Which Criterion Is More Reliable for Selecting the Distal Fusion Level in Cases of Adolescent Idiopathic Scoliosis with Structural Thoracolumbar/Lumbar Curves: Static or Dynamic?

Jae Hwan Cho, M.D., Ph.D., Chang Ju Hwang, M.D., Ph.D., Jae Woo Park, M.D., Dong-Ho Lee, M.D., Ph.D., Choon Sung Lee, M.D., Ph.D.

Originally published online December 31, 2019;
https://doi.org/10.4184/jkss.2019.26.4.132

Korean Society of Spine Surgery
SMG-SNU Boramae Medical Center, 20, Boramae-ro 5-gil, Dongjak-gu, Seoul 07061, Korea
Tel: +82-2-831-3413 Fax: +82-2-831-3414
©Copyright 2017 Korean Society of Spine Surgery
pISSN 2093-4378 eISSN 2093-4386

The online version of this article, along with updated information and services, is located on the World Wide Web at:

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

www.krspine.org
Which Criterion Is More Reliable for Selecting the Distal Fusion Level in Cases of Adolescent Idiopathic Scoliosis with Structural Thoracolumbar/Lumbar Curves: Static or Dynamic?

Jae Hwan Cho, M.D., Ph.D., Chang Ju Hwang, M.D., Ph.D., Jae Woo Park, M.D., Dong-Ho Lee, M.D., Ph.D., Choon Sung Lee, M.D., Ph.D.
Department of Orthopedic Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Study Design: Retrospective comparative study.
Objectives: To compare the reliability of 2 criteria to predict the radiological outcomes of corrective surgery in cases of adolescent idiopathic scoliosis (AIS) with structural thoracolumbar/lumbar (TL/L) curves.
Summary of Literature Review: Distal fusion level selection in AIS with structural TL/L curves is debatable.
Materials and Methods: This study included 131 AIS patients with structural TL/L curves who underwent corrective surgery in which distal fusion was stopped at L3. Whole-spine standing radiographs and bending radiographs were obtained preoperatively. The patients were divided into 2 groups according to their findings on bending radiographs (dynamic criterion) and by the last touching vertebra and the lower end vertebra (static criterion). Radiological outcomes were assessed by reviewing postoperative radiographs. Reliability tests were conducted to compare the predictability of radiological outcomes using these 2 methods. In addition, radiological parameters were compared between both criteria.
Results: Among 131 patients, 25 showed radiologically poor outcomes (19.1%). The sensitivity of the dynamic and static criteria was 0.69 and 0.50, respectively. The specificity of each criterion was 0.49 and 0.64, respectively. Overall, the dynamic criterion showed superior reliability (p=0.03). However, no significant difference in radiological parameters could be found in a comparison of both criteria.
Conclusions: Although the dynamic criterion was more sensitive for predicting poor radiological outcomes when stopping fusion at L3 in patients with structural TL/L curves, its specificity was lower than that of the static criterion. Thus, both dynamic and static criteria should be considered when selecting the distal fusion level in cases of AIS with structural TL/L curves.

Key words: Adolescent idiopathic scoliosis, Double major curve, Distal fusion level, Lower instrumented vertebra

Introduction

There are several controversial issues in the treatment of adolescent idiopathic scoliosis (AIS) although many problems have been solved.1 Distal fusion level selection in patients with thoracolumbar/lumbar (TL/L) curves (Lenke type 3C, 4C, 5C, and 6C) is one of debatable issues in AIS surgeries. Traditionally, L4 is considered as lowest instrumented vertebra (LIV) in those cases.2) However, stopping distal fusion at L3 has been successfully tried with the advent of powerful pedicle screw system, which can save one more mobile segment (L3–4). Although the clinical meaning of leaving unfused at L3–4 level...
has not been validated, potential benefit could be anticipated in that patients were generally so young.

Several studies have researched the distal fusion level selection in AIS patients with TL/L curves.\(^3\)\(^-\)\(^5\) However, no concrete criteria to select distal fusion level has been suggested. Distal junctional wedging can be a problem when stopping fusion at L3 although its clinical effect is not drawn.\(^6\) Less correction is another potential problem when distal fusion stops at L3. So, the criteria to guide distal fusion level is important.

To help the surgeons to choose appropriate LIV, several criteria have been proposed\(^3\)\(^,\)\(^4\)\(^,\)\(^7\)\(^,\)\(^8\) Those criteria can be generally divided into two, which is static or dynamic according to the usage of preoperative dynamic radiographs. Recently, 2 criteria were proposed to reveal the radiological outcomes following corrective surgery in AIS patients with structural TL/L curves. The one was a static criterion using lower end vertebra (LEV) and last touching vertebra (LTV).\(^4\) The other was a dynamic criterion using the findings on preoperative bending radiographs.\(^9\) However, superiority of those methods has not been proved.

Thus, the purpose of this study was (1) to compare the predictability of radiological outcomes by 2 methods, and (2) to compare radiological parameters between both criteria.

### Materials and Methods

#### 1. Study design and operative method

One hundred thirty-one consecutive patients who underwent corrective surgery for AIS with structural TL/L curves (Lenke 3C, 4C, 5C, and 6C) between 2009 and 2013 at our institution were included. All surgeries were conducted by single surgeon, and all patients were followed up for more than 2 years. Demographic data were obtained from electronic medical records. Radiological data were obtained from whole spine standing radiographs (anteroposterior and lateral), and side-bending radiographs preoperatively. Plain radiographs were obtained with the patients in an upright standing position in both views and with arms folded forward in lateral views. Side-bending radiographs were obtained passively by positioning a hard pillow under the apex vertebra. Patients were regularly followed up at 1, 6, and 12 months postoperatively and yearly thereafter with whole spine radiographs.

All surgeries were conducted by posterior only approach with the rod derotation method with pedicle screw fixation in every level. Direct vertebral rotation or additional anterior surgery was not performed. The selection of upper instrumented vertebra (UIV) and lower instrumented vertebra (LIV) depended on the method by Lenke et al.\(^10\) However, we tried to stop L3 distally whenever it was judged to be available. Only cases which distal fusion level was L3 were included in this study. This study was approved by the Institutional Review Board of our institution (IRB No. 2018–0668), which waived the requirement for informed consent due to the retrospective nature of the data analysis.

#### 2. Radiological Assessment

Radiological parameters included the degree of curves, trunk shift, LIV+1 tilt, LIV disc angle, apical vertebral translation (AVT), LEV, LTV, pelvic obliquity, lumbar lordosis, and thoracic kyphosis. The degree of curves was measured by Cobb angle method. The flexibility of the curve was assessed by the equation: \((\text{Cobb angle in standing position} - \text{Cobb angle in bending view}) / \text{Cobb angle in standing position} \times 100(\%)\). The correction rate was defined as follows: \((\text{Preoperative Cobb angle} - \text{postoperative Cobb angle}) / \text{preoperative Cobb angle} \times 100(\%)\). Trunk shift was measured by the distance between C7 plumb line and CSVL. LIV+1 tilt was assessed by the angle between upper endplate of vertebra just below LIV and horizontal line. LIV disc angle was defined by the angle between lower endplate of LIV and upper endplate of vertebra just below LIV. AVT was defined by the distance between center of apical vertebra and CSVL. Pelvic obliquity was measured by the angle between the line passing superior margin of both iliac crest and horizontal line. Lumbar lordosis was measured by Cobb’s angle between both upper endplates of L1 and S1. Thoracic kyphosis was measured by Cobb’s angle between upper endplate of T1 and lower endplate of T12. All radiological measurements are summarized in Fig. 1.

Radiological outcomes were divided by good or poor outcomes checked in 2 years postoperatively. Radiologically poor results were defined if more than one criterion was met with following: (1) the distance between C7 plumb line and CSVL > 20 mm, (2) LIV+1 tilt > 20°, (3) LIV disc angle > 7°,
3. Criteria to select distal fusion level

We retrospectively applied 2 criteria to select distal fusion level. The one is a static criterion, which was suggested by our previous study. If lower end vertebra (LEV) ≥ L3 and last touching vertebra (LTV) ≥ L4, then L3 is selected as appropriate LIV. If LEV ≤ L4 or LTV = L5, then L4 is selected as appropriate LIV. The other is a dynamic criterion, which was also recently published. If L3 crosses CSVL with a rotation less than grade II in side-bending radiographs, then L3 is selected as appropriate LIV. If the curve is rigid, then LIV should be L4.

4. Statistical Analysis

Radiological parameters were compared between pre- and postoperatively by paired t-test. Reliability of prediction by each criterion was assessed by sensitivity and specificity, and ROC (receiver operating characteristic) curve. Cases showing different prediction by 2 criteria (i.e., good by dynamic, and poor by static criteria) were entered into subgroup analysis.

Demographic and radiological parameters between subgroups were compared by chi-square test or independent t-test. A multivariate regression analysis was performed following a univariate analysis.

Statistical analyses were performed using the Statistical Package for Social Sciences software package (version 21.0, SPSS Inc., Chicago, IL) with p-values less than .05 considered statistically significant.

Results

1. Demographic Data and Radiological Parameters

One hundred thirty-one patients (20 male and 111 female) were included in this study, with a mean age of 15.3 ± 3.4 years. The mean height and weight of the patients were 160.0 ± 7.6 cm and 51.2 ± 9.8 kg, respectively. Body mass index was 19.9 ± 2.9. The mean follow-up period was 35.9 ± 18.4 months, and the average Risser grade was 3.5 ± 1.5. Average fusion level of operation was 9.8 ± 2.1. Pre- and postoperative radiological parameters were described in Table 1. The degree of the curve, LIV+1 tilt, AVT, and TK were improved by corrective surgery although L3–L4 disc wedging did not change.

2. Radiological outcomes

Postoperative poor radiological outcomes were shown in 25 patients (25/131, 19.1%). Fifty-one cases (38.9%) were predicted to be poor radiological outcomes by static criteria.

Table 1. Comparisons of pre- and postoperative radiological parameters

<table>
<thead>
<tr>
<th></th>
<th>Preop</th>
<th>Postop</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TL/L curve (°)</td>
<td>56.8±11.1</td>
<td>15.4±7.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MT curve (°)</td>
<td>44.0±16.2</td>
<td>11.0±5.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C7-CIVS distance (mm)</td>
<td>-10.3±13.4</td>
<td>-5.0±9.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>L3-L4 disc wedging (°)</td>
<td>3.3±2.8</td>
<td>3.7±2.7</td>
<td>0.276</td>
</tr>
<tr>
<td>LIV tilt (°)</td>
<td>18.8±6.0</td>
<td>8.5±4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>AVT (mm)</td>
<td>42.2±11.8</td>
<td>19.3±8.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TK (°)</td>
<td>26.4±9.7</td>
<td>29.2±9.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LL (°)</td>
<td>46.5±11.2</td>
<td>46.3±9.5</td>
<td>0.773</td>
</tr>
</tbody>
</table>

*Mean and standard variation in continuous variables and number of cases in categorical variables

and 72 cases (55.0%) by dynamic criteria. The relationship between 2 criteria and poor radiological outcome was shown in Table 2. The sensitivity to predict poor radiological outcomes following stopping distal fusion at L3 by dynamic and static criteria was 0.69 and 0.50, respectively (95% CI [-0.09, 0.47]).

**Table 2.** Cross-table to show the reliability of each criterion to predict radiologically poor outcome

<table>
<thead>
<tr>
<th></th>
<th>Poor outcome</th>
<th>Good outcome</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static (+)</td>
<td>13</td>
<td>38</td>
<td>51</td>
</tr>
<tr>
<td>Static (-)</td>
<td>13</td>
<td>67</td>
<td>80</td>
</tr>
<tr>
<td>Subtotal</td>
<td>26</td>
<td>105</td>
<td>131</td>
</tr>
</tbody>
</table>

*Sensitivity: 0.50, *specificity: 0.64, *accuracy: 0.61, *positive predictive value: 0.26, *negative predictive value: 0.84.

<table>
<thead>
<tr>
<th></th>
<th>Poor outcome</th>
<th>Good outcome</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic (+)</td>
<td>18</td>
<td>54</td>
<td>72</td>
</tr>
<tr>
<td>Dynamic (-)</td>
<td>8</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td>Subtotal</td>
<td>26</td>
<td>105</td>
<td>131</td>
</tr>
</tbody>
</table>

*Sensitivity: 0.69, *specificity: 0.49, *accuracy: 0.53, *positive predictive value: 0.25, *negative predictive value: 0.86.

![Fig. 2. A receiver operating characteristic curve comparing the reliability of the dynamic and static criteria.](image)

**Table 3.** Comparisons of preoperative demographic and radiological parameters

<table>
<thead>
<tr>
<th></th>
<th>Dynamic good- Static poor (N=21)</th>
<th>Dynamic poor- Static good (N=42)</th>
<th>p-value (univariate)</th>
<th>p-value (multivariate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yrs)</td>
<td>14.6±2.5</td>
<td>15.5±4.2</td>
<td>0.408</td>
<td>NA</td>
</tr>
<tr>
<td>Sex (M:F)</td>
<td>2:19</td>
<td>7:35</td>
<td>0.705</td>
<td>NA</td>
</tr>
<tr>
<td>Height (mm)</td>
<td>158.5±7.1</td>
<td>160.0±7.6</td>
<td>0.46</td>
<td>NA</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>49.3±5.5</td>
<td>51.1±9.2</td>
<td>0.34</td>
<td>NA</td>
</tr>
<tr>
<td>BMI</td>
<td>19.6±1.6</td>
<td>19.9±2.7</td>
<td>0.619</td>
<td>NA</td>
</tr>
<tr>
<td>Risser grade</td>
<td>3.1±1.9</td>
<td>3.3±1.5</td>
<td>0.787</td>
<td>NA</td>
</tr>
<tr>
<td>TL/L curve (°)</td>
<td>57.3±8.1</td>
<td>58.4±10.4</td>
<td>0.680</td>
<td>NA</td>
</tr>
<tr>
<td>TL/L flexibility (%)</td>
<td>54.6±16.0</td>
<td>46.7±17.5</td>
<td>0.090</td>
<td>0.047</td>
</tr>
<tr>
<td>MT curve (°)</td>
<td>50.9±12.7</td>
<td>41.5±14.4</td>
<td>0.014</td>
<td>0.229</td>
</tr>
<tr>
<td>C7-CSVL distance (mm)</td>
<td>-10.1±11.5</td>
<td>-10.7±15.5</td>
<td>0.858</td>
<td>NA</td>
</tr>
<tr>
<td>L3-L4 disc wedging (°)</td>
<td>2.0±1.6</td>
<td>3.5±2.9</td>
<td>0.011</td>
<td>0.289</td>
</tr>
<tr>
<td>LIV tilt (°)</td>
<td>16.8±4.1</td>
<td>21.5±5.2</td>
<td>0.001</td>
<td>0.613</td>
</tr>
<tr>
<td>AVT (mm)</td>
<td>35.0±8.2</td>
<td>48.6±10.3</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pelvic obliquity (°)</td>
<td>2.0±1.4</td>
<td>2.4±2.0</td>
<td>0.287</td>
<td>NA</td>
</tr>
<tr>
<td>TK (°)</td>
<td>25.2±9.3</td>
<td>26.5±10.0</td>
<td>0.644</td>
<td>NA</td>
</tr>
<tr>
<td>LL (°)</td>
<td>42.5±8.2</td>
<td>47.7±12.3</td>
<td>0.087</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*Data represent mean and standard deviation.
*Negative means trunk is shifted to the left side.
3. Subgroup analysis

Twenty-one cases were predicted to be good outcomes by dynamic, and poor outcomes by static criteria. Whereas, 42 cases were predicted to be poor outcomes by dynamic, and good outcomes by static criteria. Preoperative demographic and radiological data were compared between the 2 groups in Table 3. The TL/L flexibility (54.6° vs. 46.7°, p=0.047), AVT (35.0 mm vs. 48.6 mm, p=0.001) and LL (42.5° vs 47.7°, p=0.019) were the radiological parameters exhibiting significant differences between 2 groups based on multivariate analyses. However, postoperative radiological parameters revealed no differences between 2 groups (Table 4).

Two representative cases were shown in Fig. 3 and Fig. 4.

Discussion

The selection of the distal fusion level in AIS with a structural...
TL/L curve (Lenke type 3C, 4C, 5C, and 6C) has not been fully elucidated. It was recommended to extend to L4 before using pedicle screws. However, stopping at L3 has been tried along with the development of strong fixation by pedicle screws and rods. There are several reasons not to extend distal fusion into L4. First, saving motion segment can be a meaningful because patients usually are young. Second, extension to L4 could increase the risk of adjacent disc degeneration at L4-L5 and L5-S1, which were most commonly prone to be degenerated. Third, not a few proportions of patients with AIS showed fused L5-S1, which was previously published. This means only 1 segment is left when they were fused to L4.

However, stopping L3 distally could cause worries about incomplete correction, progressive worsening of distal curves, or junctional problems (aggravation of L3-L4 wedge angle). Although stopping fusion at L3 might cause less correction in terms of Cobb’s angle and AVT, good clinical outcomes and no radiological worsening could be expected in our previous study. In fact, many authors suggested there are no differences of clinical outcomes according to distal fusion level which is L3 or L4. However, controversial results have been issued in terms of radiological outcomes.

Before further discussion, it needs to make a definition for poor radiological outcomes. Traditionally, C7-CSVL distance more than 2cm has been considered to be poor outcomes, which is probably not disputed by most surgeons. Furthermore, progression of distal lumbar curve is somewhat reasonable because that means a compensation. However, other criteria such as LIV+1 tilt or LIV disc angle (L3-L4 disc wedge angle) could be controversial even we also adopted those criteria in this study. Is there any difference between patients with 5° of L3-L4 disc wedging and 10° of it? In recent study, LIV+1 tilt angle improved postoperatively while L3-L4 disc wedging did not change if fusion stops at L3. However, it was also reported that increased L3-L4 disc wedging is correlated with aggravation of lumbar compensatory curves. Although those articles did not mention about clinical outcomes, and showed only short-term outcomes, it is definite that postoperatively observed remaining L3-L4 disc wedging can be a major issue when L3 is selected as a LIV. Further long-term, well-designed study could reveal the importance of this phenomenon in near future.

To select more appropriate LIV, many criteria have been suggested. We selected 2 criteria to compare the reliability to select LIV to avoid poor radiological outcomes. The one is a static criterion based on LEV and LTV measured on standing radiographs. The other is a dynamic criterion based on the findings by using side-bending radiographs. Our results showed better sensitivity by the dynamic criterion but better specificity by the static one. That means that if we use the dynamic one, the probability to extend the fusion into L4 could increase. Conversely, the static criterion could give more chance to stop at L3 regardless of more possible radiological problems. Although overall reliability to predict poor radiological outcomes was superior by the dynamic criterion, both methods are thought to be important to consider because the predictability is thought to be not so high by only one method.

Many studies focused the condition to save L3-4 segments including 2 criteria we used in this study. It was recommended to extend fusion into L4 in rigid curves. Side-bending radiographs or traction x-ray under anesthesia were used to determine the rigidity of the curve. There were several reports to support the advantage of stopping at L3. It was reported that severe disc degeneration was observed in patients with the lowest fusion level at L4 or lower. However, no considerable differences of disc and facet joint degeneration between L3 and L4 groups were suggested by another study. This means that the advantage of stopping at L3 is unclear clinically and radiologically regardless of its theoretical benefits. No definitive conclusion has been made about the LIV selection in AIS patients with structural TL/L curves.

Interestingly, several radiological parameters showed some important findings to consider. The more TL/L curve is flexible, the higher chance to stop distal fusion at L3 can be expected by the dynamic one, which was supported by recently published study. Less AVT is also related with the selection of L3 by the dynamic one. If AVT is greater, then L3 vertebral body usually could not cross the CSVL by side bending. That might be the explanation of the finding. Less LL is related with more compensation at the lumbar curve, which means more flexibility. So, this is thought be also reasonable explanation for the difference of LL between 2 subgroups.

There are some limitations in this study, including those inherent to the retrospective design. First, there is intrinsic selection bias because we did not select LIV using strict criteria. We tried to select L3 as LIV whenever it could be available. So,
we could not include the patients who underwent fusion to L4 in this study. More reliable results could have been derived if we have enough cases that LIV was L4. Second, clinical outcomes such as SRS–22 scores were not included in the analysis. Patients showing radiological poor result such as 10 degrees of L3–L4 disc wedging frequently did not complain at all. Third, different result could be derived if we perform direct vertebral rotation. However, because the number of patients was enough and the LIV was L3 for all included cases, it could be enough to reveal the general predictability of 2 different criteria for structural TL/L curves.

Conclusions

In summary, dynamic criteria showed higher sensitivity and static criteria showed higher specificity to predict radiological poor outcomes in distal fusion level selection for patients with AIS with TL/L structural curves. Overall reliability by dynamic criteria was superior to that by static criteria. The main differences between 2 criteria could be derived by preoperative TL/L curve flexibility, AVT, and LL. It is required to compare both type of criteria before selecting distal fusion level because they have merits and limitations.

Acknowledgments

This study did not receive any funding.

REFERENCES


구조적 흉요부/요부 만곡 청소년기 특발성 측만증에서 원위부 유합 범위를 결정하는데 있어서 동적 혹은 정적 기준의 신뢰도 비교
조재환 ∙ 황창주 ∙ 박재우 ∙ 이동호 ∙ 이춘성
울산대학교 의과대학 정형외과학교실

연구 계획: 후향적 비교 연구
목적: 구조적 흉요부/요부 만곡 청소년기 특발성 측만증 환자 수술 후 방사선학적 결과를 예측하는 2가지 기준의 신뢰도를 비교하기 위함

대상 및 방법: 원위 유합을 L3까지 수술한 구조적 흉요부/요부 만곡 청소년기 특발성 측만증 환자 131명을 대상으로 하였다. 수술전 전신 기립 자세와 측부 굽힘 방사선 사진을 촬영하였다. 환자들은 측부 굽힘 사진 상의 기준(동적)과, Last touching vertebra와 lower end vertebra에 따른 기준(정적)에 따라 분류하였고, 숭후 방사선학적 결과에 따라 분석하였다. 두 기준에 따른 방사선학적 결과 예측의 신뢰도를 분석하였고, 여러 방사선학적 인자들의 차이를 분석하였다.

결과: 131명의 환자중 25명이 불량한 결과를 보였다(19.1%), 동적 기준과 정적 기준의 민감도는 각각 0.69와 0.50이었고, 특이도는 0.49와 0.64였다. 전반적으로는 동적 기준이 우월한 신뢰도를 보여주었다(p=0.03). 하지만, 두 기준에 따른 방사선학적 인자들의 차이는 없었다. 결론: L3에서 유합을 멈췄을 때 동적 기준이 더 불량한 예후를 예측하는데 민감함을 보이지만, 특이도는 정적 기준보다 낮았다. 구조적 흉요부/요부 만곡 청소년기 특발성 측만증 환자에서 원위 유합 범위 결정에 두 기준 모두를 고려하는 것이 좋을 것으로 생각된다.

색인 단어: 청소년기 특발성 측만증, 이중 대만곡, 원위 유합, 하위 고정 척추

약칭 제목: 구조적 흉요부/요부 만곡 청소년기 특발성 측만성 환자의 원위 유합 범위 결정