Nonoperative Treatment of Spondylolysis and Grade I Spondylolisthesis in Children and Young Adults

A Meta-analysis of Observational Studies

Guy Klein, BS,* Charles T. Mehlman, DO, MPH,† and Mary McCarty, BS‡

Background: The incidence of spondylolysis is at least 6% by the end of childhood, and painful lesions are not infrequent. The most common treatments for spondylolysis are nonoperative in nature and include bracing, activity restriction, and therapeutic exercises. These treatments have been used either alone or in concert. The aim of this meta-analysis was to identify and summarize the evidence from the literature on the effectiveness of nonoperative treatment for spondylolysis (including those with up to 25% spondylolisthesis) in children and young adults.

Methods: A comprehensive literature search identified articles meeting the following inclusion criteria: (1) the target population was children and young adults with spondylolysis (including those with up to 25% spondylolisthesis); (2) the treatment intervention was nonoperative; (3) minimum follow-up was 1 year in studies using clinical parameters as the primary outcome; and (4) the studies included at least 10 subjects. Outcome data from eligible studies were pooled into 1 of 2 groups: clinical outcome or radiographic evidence of a union of the pars defects.

Results: Fifteen observational studies measuring the clinical outcome had a weighted and pooled success rate of 83.9% in 665 patients. A subgroup analysis comparing the clinical outcome of patients treated with a brace to patients treated without a brace was not significantly different (P = 0.75). Ten studies evaluating radiographic healing of the defects had a pooled success rate of 28.0% (n = 847). A subgroup analysis showed that unilateral defects healed at a pooled and weighted rate of 71% (n = 92), significantly more than bilateral defects at 18.1% (n = 446, P < 0.0001). An additional subgroup analysis showed acute defects healed at a rate of 68.1% (n = 236), significantly more than progressive lesions (28.3%, n = 224, P < 0.0001) and terminal lesions (n = 217, P < 0.0001), of which not one defect healed.

Conclusions: A meta-analysis of observational studies suggests that 83.9% of patients treated nonoperatively will have a successful clinical outcome after at least 1 year. Bracing does not seem to influence this outcome. In contrast to the high rate of success with clinical parameters, most defects did not heal with nonoperative treatment suggesting that a successful clinical outcome does not depend on healing of the lesion. Lesions diagnosed at the acute stage were more likely to heal after nonoperative treatment as were unilateral defects when compared with bilateral defects.

Level of Evidence: Meta-analysis of level IV studies. Therapeutic level IV.

Key Words: spine, spondylolysis, spondylolisthesis, children

(J Pediatr Orthop 2009;29:146–156)

Spondylolysis is remarkably common in children and adults with an incidence of at least 6% by the end of childhood. Lesions that become painful are not infrequent, and this is illustrated by the fact that spondylolysis is the most common identifiable cause of low back pain in adolescent athletes. Although genetics may play a role in predisposing one to spondylolysis, the current literature suggests that it is not a congenital condition. The pathogenesis of spondylolysis involves disrupted ossification of the pars interarticularis due to repetitive mechanical stress. Prevalent and often painful spondylolysis results in a significant use of medical resources related to its diagnosis and treatment.

Nonoperative treatments for spondylolysis are most common and include bracing, activity restriction, and therapeutic exercises, either used alone or concomitantly. The literature evaluating nonoperative treatment of spondylolysis is overwhelmingly observational; thus, many aspects of its treatment remain controversial. In addition, although a large number of reviews on this topic have been written, all have been narrative in nature. The purpose of our study was to perform a meta-analysis of studies focusing on nonoperative treatment of spondylolysis (including those with low-grade spondylolisthesis) in children and young adults.

METHODS

Eligibility Criteria

Articles were identified which met the following eligibility criteria: (1) a target population of children and young adults with spondylolysis (including those with up to 25% spondylolisthesis), (2) nonoperative treatment intervention, (3) minimum of 1-year follow-up in studies using clinical parameters as the primary outcome, and (4) the studies included at least 10 subjects.
TABLE 1. Criteria for Clinical Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Excellent</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No pain</td>
<td>Occasional aching with vigorous activity</td>
<td>Pain with vigorous activity</td>
<td>Pain during activities of daily living even with brace</td>
</tr>
<tr>
<td></td>
<td>No brace requirement</td>
<td>No brace requirement</td>
<td>Occasional use of brace</td>
<td>Candidate for fusion</td>
</tr>
<tr>
<td></td>
<td>Full activities, including sports</td>
<td>Full activities, including sports</td>
<td>Activities of daily living without pain</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Steiner and Micheli. 21

Study Identification

The PubMed MEDLINE database was searched (1950 to November 2007) for the following terms and Boolean operators: (spondylolysis OR spondylolisthesis OR pars interarticularis) and (treatment OR conservative OR nonoperative OR brace). Additional articles were identified using the “related articles” search feature on prominent articles that met our eligibility criteria. The Cochrane Central Register of Controlled Trials and the Cochrane Database of Systematic Reviews were searched for clinical trials and systematic reviews (1800 to November 2007) by searching the terms (spondylolysis OR spondylolisthesis OR pars interarticularis). CINAHL was searched (1982 to November 2007) using the terms (spondylolysis OR spondylolisthesis OR pars interarticularis) AND (treatment). In addition, we searched the abstracts of the annual meetings of the Scoliosis Research Society (2001–2006) (www.spineuniverse.com/displayarticle.php/article2591.html) and the American Association of Neurological Surgeons/Congress of Neurological Surgeons Joint Section on SpineUniverse (www.spineuniverse.com/displayarticle.php/article1229.html). The programs for the annual meetings of the Pediatric Orthopaedic Society of North America (2002 to 2007) were hand searched. The bibliographies of relevant chapters in 2 major textbooks in Pediatric Orthopaedics were searched (Lovell and Winter’s Pediatric Orthopaedics, edited by Morrissy and Weinstein, and Tachdjian’s Pediatric Orthopaedics, edited by Herring). To identify recently completed or ongoing trials, we searched the National Research Register (www.nrr.nhs.uk), Current Controlled Trials (www.controlled-trials.com), www.clinicaltrials.gov, the Australian Clinical Trials Registry (www.actr.org.au), and the University Hospital Medical Information Network Clinical Trials Registry (www.umin.ac.jp/ctr). Full text was obtained for the articles that merited further review based on title and abstract. The bibliographies of these articles and reviews of spondylolysis or spondylolisthesis were screened for additional studies to be evaluated. Finally, we used the Science Citation Index to screen articles that had cited the 22 studies that ultimately met our inclusion criteria. Our exhaustive English language search did not limit study inclusion by the year of publication and included articles written in other languages that were found in our search.

Assessment of Study Quality and Data Abstraction

Two authors (G.K. and C.T.M.) independently assessed the full text of 77 potentially eligible studies using a standardized evidence-based review form. This form assessed the eligibility criteria, type of study, level of evidence (according to the Oxford Centre for Evidence-Based Medicine), demographic information, methodology, intervention, and reported outcomes. In addition, the same 2 authors independently abstracted data from the studies using the review form. Disagreements between authors were resolved with discussion and additional review. Attempts were made to contact 25 authors of past and current studies to request additional data. Additional data on 6 published studies were obtained in this fashion. The data extracted included sample size, number of cases with a good or excellent clinical result, primary outcome, patient demographics, treatment protocol, and percentage of lesions at L5. In studies with a radiographic outcome, the number of unions and defects was also abstracted. The definition of union used by the authors of various studies ranged from evidence of healing on plain films only to computed tomographic (CT) evidence of bony union. Studies were included provided plain radiographs, CT, or both were used for diagnostic purposes.

Data Analysis

Outcome data from eligible studies were pooled into 1 of 2 groups: functional outcome or radiographic evidence of a union of the pars defects. In the clinical outcome pool, treatment success was considered as either a good or excellent outcome, as defined by the scale of Steiner and Micheli 21 (Table 1) or a comparable measure by the study’s authors. These criteria dichotomize clinical/functional outcomes into either successful (excellent or good results) or unsuccessful (fair and poor results). A successful result can be further typified by occasional or no pain with vigorous activities, no brace requirement, and full unrestricted activities. An unsuccessful result can be further typified by more frequent pain with vigorous activities or pain with activities of daily living, a brace requirement, and patients considered candidates for spinal fusion. In the radiographic pool, defects were characterized when possible as acute (early), progressive, or chronic (chronic) according to the Tokushima classification as defined by Morita et al 22 and Fujii et al 23 (Table 2; Fig. 1). We analyzed the data using the random effects model of DerSimonian and Laird, 24 with 95% confidence intervals (CI) for each outcome. Odds ratios were calculated to compare the effects of subgroups on clinical and radiographic outcomes.

TABLE 2. Tokushima Classification for Grading Defects as Acute (Early), Progressive, and Terminal (Chronic)

<table>
<thead>
<tr>
<th></th>
<th>Acute (Early)</th>
<th>Progressive</th>
<th>Terminal (Chronic)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hair-line defect also referred to as “a fissure in the pars” or “focal bony absorption”</td>
<td>The defect is moderately wide; the edges are now round</td>
<td>The defect is wide with sclerotic changes</td>
</tr>
</tbody>
</table>

Adapted from Morita et al and Fujii et al. 23

TABLE 2. Tokushima Classification for Grading Defects as Acute (Early), Progressive, and Terminal (Chronic)

<table>
<thead>
<tr>
<th></th>
<th>Acute (Early)</th>
<th>Progressive</th>
<th>Terminal (Chronic)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hair-line defect also referred to as “a fissure in the pars” or “focal bony absorption”</td>
<td>The defect is moderately wide; the edges are now round</td>
<td>The defect is wide with sclerotic changes</td>
</tr>
</tbody>
</table>

Adapted from Morita et al and Fujii et al. 23
Inverted funnel plots were constructed for each outcome to assess possible publication bias.25

**Evaluation of Heterogeneity**

Before analyzing the data, we developed a hypothesis that would account for heterogeneity between studies. We hypothesized that differences in population characteristics (including unilateral vs bilateral and acute vs terminal lesions), treatment (including bracing vs no bracing), and methodology (use of plain radiograph vs CT) would account for observer heterogeneity. We used the $I^2$ statistic (a tool aimed at describing the proportion of total variation in study estimates considered to be due to heterogeneity) to quantify the amount of inter-study variability.26,27 Consistent with the literature, we have assigned the adjectives of low, moderate, and high heterogeneity to $I^2$ values of 25%, 50%, and 75%, respectively.27

**Statistical Analysis**

Meta-analysis and statistical analysis were conducted using Comprehensive Meta-Analysis v2.0 (Biostat, Englewood, NJ). Confidence limits were set at 95%, and data were pooled in a random effects model. Study weights were calculated in Comprehensive Meta-Analysis using 1/(variance + tau^2), an equation that takes into account sample size, intrastudy variability, and interstudy variability. In addition, we used Epi Info v6 (CDC, Atlanta, Ga) for statistical analysis (Fisher exact test). Probability values (P values) less than or equal to 0.05 were considered to be statistically significant. Odds ratios with 95% CI were also calculated where appropriate.

**RESULTS**

**Literature Search**

Our search identified 3,863 articles, of which 77 were selected for a full-text evaluation. Seventy-three of the 77 studies were identified via PubMed, although many of these were also identified through the other search techniques. Four additional studies were identified by screening bibliographies of the retrieved articles. After 2 authors independently performed a full-text review, 55 of the 77 articles were excluded from our study; 24 reported on less than 10 cases, 14 had a population older than our targeted demographic, 5 did not report on treatment outcome, 4 reported on other conditions, 3 had a follow-up of less than 1 year with a clinical outcome, 3 articles reported incomplete data, and 2 articles reported on cases that had been reported in other publications. Ultimately, 22 studies met the inclusion criteria of this review.

**Study Characteristics**

Of the 22 articles selected for inclusion, 16 studies measured the functional outcome17,21,28–39 and 10 studies measured radiographic evidence of healing of the pars defects17,20,22,23,30,35,40–43 Four articles reported both functional and radiographic outcomes.17,30,35,43 Tables 3 and 4 present the eligibility criteria, outcomes, and methodology of the studies grouped by clinical and radiographic outcome, respectively. For studies that measured clinical outcome, the nature of the data was such that we could ascertain overall success rate and conduct a subgroup analysis based on treatment type (brace vs no brace). For studies with a radiographic outcome, the nature of the data was such that we could conduct subgroup analyses comparing healing between unilateral and bilateral defects and acute, progressive, and terminal lesions.

Of the 16 articles that evaluated a clinical outcome, 12 were retrospective case series, 3 were prospective case series, and 1 was a randomized controlled trial. Eight of the 10 studies that evaluated a radiographic outcome were retrospective case series and the other 2 were prospective case series. The treatment interventions in the 22 observational studies were combinations of bracing, activity restriction, and therapeutic exercises. The one randomized controlled trial compared the functional results of an exercise program training deep abdominal muscles and lumbar multifidus to a control group that performed physician-directed exercise such as swimming, walking, and gym work. As stated in our inclusion criteria, we did not include any study measuring functional outcome that had a follow-up period of less than 1 year. Five of the 10 studies in the radiographic outcome pool included cases with less than 1-year follow-up.

The sample size of studies with a clinical outcome ranged from 12 to 73 patients. The average age in the clinical pool was 14.8 years at diagnosis, 88% of patients had their
TABLE 3. Characteristics of Studies Evaluating Clinical Outcome

<table>
<thead>
<tr>
<th>Study Author</th>
<th>Design</th>
<th>Intervention</th>
<th>Length of Follow-up (Range)</th>
<th>Mean Age of Subjects (Range), y</th>
<th>No. Patients with Defects at L5</th>
<th>Positive Treatment (Excellent or Good)</th>
<th>Sample Size</th>
<th>Unweighted Treatment Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson et al</td>
<td>CS</td>
<td>TLSO, physical therapy, activity restriction</td>
<td>1 y</td>
<td>13 (5–17)</td>
<td>34</td>
<td>31</td>
<td>34</td>
<td>91%</td>
</tr>
<tr>
<td>Bell et al</td>
<td>CS</td>
<td>TLSO 23 h/d, exercises, activity restriction</td>
<td>Minimum of 1 y</td>
<td>11.4 (5.7–15.2)</td>
<td>27</td>
<td>28</td>
<td>28</td>
<td>100%</td>
</tr>
<tr>
<td>Blanda et al</td>
<td>CS</td>
<td>LSO maintaining lumbar lordosis, exercises, activity restriction, posttreatment PT</td>
<td>Minimum of 2 y</td>
<td>15.5 (11–20)</td>
<td>53</td>
<td>60</td>
<td>62</td>
<td>97%</td>
</tr>
<tr>
<td>Debnath et al</td>
<td>Prospective case series</td>
<td>LSO, organized PT, activity restriction</td>
<td>Survey of symptoms at 2 y</td>
<td>20 (14–35)</td>
<td>24</td>
<td>34</td>
<td>42</td>
<td>81%</td>
</tr>
<tr>
<td>d’Hemecourt et al</td>
<td>CS</td>
<td>TLSO 23 h/d, physical therapy, activity restriction</td>
<td>4 y (2–9)</td>
<td>15.7 (9–19)</td>
<td>48</td>
<td>56</td>
<td>73</td>
<td>77%</td>
</tr>
<tr>
<td>El Rassi et al</td>
<td>CS</td>
<td>Combinations of TLSO, PT, and activity restriction</td>
<td>Minimum of 2 y</td>
<td>13.1 (9–18)</td>
<td>49</td>
<td>51</td>
<td>57</td>
<td>89%</td>
</tr>
<tr>
<td>Frennered et al</td>
<td>CS</td>
<td>Combinations of bracing, activity restriction, and PT, and no treatment</td>
<td>7 y (1–15.7)</td>
<td>12.2 (1.6–16.9)</td>
<td>43</td>
<td>26</td>
<td>43</td>
<td>60%</td>
</tr>
<tr>
<td>Lets et al</td>
<td>Prospective case series</td>
<td>Various combination of TLSO 23 h/d, rest, plaster vest, exercises</td>
<td>1 y</td>
<td>13.4 (9–16)</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>92%</td>
</tr>
<tr>
<td>Miller et al</td>
<td>Prospective case series</td>
<td>Nonrigid lumbar bracing 23 h/d, flexion only trunk strengthening after 2–4 wk relative rest, activity restriction</td>
<td>7–11 y</td>
<td>14.6 (12–20)</td>
<td>23</td>
<td>29</td>
<td>32</td>
<td>91%</td>
</tr>
<tr>
<td>Pizzutillo and Hummer</td>
<td>CS</td>
<td>Varied forms of conservative treatment, including rest, traction, exercises, TLSO, corsets, and casts</td>
<td>5.5 y (1–14.8)</td>
<td>14.3 (6.5–21)</td>
<td>53</td>
<td>41</td>
<td>58</td>
<td>71%</td>
</tr>
<tr>
<td>Seitsalo</td>
<td>CS</td>
<td>Activity restriction and exercises in all pts; brace in only 5 pts</td>
<td>15.3 y (5–25)</td>
<td>13.8 ± 3.6</td>
<td>71</td>
<td>63</td>
<td>72</td>
<td>88%</td>
</tr>
<tr>
<td>Skinner and Micheli</td>
<td>CS</td>
<td>Boston Brace for 6 mo 23 h/d, PT</td>
<td>2.5 y (1–6.3)</td>
<td>16 (7–32)</td>
<td>61</td>
<td>52</td>
<td>67</td>
<td>78%</td>
</tr>
<tr>
<td>Sys et al</td>
<td>CS</td>
<td>Boston Brace 23 h/d, hamstring stretching, ab strengthening, pelvic tilts</td>
<td>Minimum of 1 y</td>
<td>18.1 (15–27)</td>
<td>11</td>
<td>17</td>
<td>18</td>
<td>94%</td>
</tr>
<tr>
<td>Takemitsu et al</td>
<td>CS</td>
<td>Combinations of bracing, activity restriction, and PT</td>
<td>1–10 y</td>
<td>12.3 (7–16)</td>
<td>19</td>
<td>21</td>
<td>21</td>
<td>100%</td>
</tr>
<tr>
<td>Turner and Bianco</td>
<td>CS</td>
<td>18 pts no treatment, 13 pts treated with supportive garments, activity restriction, and muscle strengthening exercises</td>
<td>1–9 y</td>
<td>19 y or younger</td>
<td>55</td>
<td>30</td>
<td>46</td>
<td>65%</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>582</td>
<td>550</td>
<td>665</td>
<td>82.7%</td>
</tr>
<tr>
<td>O’Sullivan et al</td>
<td>RCT</td>
<td>Experimental group: 10 exercises training of internal obliques, transversus abdominis, lumbar multifidi</td>
<td>2.5 y</td>
<td>31 (16–49)</td>
<td>36</td>
<td>36</td>
<td>42</td>
<td></td>
</tr>
</tbody>
</table>

CS indicates retrospective case series; LSO, lumbosacral orthosis; PT, physical therapy; pts, patients; RCT, randomized controlled trial; TLSO, thoracolumbar sacral orthosis.
TABLE 4. Characteristics of Studies Evaluating Radiographic Outcome

| Study Author          | Type of Study | Intervention                                                                 | Healing of Defects Measured by | Mean Age (Range), y | N. Patients with Defects at L5 | No. Patients Acute Healed | Acute Total | Progressive Healed | Progressive Total | Terminal Healed | Terminal Total | Bilateral Healed | Bilateral Total | N. Defects That Healed | Total Defects | Unweighted Percentage of Healed Defects | Unweighted subtotals |
|-----------------------|---------------|-----------------------------------------------------------------------------|-------------------------------|-------------------|--------------------------------|----------------------------|-------------|-------------------|-------------------|-----------------|----------------|-------------------|-----------------|----------------------------|------------------|-------------------------------------------------|
| Blanda et al<sup>10</sup> | CS            | LSO maintaining lumbar lordosis, exercises, activity restriction, posttreatment PT | Plain radiograph              | 15.5 (11–20)      | 53                             | 63                          | 20          | 25                | 6                | 74              | 26             | 99                | 26              | 99%                                               | 26%                          |
| Daniel et al<sup>20</sup> | CS            | TLSO, activity restriction, NSAIDs                                           | Plain radiograph              | 21 (13–31)         | 20                             | 29                          | 2           | 29                | 7%                | 97%             | 2              | 12                | 12%             | 2%          | 97%                                               | 26%                          |
| Edelmann<sup>41</sup>  | CS            | Milwaukee or Boston Brace                                                   | Plain radiograph              | 14.1 (10–18)       | 12                             | 12                          | 2           | 2                 | 7%                | 97%             | 2              | 12                | 17%             | 2%          | 16.7%                                             | 17%                          |
| Fujii et al<sup>23</sup> | CS            | Activity restriction, Darnon corset for 3 mo, isometric exercises for trunk muscles | CT                            | 13.8 (7–17)        | 114                            | 134                         | 31          | 50                | 9                 | 103             | 0              | 86                | 99%             | 26              | 99%                                               | 26%                          |
| Jackson et al<sup>20</sup> | CS            | Activity restriction                                                        | Plain radiograph              | 16.1 (11–19)       | 14                             | 15                          | 0           | 2                 | 0                 | 10              | 0              | 12                | 9%              | 2%          | 97%                                               | 26%                          |
| Letts et al<sup>17</sup> | Prospective case series          | Various combination of TLSO 23 h/d, rest, plaster vest, exercises           | Plain radiograph              | 13.4 (9–16)         | 12                             | 12                          | 5           | 10                | 0                 | 8               | 5              | 18                | 28%             | 2%          | 97%                                               | 26%                          |
| Miller et al<sup>46</sup> | Cohort        | Nonsurgical lumbar bracing 23 h/d, flexion only trunk strengthening after 2–4 wk relative rest, activity restriction | CT                            | 14.6 (12–20)       | 7                              | 11                          | 4           | 8                 | 0                 | 10              | 4              | 4                 | 18%             | 2%          | 97%                                               | 26%                          |
| Edelmann<sup>41</sup>  | CS            | Activity restriction, lumbar corset for 3–6 mo, after defect healed patients wore a lumbar sacral brace promoting flexion | Plain radiograph and/or CT    | 11.9 (5–8)          | 168                            | 185                         | 87          | 119               | 42                | 109             | 28             | 103              | 306             | 129         | 346                                               | 37.3%                        |
| Sairyo et al<sup>22</sup> | CS            | Activity restriction, plastic thoracolumbar brace                           | CT                            | 15 (11–18)          | 9                              | 13                          | 12          | 14                | 3                 | 12              | 0              | 3                 | 52%             | 2%          | 97%                                               | 26%                          |
| Sys et al<sup>43</sup>  | CS            | Boston Brace 23 h/d, hamstring stretching, ab strengthening, pelvic tibias   | CT                            | 17.2 (12–27)        | 19                             | 34                          | 28          | 45                | 11                | 34              | 28             | 45                | 62%             | 2%          | 97%                                               | 26%                          |
| Total                 |               |                                                                              |                               | 416                            | 508                           | 162                         | 236         | 354               | 224              | 217             | 68             | 92               | 126             | 446         | 251                                               | 847                          | 29.6%                                              | 29.6%                                      |

CS indicates retrospective case series; LSO, lumbosacral orthosis; NSAIDs, nonsteroidal anti-inflammatory drugs; PT, physical therapy; TLSO, thoracolumbar sacral orthosis.
defect(s) at the level of L5, and 70% of these patients were male. In the radiographic pool, the sample size of defects ranged from 12 to 346 defects (a bilateral spondylolysis was counted as 2 defects). The average age in the radiographic pool was 14.6 years, 82% of the defects occurred at L5, and 82% of the patients were male.

Clinical/Functional Outcome Pool

Fifteen observational studies reported on a total of 665 patients meeting our eligibility criteria. The mean age of patients included in studies included in the clinical/functional outcome group ranged from 11.4 to 18.1 years. The overall age range of patients included in the pooled clinical/functional outcome meta-analysis was from 1.6 to 35 years of age. The unweighted rate of success in individual studies ranged from 60% to 100%. When all 665 patients where pooled together in a random effects model with CI of 95%, the weighted rate of treatment success was 83.9% (95% CI; Fig. 2). The $I^2$ value for this analysis was 69.6%, indicating a moderate to high level of heterogeneity between studies.

FIGURE 2. Effect of nonoperative treatment on clinical outcome. There were 83.9% of subjects that had a successful clinical outcome after a minimum of 1-year follow-up. Asterisk (*) represents no success rate displayed because there were less than 20 subjects included.

FIGURE 3. Effect of nonoperative treatment with and without bracing on clinical outcome. In a subgroup analysis, 89.0% of subjects treated with a brace had a successful clinical outcome compared with 85.8% of subjects treated without a brace (no significant difference, $P = 0.75$). Asterisk (*) represents no success rate displayed because there were less than 20 subjects included.
We performed a subgroup analysis comparing 137 patients treated without a brace (not treated or only treated with activity restriction and/or exercises) to 334 patients whose treatment included bracing. The pooled and weighted rate of treatment success was 86% in patients without a brace and 89.0% in the pool of patients treated with a brace (Fig. 3). These were not significantly different ($P = 0.75$). The $I^2$ values for treatment without bracing and for treatment with bracing were 0% and 37.2% (implying no and low to moderate heterogeneity, respectively).

The one randomized controlled trial meeting our inclusion criteria showed a significant reduction in pain intensity and functional disability when comparing the effects of an exercise program training deep abdominal muscles and lumbar multifidus ($n = 21$) to a control group ($n = 21$), which performed physician-directed exercise such as swimming, walking, and gym work.

Radiographic Outcome Pool

Ten studies reported radiographic evidence of healing of the pars defects at unweighted rates ranging from less than 10% to 62%. The mean age of patients included in the radiographic outcome group ranged from 13.4 to 21 years. The overall range of patients included in the pooled radiographic meta-analysis was from 6 to 31 years of age. Of a total of 847 defects (a bilateral spondylolysis was considered as 2 defects), 251 defects demonstrated radiographic evidence of healing with a pooled and weighted success rate of 28.0% (Fig. 4). The $I^2$ value for this analysis was 85.7%, indicating high heterogeneity between studies. A subgroup analysis performed physician-directed exercise such as swimming, walking, and gym work.

---

FIGURE 4. Effect of all nonoperative treatment on radiographic outcome. There were 28.0% of all defects that healed with nonoperative treatment. Asterisk (*) represents no success rate displayed because there were less than 20 subjects included.

FIGURE 5. Healing of unilateral versus bilateral defects. In a subgroup analysis, 71.1% of unilateral defects healed, significantly more than bilateral defects (18.1%, $P < 0.0001$). Asterisk (*) represents no success rate displayed because there were less than 20 subjects included.

152
comparing radiographic healing in unilateral and bilateral defects identified 68 of 92 unilateral defects healed (71% healing rate in the random effects model). This is significantly greater than the 126 (18.1%) of 446 bilateral defects that healed (P < 0.0001; odds ratio, 7.20; CI, 4.21–12.37; Fig. 5).

In a separate subgroup analysis, we evaluated the effect of defect chronicity (categorized as acute, progressive, or chronic according to the Tokushima classification) on radiographic evidence of healing. A total of 162 of 236 acute lesions healed (a weighted rate of 68.1% in the random effects model), 54 (28.3%) of 224 progressive lesions healed, and 0 of the 217 of the terminal lesions healed (Fig. 6). Acute lesions were significantly more likely to heal compared with nonacute lesions (progressive and terminal) (P < 0.0001; odds ratio, 15.68; CI, 10.37–23.80) and were also significantly more likely to heal when compared with progressive lesions alone (P < 0.0001; odds ratio, 6.89; CI, 4.48–10.64). Progressive lesions were significantly more likely to heal than terminal lesions (P < 0.0001).

I² values for acute, progressive, and terminal defects were 32.3%, 90.9%, and 25.3%, respectively, indicating low, high, and low levels of heterogeneity.
Of the 10 studies in the radiographic outcome pool, only Fujii et al.\textsuperscript{23} reported on the relationship between slip percent and radiographic evidence of bony union. Fujii et al. found 3 unions of 65 total defects with greater than 5% spondylolisthesis (4.6%) and 37 unions of 174 total defects with less than 5% spondylolisthesis (21.3%). The rate of bony healing was significantly greater for defects with less than 5% spondylolisthesis compared with defects with more than 5% spondylolisthesis ($P < 0.005$; odds ratio, 5.58; CI, 1.57–23.63).

A funnel plot of the 15 observational studies with a clinical outcome suggests the presence of publication bias favoring reporting of cases with a positive outcome (Fig. 7). A funnel plot of the 10 studies evaluating healing of the defects suggests a slight publication bias favoring reporting of cases with a negative outcome (Fig. 8).

DISCUSSION

One of the primary findings of this meta-analysis is that within the context of at least 1 year of follow-up, children and young adults with spondylolysis (including those with up to 25% spondylolisthesis) treated nonoperatively are able to return to pain-free or near pain-free, unrestricted activities approximately 84% of the time. Analyses of radiographic studies indicate that acute lesions have a nearly 70% rate of healing, whereas terminal lesions do not achieve a bony union via nonoperative methods, and thus, the Tokushima classification would seem to influence radiographic prognosis. Furthermore, unilateral defects have been shown to heal significantly more often than bilateral lesions (71.1% vs 18.1%). Although a stable, bony union is the ideal goal of therapy, there are abundant reports of asymptomatic children with persistent pars defects.\textsuperscript{45,46} Remarkably, there was no significant difference in clinical outcome between patients treated with and without bracing.

Strengths and Limitations of the Study

Our results were derived almost exclusively from observational studies, which may be biased despite the care and integrity of their authors. Furthermore, most cases with a clinical outcome had a follow-up less than 10 years (range, 1–25 years). Thus, on balance, this analysis focused on relatively short-term follow-up results.

When all clinical data were pooled together (Fig. 2), there was a moderate to high level of heterogeneity between studies (69.6%), most likely due to the variability in treatment, follow-up time, and lesions included among the studies in this analysis. On the contrary, the subgroup analysis of the clinical effect of treatment including bracing ($F = 37.2\%$) compared with treatment without bracing ($F = 0\%$) had low levels of variability that strengthen the generalizability of this particular analysis. In the pool of all radiographic outcomes, there was a high degree of heterogeneity ($F = 85.7\%$). This can be understood when one considers that the studies pooled in this particular analysis varied widely in regard to the severity of the defects treated (acute vs terminal and unilateral vs bilateral) and the imaging study used to evaluate them (CT vs plain radiographic). Much of this heterogeneity can be explained by the subgroup analysis exploring the effects of chronicity, where acute defects ($F = 32.4\%$) and terminal defects ($F = 25.3\%$) had significantly different outcomes and low levels of heterogeneity.

In addition, the variability in treatment, the large span of time across which eligible studies were published (36 years), and the many countries of origin (8 from 4 continents: Australia, Belgium, England, Finland, Germany, Japan, Spain, Sweden, and the United States) add to the generalizability of our results. Furthermore, our rigorous search strategy across multiple databases, bibliographies, and meetings has greatly reduced the possibility for study selection bias. Funnel plots of the clinical and radiographic outcome groups suggest publication bias. However, it has been noted that publication bias is of less importance in meta-analyses of observational studies in situations where numerous sources of heterogeneity exist.\textsuperscript{47}

Clinical/Functional Outcome Pool

The 15 observational studies measuring clinical parameters had a pooled success rate of 83.9% with nonoperative management of spondylolysis. To determine the effectiveness of bracing, we compared patients treated with a brace to patients treated without a brace. As shown in Figure 3, these data were not significantly different and imply that bracing is not responsible for the improvement in clinical parameters reported with conservative treatment. Although these data are based on observational reports, we interpret this as a strong impetus to reconsider the status of bracing as standard treatment for spondylolysis. A large randomized controlled trial comparing activity restriction to bracing combined with activity restriction may clarify the discrepancy between the current accepted standard of care and the lack of published data supporting brace treatment when compared with other conservative treatments. On a related note, the use of a corset with activity restriction has been associated with high rates of healing of acute defects in 2 large case series.\textsuperscript{22,23} However, the use of a corset was discontinued in later studies by the same group of researchers\textsuperscript{42} after finite element analysis suggested a link between trunk rotation and spondylolysis.\textsuperscript{42}

Based on these findings, they categorically changed their nonoperative treatment protocol to include antilordotic thoracolumbar sacral orthosis bracing to restrict trunk rotation (K. Sairyo, personal communication, 2007). However, lumbo-sacral orthoses and corsets have not been shown to effectively minimize motion distal to L4-L5.\textsuperscript{48,49} Combined with the often benign natural history of spondylolysis,\textsuperscript{1} this may explain the results of the subgroup analysis comparing the clinical outcome of treatment with bracing to treatment without bracing. This implies that it is not the corset or brace causing the healing of defects nor the plastic brace causing the relief of symptoms, but the activity restriction and/or the commonly benign natural history of the condition.\textsuperscript{1} As Rush\textsuperscript{50} astutely commented, one must cautiously interpret the results of case series so as not to make the common error of post hoc ergo propter hoc (after this, therefore because of this). Thus, it is impossible to be certain that the clinical and radiographic outcomes reported in this study are due to true treatment effect or if they would have also occurred in the absence of any formal treatment at all.
In the randomized controlled trial of O’Sullivan et al, the experimental group training lumbar multifidus and deep abdominal muscles showed a significant reduction in pain intensity and functional disability levels when compared with a control group undergoing exercise such as swimming, walking, and gym work. The control group used was less than ideal and diminishes the value of this study when one considers that not a single study out of the other 21 reviewed in this article included such a general exercise plan as treatment. The combination of activity restriction, physical therapy, and bracing has been the standard of care that the experimental group would ideally have been compared against. The difference in pain and function between groups may not be due to a benefit from the therapeutic exercises but from the potential pain caused by undergoing exercises such as swimming and gym work. Repeated hyperextension in the breast stroke and butterfly stroke has been implicated as a potential cause of spondylolysis.31

Radiographic Outcome Pool
When we pooled the defects from all 10 studies in the radiographic pool, only 28% of the defects healed. However, this evaluation of the data was influenced by counting a bilateral defect as 2 defects and a unilateral lesion as 1 defect. When we compared healing of unilateral defects against bilateral defects, it was evident that unilateral defects are significantly more likely to heal with success rates of 71.1% and 18.1%, respectively (P < 0.0001; Fig. 5). In addition, when we compared defects pooled into acute, progressive, or terminal categories, it was evident that although a defect treated during its acute stage has a 68.1% chance of healing, progressive fractures are unlikely to heal (28.1%) and terminal defects will not heal with conservative treatment (Fig. 6). All 3 of these subgroups were significantly different from one another (P < 0.0001). We acknowledge that the Tokushima classification of acute, progressive, and terminal defects has not been validated.

Correlation of Positive Clinical Outcome With Radiographic Evidence of Healing of the Pars Defect
Based on the study of Fredrickson et al,1 a cohort of 30 patients who were found to have spondylolysis out of a population of 500 first-grade children has illustrated that a defect in the pars interarticularis is often asymptomatic. Of the 30 patients with spondylolysis, only 1 pain episode was noted in any of the subjects during childhood and adolescence. At a 45-year follow-up, this cohort showed no significant difference in pain and function when compared with the age-matched general population as measured by the Short form-36.46 Furthermore, the study of Muschik et al reported that during the observation of 86 active young athletes with spondylolysis or spondylolisthesis for an average of 4.8 years, not 1 of the children had symptoms.35 In addition, Blanda et al30 reported 37 patients with a successful functional outcome in the presence of a defect. On the other hand, Miller et al55 demonstrated the expected result that patients with a defect had greater low back pain on average than those without a defect. It is not clear why only a small percentage of pars defects are symptomatic.1,32

In summary, nonoperative treatment consisting of bracing, therapeutic exercises, and activity restriction has shown a high rate of success when measured by functional outcome. Bracing does not seem to influence the clinical outcome of treatment. In contrast to the clinical outcome, nonoperative treatment often does not result in bony union of the defects. Bony healing has been more likely to occur when conservative treatment was started early. Diagnostic techniques with greater sensitivity may identify patients earlier and ultimately lead to better outcomes. Nonoperative treatment has been unsuccessful in creating a bony union in terminal defects but may alleviate symptoms.

ACKNOWLEDGMENTS
The authors are deeply grateful to Koichi Sairyo, MD, PhD, and Seppo Seitsalo, MD, PhD, for generously sharing their expertise regarding spondylolysis and spondylolisthesis and additional data on published studies. They are also grateful to Joseph Congeni, MD; Karin Frenmered, MD; John Sarwark, MD; and Jan Sys, MD, for providing additional data on their published research. In addition, the authors thank Raghlav Tadepalli for his assistance on this project.

REFERENCES