed in eight of the 11 analyzed domains (including Overall Satisfaction analyzed separately from Parent and Child Satisfaction): Pulmonary Function (28%), Physical Function (25%), Overall Satisfaction (22%), Transfer (20%), Financial Burden (22%), Daily Living (18%), Fatigue/Energy Level (18%), and Emotion (16%) (Table 1). In terms of frequency, marked proportion of the study groups showed ceiling effects  $\geq 15\%$  in Pulmonary Function (97% of the groups), Transfer (79% of the groups), Financial Burden (77% of the groups), and Physical Function (71% of the groups), whereas General Health (16% of the groups) and Parental Burden (17% of the groups) were least affected (Table 1).

## Discussion

In this review, the EOSQ-24 demonstrated substantial ceiling effects across multiple domains. Ceiling effects of  $\geq 15\%$  were identified in eight of the domains: Pulmonary Function, Transfer, Physical Function, Daily Living, Fatigue/Energy Level, Emotion, Financial Burden, and Overall Satisfaction. Domains such as Pain/Discomfort and General Health demonstrated lower ceiling effects. These domains include items assessing subjective symptom intensity and overall health perception across a broader severity range, which may improve person–item targeting and preserve variability at higher functioning levels. These results indicate that ceiling

**Fig. 1** Flowchart of the study selection process



**Table 1** Weighted mean scores and mean ceiling effects for each EOSQ-24 item

Item	Mean score	SD	Mean ceiling effect	SD	Percentage of study groups with ceiling effect minimum of 15%
General health	74.1	5.8	7.7	5.0	16
Pain/discomfort	73.8	8.0	13.0	7.8	38
Pulmonary function	88.3	7.3	28.1	11.8	97
Transfer	79.6	12.7	20.4	13.6	79
Physical function	76.6	16.5	24.7	14.8	71
Daily living	66.3	21.6	17.5	15.1	55
Fatigue and energy level	74.0	10.5	17.9	11.3	55
Emotion	76.7	8.4	15.9	7.9	44
Parental burden	71.6	10.8	9.4	9.7	17
Financial burden	78.3	10.8	22.1	11.4	77
Child satisfaction	70.1	8.1	12.1	7.6	25
Parental satisfaction	71.0	8.1	13.7	7.7	42
Overall satisfaction	69.5	28.8	21.9	22.1	33

effects are widespread in the EOSQ-24 and may limit its ability to detect treatment-related improvements, particularly in higher-functioning patients. Notably, Pulmonary Function, Transfer, Financial Burden, and Physical Function exceeded the 15% threshold in most study groups, underscoring the consistency of these findings across studies.

The Pulmonary Function domain demonstrated the most pronounced ceiling effect, with nearly all study groups exceeding the 15% threshold. Items within this domain primarily assess overt respiratory compromise rather than exertional limitation, nocturnal symptoms, or early functional decline. As a result, many patients without advanced pulmonary impairment may cluster at the maximum score despite clinically meaningful differences in respiratory status. This ceiling clustering likely reflects item content rather than absence of pulmonary morbidity and limits the domain's sensitivity in higher-functioning patients. The EOSQ-24 Pulmonary Domain may therefore be most informative in patients with severe neuromuscular or syndromic EOS, while offering limited discrimination in idiopathic or moderately affected cohorts. Several previous studies have noted the marked ceiling effect in the Pulmonary Function domain [5, 8–10]. Clinical pulmonary function testing should be implemented to adequately evaluate pulmonary function related outcomes in patients with EOS [26].

The presence of ceiling effects beyond the commonly accepted 15% threshold raises concerns about the responsiveness of the EOSQ-24. Once patients score near the maximum, further improvements cannot be captured, potentially underestimating treatment benefits. Moreover, when marked ceiling effects are present, comparisons between groups may become unreliable, as differences in outcomes cannot be fully expressed. When baseline scores approach the maximum, even clinically meaningful improvement cannot be captured, reducing standardized response means and effect sizes and increasing the risk of false-negative conclusions. In clinical decision making, this may obscure treatment benefits, particularly when comparing surgical techniques or longitudinal outcomes in higher-functioning patients. These limitations are consistent with those described for other PROMs and highlight the need to complement EOSQ-24 with objective measures to ensure adequate assessment across the full spectrum of disease severity.

This review has limitations. In some studies, mean and SD were approximated from medians and IQRs, which may have introduced imprecision; however, the principal conclusions of this study remain unaffected. Heterogeneity in study populations and designs may also have contributed to variability. The  $\geq 15\%$  threshold was adopted a priori based on widely used PROM quality criteria rather than derived empirically. Proxy reporting may contribute to ceiling effects, as caregivers may be less sensitive to subtle functional or symptomatic changes, particularly in

higher-functioning children. Respondent type was inconsistently reported and rarely stratified across studies, precluding quantitative comparison of ceiling effects by respondent. Some studies reported aggregated satisfaction scores that likely represent mixed constructs. Although analyzed separately, this heterogeneity may obscure domain-specific ceiling behavior. The reliance on distribution-based estimation rather than individual-level data introduces uncertainty, particularly given the bounded and non-normal nature of EOSQ-24 scores. However, the near-universal presence of marked ceiling effects across studies and domains suggests that these findings are not solely attributable to distributional assumptions. Ceiling effects are likely influenced by patient age, etiology, disease severity, treatment stage, and respondent type. Although the included studies encompassed a wide spectrum of EOS populations, subgroup-specific analyses were not feasible due to inconsistent reporting of stratified EOSQ-24 summary statistics. Future studies using individual-level data are needed to distinguish ceiling effects driven by case-mix characteristics from those reflecting inherent limitations of item content. Nonetheless, the consistent presence of ceiling effects across studies suggests the presence of ceiling effects is an inherent characteristic of the EOSQ-24 rather than a study-specific phenomenon.

## Conclusions

Ceiling effects are common in the EOSQ-24, particularly in functional and pulmonary items, and may limit its ability to capture treatment-related change. The presence of ceiling effects should be reported and recognized as a limitation in clinical studies. In practice, reliance on EOSQ-24 alone may underestimate treatment benefits in higher-functioning patients. Complementary use of objective measures, particularly pulmonary function testing, is recommended in future studies. When PROMs are used as primary outcomes, awareness of domain-specific ceiling effects is essential for interpretation.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s43390-026-01293-2>.

**Author contributions** Dr. Saarinen: Data collection, statistical analysis, writing-original draft preparation, and approval of final version of manuscript). All authors agree to be accountable for the work.

**Funding** Open Access funding provided by University of Turku (including Turku University Central Hospital). Research funding from Päivikki and Sakari Sohlberg Foundation and Vappu Uuspuu Foundation.

**Data availability** Data is available for a reasonable request from the corresponding author.

## Declarations

**Conflict of interest** None declared.

**Ethical approval** Not applicable (literature review).

**Informed consent** Not applicable (literature review).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Matsumoto H, Williams B, Park HY, Yoshimachi JY, Roye BD, Roye JDP, Akbarnia BA, Emans J, Skaggs D, Smith JT, Li Y, Burke MC, Gagnier J, Caird MS, Farley FA, Hanbali Y, Perry T, Hanif A, Matsomotu H, Musmar H, Bader K, Ahmad AA, Mladenov K, Braunschweig L, Behrend J, Lorenz HM, von Deimling U, Hell AK, Matsumoto H, Williams B, Park HY, Yoshimachi JY, Roye BD, Roye JDP, Akbarnia BA, Emans J, Skaggs D, Smith JT (2018) The final 24-item early onset scoliosis questionnaires (EOSQ-24): validity, reliability and responsiveness. *J Pediatr Orthop* 38:144. <https://doi.org/10.3171/2019.1.PEDS18704>
- Williams BA, Matsumoto H, McCalla DJ, Akbarnia BA, Blakemore LC, Betz RR, Flynn JM, Johnston CE, McCarthy RE, Roye JDP (2014) Development and initial validation of the Classification of Early-Onset Scoliosis (C-EOS). *JBJS* 96:1359–1367
- Ramo BA, McClung A, Jo CH, Sanders JO, Yazsay B, Oetgen ME (2021) Effect of etiology, radiographic severity, and comorbidities on baseline parent-reported health measures for children with early-onset scoliosis. *J Bone Jt Surg* 103:803–811. <https://doi.org/10.2106/JBJS.20.00819>
- Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60:34–42
- Haapala H, Salonen A, Suominen E, Syvänen J, Repo J, Matsumoto H, Ahonen M, Helenius I, Saarinen A (2024) Reliability and validity of the adapted Finnish version of the early onset scoliosis questionnaire (EOSQ-24). *Spine Deform* 12:1137–1143. <https://doi.org/10.1007/s43390-024-00861-8>
- Demirkiran HG, Kinikli GI, Olgun ZD, Kamaci S, Yavuz Y, Vitale MG, Yazici M (2015) Reliability and validity of the adapted Turkish version of the early-onset scoliosis-24-item questionnaire (EOSQ-24). *J Pediatr Orthop* 35:804–809
- Hanbali Y, Perry T, Hanif A, Matsomotu H, Musmar H, Bader K, Ahmad AA (2019) Reliability and validity of the Arabic version of the Early Onset Scoliosis 24 Items Questionnaire (EOSQ-24). *SICOT J*. <https://doi.org/10.1051/sicotj/2019001>
- Li Z, Yue Y, Matsumoto H, Vitale MG, Roye JDP, Song X, Xu N (2020) Reliability and validity of the simplified Chinese version of the Early Onset Scoliosis-24-Item Questionnaire (EOSQ-24). *Transl Pediatr* 9:513
- Green-Petersen I, Cheng T, Blixt S, Gerdhem P (2025) Translation and validation of the Swedish version of the early-onset scoliosis 24-item questionnaire. *Spine Deform* 13:1253–1261. <https://doi.org/10.1007/s43390-025-01064-5>
- Hwang CJ, Chang SY, Park K-B, Lee D-H, Cho JH, Matsumoto H, Park S (2025) Validation of the Korean version of the 24-item early-onset scoliosis questionnaire: a multicenter study from the Korean Research Society of Spinal Deformity. *Clin Orthop Surg* 17:123–129. <https://doi.org/10.4055/cios24210>
- Wan X, Wang W, Liu J, Tong T (2014) Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 14:135
- Saarinen A, Pakarinen O, Vaajala M, Liukkonen R, Ponkilainen V, Kuitunen I, Uimonen M (2024) Randomized controlled trials reporting patient-reported outcomes with no significant differences between study groups are potentially susceptible to unjustified conclusions—a systematic review. *J Clin Epidemiol*. <https://doi.org/10.1016/j.jclinepi.2024.111308>
- Lemans JVC, Wijdicks SPJ, Castelein RM, Kruyt MC (2021) Spring distraction system for dynamic growth guidance of early onset scoliosis: two-year prospective follow-up of 24 patients. *Spine J* 21:671–681. <https://doi.org/10.1016/j.spinee.2020.11.007>
- Matsumoto H, Auran E, Fields MW, Hung CW, Hilaire TS, Roye B, Sturm P, Garg S, Sanders J, Oetgen M, Levine S, Roye D, Vitale M (2020) Serial casting for early onset scoliosis and its effects on health-related quality of life during and after discontinuation of treatment. *Spine Deform* 8:1361–1367. <https://doi.org/10.1007/s43390-020-00175-5>
- Sleem A (2022) Health-related quality of life assessment in idiopathic early-onset scoliosis treated by serial Mehta casting. *Int J Spine Surg* 16:554–558. <https://doi.org/10.14444/8253>
- Shaw KA, Ramo B, McClung A, Thornberg D, Yazsay B, Sturm P, Jo C-H, Oetgen ME (2023) Impact of surgical treatment on parent-reported health related quality of life measures in early-onset scoliosis: stable but no improvement at 2 years. *Spine Deform* 11:213–223. <https://doi.org/10.1007/s43390-022-00572-y>
- Lemans JVC, Top A, Tabeling CS, Scholten EP, Stempels HW, Schlösser TPC, Castelein RM, Kruyt MC (2024) Health-related quality of life in early onset scoliosis patients treated with the spring distraction system: what to expect in the first 2 years after surgery. *Spine Deform* 12:489–499. <https://doi.org/10.1007/s43390-023-00777-9>
- Saarinen A, Andras L, Boachie-Adjei O, Cahill P, Guillaume T, Snyder B, Sponseller P, Sturm P, Vitale M, Helenius I (2025) High preoperative body mass index is associated with implant breakage in patients treated with magnetically controlled growing rods for early-onset scoliosis. *J Pediatr Orthop*. <https://doi.org/10.1097/BPO.0000000000002988>
- Matsumoto H, Fano AN, Ball J, Roye BD, George A, Garg S, Erickson M, Samdani A, Skaggs D, Roye DP, Vitale MG (2022) Uncorrected pelvic obliquity is associated with worse health-related quality of life (HRQoL) in children and their caregivers at the end of surgical treatment for early onset scoliosis (EOS). *J Pediatr Orthop* 42:e390–e396. <https://doi.org/10.1097/BPO.0000000000002096>
- Matsumoto H, Skaggs DL, Akbarnia BA, Pawelek JB, Hilaire TS, Levine S, Sturm P, Perez-Grueso FJS, Luhmann SJ, Sponseller PD, Smith JT, White KK, Vitale MG (2021) Comparing health-related quality of life and burden of care between early-onset scoliosis patients treated with magnetically controlled growing rods and traditional growing rods: a multicenter study. *Spine Deform* 9:239–245. <https://doi.org/10.1007/s43390-020-00173-7>

21. Roye BD, Fano AN, Matsumoto H, Fields MW, Emans JB, Sponseller P, Smith JT, Thompson GH, White KK, Vitale MG (2022) The impact of unplanned return to the operating room on health-related quality of life at the end of growth-friendly surgical treatment for early-onset scoliosis. *J Pediatr Orthop* 42:17–22. <https://doi.org/10.1097/BPO.0000000000002006>
22. Sborov KD, Agarwal M, Heffernan MJ, Anari JB, Roye B, Parent S, Miyanji F, Poon SC (2025) Complications in vertebral body tethering: what are the short term effects on patient reported outcomes? *Spine Deform*. <https://doi.org/10.1007/s43390-025-01107-x>
23. Gomez JA, Ge DH, Boden E, Hanstein R, Alvandi LM, Lo Y, Hwang S, Samdani AF, Sponseller PD, Garg S, Skaggs DL, Vitale MG, Emans J (2022) Posterior-only resection of single Hemivertebrae with 2-level versus >2-level fusion: can we improve outcomes? *J Pediatr Orthop* 42:354–360. <https://doi.org/10.1097/BPO.0000000000002165>
24. Saarinen AJ, Flynn JM, Thompson GH, Emans JB, Sturm PF, Sponseller PD, Helenius IJ (2025) Traditional versus magnetically controlled growing rods for idiopathic early-onset scoliosis: outcomes at 5-year follow-up. *Spine J*. <https://doi.org/10.1016/j.spinee.2025.07.011>
25. Saarinen AJ, Sponseller P, Thompson GH, White KK, Emans J, Cahill PJ, Hwang S, Helenius I (2024) Outcomes after completing growth-friendly surgical treatment for early-onset scoliosis in patients with skeletal dysplasia. *Bone Jt J* 106:596–602. <https://doi.org/10.1302/0301-620X.106B6.BJJ-2023-1417.R2>
26. Manning AE, Redding G, Bauer J, Yaszay B, Yang S (2025) Pulmonary function in early-onset scoliosis: a literature review. *JBJS Rev*. <https://doi.org/10.2106/JBJS.RVW.25.00122>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.