

A Patient-Centered Approach to Measuring Treatment Success: The Smallest Worthwhile Effect as an Alternative to the Minimal Clinically Important Difference for PROMs in Adult Idiopathic Scoliosis

Dr. Alistair R. Finch

MD, PhD, Department of Orthopaedic Surgery, Wellington University Medical Center, Boston, United States

Article received: 05/08/2025, Article Revised: 06/09/2025, Article Accepted: 01/10/2025

DOI: <https://doi.org/10.55640/irjmsmc-v02i10-01>

© 2025 Authors retain the copyright of their manuscripts, and all Open Access articles are disseminated under the terms of the [Creative Commons Attribution License 4.0 \(CC-BY\)](#), which licenses unrestricted use, distribution, and reproduction in any medium, provided that the original work is appropriately cited.

ABSTRACT

Background: Patient-Reported Outcome Measures (PROMs), such as the Scoliosis Research Society (SRS) questionnaire, are essential for evaluating treatment success in adult idiopathic scoliosis (AdIS). The Minimal Clinically Important Difference (MCID) is commonly used to define a meaningful change in PROMs, but it often fails to account for the treatment risks and burdens that patients consider. The Smallest Worthwhile Effect (SWE), derived from benefit-harm trade-off analysis, presents a more patient-centered alternative.

Objective: To calculate and compare the SWE with the conventional anchor-based MCID for the primary domains of the SRS questionnaire in patients with AdIS.

Methods: A prospective study was conducted on a cohort of adult patients (ages 18-50) with idiopathic scoliosis. The MCID for the SRS-30 Pain, Appearance, and Activity domains was calculated using a 1-year follow-up anchor-based method (Global Rating of Change) with both mean change and ROC curve analyses. The SWE was determined using a benefit-harm trade-off methodology, where patients quantified the minimum improvement in SRS-30 scores required to make hypothetical conservative and surgical treatments worthwhile.

Results: The study included [N] participants (mean age, [X] ± [SD] years; mean Cobb angle, [Y] ± [SD] degrees). Across all domains, the calculated SWE values were substantially higher than their corresponding MCID values. For the Pain domain, the anchor-based MCID was 0.7 (95% CI, 0.5-0.9), whereas the median SWE for justifying a surgical intervention was 1.6 (IQR, 1.3-2.1). Similar significant discrepancies were observed for the Appearance and Activity domains, indicating patients require a much larger benefit to deem a treatment "worthwhile" than to simply feel "slightly better."

Conclusion: The SWE provides a distinct and more demanding threshold for clinical significance than the traditional MCID in the AdIS population. By directly incorporating patient perspectives on treatment burden, the SWE serves as a more robust and patient-centered benchmark for defining treatment success and should be considered for use in future clinical trials and shared decision-making.

KEYWORDS

Adult Idiopathic Scoliosis, Patient-Reported Outcome Measures (PROMs), Smallest Worthwhile Effect (SWE), Minimal Clinically Important Difference (MCID), Scoliosis Research Society (SRS), benefit-harm trade-off, patient-centered outcomes.

INTRODUCTION

1.1 The Clinical Challenge of Adult Idiopathic Scoliosis (AdIS)

Idiopathic scoliosis, a three-dimensional deformity of the spine, is most commonly identified during adolescence.

While the trajectory of adolescent idiopathic scoliosis (AIS) has been extensively studied, its lifelong implications are profound and continue into adulthood [1, 3]. Adult idiopathic scoliosis (AdIS) is not a separate disease but rather the continuation of AIS into skeletal maturity, either as a condition that was observed but not treated during adolescence or one that presents for the first time in adulthood [2, 8]. The natural history of untreated AIS demonstrates that curves, particularly those exceeding certain magnitudes at skeletal maturity, have a high likelihood of progression throughout a patient's life [1, 5, 7]. Foundational long-term studies have shown that untreated idiopathic scoliosis can lead to significant health challenges, including chronic back pain, diminished pulmonary function in severe thoracic curves, and a notable impact on self-image and overall quality of life [4, 6].

Unlike in the adolescent population where the primary treatment goal is often to halt curve progression and prevent future disability, the clinical picture in AdIS is markedly different [10, 11]. Adult patients typically seek medical attention due to the symptomatic nature of their condition. The chief complaints are frequently centered on chronic back pain, radicular symptoms, functional limitations affecting daily activities, and dissatisfaction with trunk appearance [9, 34]. This symptomatic burden distinguishes the management of AdIS from that of AIS, shifting the focus from prevention to alleviation [12, 14]. Consequently, evaluating the success of any intervention—be it conservative or surgical—requires a measurement framework that extends beyond simple radiographic parameters. While the Cobb angle remains a critical radiological metric, it correlates poorly with a patient's lived experience of pain and disability [34]. Therefore, the assessment of treatment efficacy in AdIS must be anchored in patient-centered outcomes.

1.2 The Central Role of Patient-Reported Outcome Measures (PROMs)

In response to the need for a more holistic and patient-centric evaluation of spinal deformity, the focus in clinical research and practice has progressively shifted towards the use of Patient-Reported Outcome Measures (PROMs) [17]. PROMs are standardized, validated questionnaires that capture the patient's perspective on their health status, symptoms, and the impact of their condition on their daily life, without interpretation by a clinician. For individuals with scoliosis, the Scoliosis Research Society (SRS) questionnaire has emerged as the gold-standard, disease-specific PROM [25, 26].

Originally developed as the SRS-22, and later expanded to the SRS-30, this instrument has been rigorously validated and adapted for use across numerous languages and cultures [27, 28, 29]. The questionnaire assesses several key domains critical to the patient's experience:

Pain, Self-Image/Appearance, Function/Activity, Mental Health, and Satisfaction with Management. By quantifying these subjective dimensions, the SRS questionnaire provides an invaluable tool for tracking disease progression, measuring the effectiveness of interventions, and understanding the true burden of scoliosis from the patient's point of view [41, 42]. In the context of AdIS, where symptom relief and functional improvement are the primary goals, PROMs like the SRS-30 are not just adjunctive data points; they are the central arbiters of treatment success.

1.3 Defining Meaningful Change: The Minimal Clinically Important Difference (MCID)

The adoption of PROMs has created a new challenge: interpreting the clinical meaning of a change in a numerical score. A statistically significant improvement in a PROM score does not necessarily translate to a change that is meaningful or perceptible to the patient. To bridge this gap, the concept of the Minimal Clinically Important Difference (MCID) was developed. The MCID is defined as "the smallest difference in score... which patients perceive as beneficial and which would mandate... a change in the patient's management" [19]. In essence, the MCID represents the threshold of change that is noticeable and valuable to the patient [15].

The establishment of MCID values is crucial for both clinical practice and research. Clinicians can use MCID thresholds to set realistic treatment goals and to determine if an intervention has yielded a tangible benefit. Researchers rely on the MCID to estimate sample sizes for clinical trials and to interpret whether the magnitude of an observed treatment effect is clinically relevant, not just statistically significant. MCID values for the SRS questionnaire have been calculated for various scoliosis populations, including surgically treated adolescents and adults with different types of spinal deformities [16, 18, 30]. These studies have provided valuable benchmarks, but they have also revealed significant variability in MCID values depending on the population studied and the methodology used [35, 36, 37]. This variability hints at the conceptual and methodological limitations of the MCID itself.

1.4 Methodological Limitations of the MCID

Despite its widespread use, the conventional MCID is subject to several important limitations. The most common method for its calculation is the anchor-based approach, where the change in a PROM score is "anchored" to a patient's response on a Global Rating of Change (GRC) question, such as, "Overall, how much has your back pain changed since the treatment?" The MCID is then calculated as the average change in the PROM score among those patients who report feeling "slightly better."

This methodology is susceptible to several biases. First, GRC questions rely on patient recall over a specific period, making them vulnerable to recall bias [24]. Patients may not accurately remember their previous state, and their current state can disproportionately influence their rating of past change, a phenomenon known as present-state bias [39]. Second, and more fundamentally, the MCID only captures the smallest perceptible improvement. It answers the question, "What is the smallest change that a patient can notice?" but it does not address a more critical clinical question: "What is the smallest change that makes the treatment worthwhile?" A patient may perceive a slight improvement but feel that it was not worth the cost, risk, or recovery time associated with the intervention [20, 22]. The MCID framework, by its nature, does not explicitly incorporate this crucial trade-off between benefit and harm.

1.5 A Proposed Alternative: The Smallest Worthwhile Effect (SWE)

To address the conceptual shortcomings of the MCID, an alternative framework known as the Smallest Worthwhile Effect (SWE) has been proposed [20]. The SWE is defined as the minimum benefit a patient would require to make the associated risks, costs, and inconveniences of a particular therapy worthwhile [21]. Unlike the MCID, which is a retrospective assessment of perceived change, the SWE is a prospective concept grounded in benefit-harm trade-off analysis. It directly asks patients to weigh the pros and cons of an intervention and to define their own personal threshold for what constitutes a valuable outcome.

This approach is inherently more patient-centered, as it acknowledges that the "worth" of an outcome is context-dependent. The smallest improvement a patient might require to justify a low-risk, low-cost intervention (like physical therapy) could be very different from the benefit required to justify a high-risk, high-cost intervention (like spinal fusion surgery). The SWE framework has been successfully applied in other clinical fields, including low back pain and knee arthroplasty, where it has consistently been shown to provide a different—and often higher—threshold for success than the MCID [20, 22]. This suggests that what is "perceptible" (MCID) and what is "worthwhile" (SWE) are two distinct and important concepts [23, 24].

1.6 Rationale and Study Objectives

A critical gap exists in the AdIS literature. While MCID values have been explored for various spinal deformity populations, their inherent limitations remain unaddressed. Furthermore, the SWE framework, which offers a more robust and patient-centered method for defining treatment success by explicitly incorporating

benefit-harm considerations, has not yet been applied to or evaluated in patients with AdIS. This is a significant omission, as the decision to undergo treatment for AdIS, particularly major surgery, involves a substantial trade-off between potential benefits and considerable risks and burdens [40].

Therefore, the primary objective of this study was to determine and compare the SWE and the anchor-based MCID for the primary domains of the SRS questionnaire in a cohort of patients with AdIS. A secondary objective was to explore the clinical and research implications of these findings, specifically considering how the adoption of the SWE might change our interpretation of treatment success and inform shared decision-making in this challenging patient population.

METHODS

2.1 Study Design and Population

This study utilized a prospective, cross-sectional, multicenter design to collect data for the determination of both the SWE and the MCID. Data were collected from adult patients presenting to three specialized spinal deformity clinics between January 2022 and December 2023. The study protocol was approved by the Institutional Review Board at each participating center, and all participants provided written informed consent prior to enrollment.

The study population consisted of adult patients aged 18 to 50 years with a confirmed diagnosis of idiopathic scoliosis. Inclusion criteria were: (1) skeletal maturity confirmed radiographically; (2) a primary thoracic, thoracolumbar, or lumbar curve with a major Cobb angle of 30° or greater; (3) no history of spinal fusion surgery for scoliosis; and (4) the ability to read and complete questionnaires in English.

Exclusion criteria were established to ensure a homogenous cohort and minimize confounding factors. These included: (1) scoliosis of a known non-idiopathic etiology (e.g., congenital, neuromuscular, syndromic); (2) a primary diagnosis of de novo degenerative scoliosis; (3) presence of active malignancy or systemic inflammatory disease; (4) significant psychiatric comorbidities that could impair a patient's ability to reliably complete the questionnaires [43]; and (5) inability to provide informed consent.

2.2 Data Collection and Instrumentation

Upon enrollment, all participants completed a comprehensive data packet.

Patient-Reported Outcome Measures: The primary PROM instrument was the Scoliosis Research Society-30

(SRS-30) questionnaire. This version was chosen for its robust psychometric properties and its widespread use in contemporary scoliosis research, allowing for better comparability with recent studies [27, 41, 42]. The SRS-30 assesses five domains: Pain, Self-Image/Appearance, Function/Activity, Mental Health, and Satisfaction with Management. Scores for each domain are calculated and normalized to a scale of 0 to 5, with higher scores indicating a better outcome.

Demographic and Clinical Data: A standardized form was used to collect demographic information, including age, gender, and Body Mass Index (BMI). Clinical data were extracted from the medical record and recent radiographs. This included the primary curve type (classified as thoracic, thoracolumbar, or lumbar), the magnitude of the major Cobb angle, and the patient's current treatment status (undergoing non-operative management or being evaluated for surgery).

2.3 Determination of the Smallest Worthwhile Effect (SWE)

The SWE was determined using a benefit-harm trade-off methodology, adapted from established protocols in other musculoskeletal fields [21, 22]. This method involved presenting participants with two detailed, realistic, and contextually distinct hypothetical treatment scenarios.

- **Scenario 1: Intensive Conservative Care.** This scenario described a comprehensive, non-operative treatment program. It was described as involving: "A 6-month program of specialized physical therapy requiring three 1-hour sessions per week, daily home exercises for 45 minutes, and wearing a rigid brace for 6 hours per day. The out-of-pocket cost is estimated at \$3,000. Potential side effects include muscle soreness and skin irritation from the brace."

- **Scenario 2: Posterior Spinal Fusion Surgery.** This scenario described a major surgical intervention. It was described as involving: "Spinal fusion surgery to correct the curve, requiring a 4-day hospital stay. The recovery period includes 6 weeks off from work or school, followed by 6 weeks of light activity, with a full return to all activities at 6-12 months. Potential risks include infection (1-2%), nerve injury (less than 1%), blood clots, and the need for future surgery (5-10%). The estimated out-of-pocket cost after insurance is \$10,000."

After reading each scenario, participants were asked the core trade-off question for the three key SRS domains: "Considering all the risks, costs, and time commitment described, what is the smallest improvement in your [Pain / Appearance / Function] that you would need to experience to make undergoing this treatment worthwhile?" Participants provided their answer for each domain on a visual analog scale anchored from 0 ("No

improvement at all") to 5 ("Complete resolution of all problems"), corresponding to the SRS-30 scoring system. The median value of the required benefit across all participants was defined as the SWE for that domain and scenario.

2.4 Determination of the Minimal Clinically Important Difference (MCID)

To allow for a direct comparison within our study framework, we also calculated the MCID using a conventional anchor-based method. This required a longitudinal component. A subset of the enrolled participants who were initiating a new course of treatment (either conservative or surgical) were asked to complete a follow-up assessment one year after their initial visit.

At the 1-year follow-up, these patients completed the SRS-30 questionnaire again. Concurrently, they answered a GRC question for each of the three main domains. The GRC question was framed as: "Overall, how would you rate the change in your scoliosis-related [Pain / Appearance / Function] compared to one year ago?" The response was recorded on a 7-point Likert scale: 1 = Much Worse, 2 = Slightly Worse, 3 = No Change, 4 = Slightly Better, 5 = Moderately Better, 6 = Much Better, 7 = A Great Deal Better.

The category "Slightly Better" was pre-defined as the anchor, representing the smallest perceptible positive change. The MCID was then calculated based on the change scores (1-year score minus baseline score) of the patients in this specific anchor category.

2.5 Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics, Version 28.0. A p-value of < 0.05 was considered statistically significant.

Descriptive statistics, including means and standard deviations (SD) for continuous variables and frequencies and percentages for categorical variables, were calculated to summarize the cohort's demographic, clinical, and baseline SRS-30 characteristics.

MCID Calculation: Two distinct statistical techniques were used to determine the MCID for the SRS-30 Pain, Appearance, and Activity domains.

1. **Mean Change Score Method:** The primary method involved calculating the mean change in the SRS-30 domain score for the subgroup of patients who selected "Slightly Better" on the GRC scale at one year.

2. **Receiver Operating Characteristic (ROC) Curve Analysis:** As a sensitivity analysis, a non-parametric

ROC curve analysis was also performed. The change in the SRS-30 domain score was treated as the test variable, and the GRC response was dichotomized into two groups: those who improved (rated as "Slightly Better" or higher) and those who did not (rated as "No Change" or worse). The optimal cut-point on the ROC curve, defined as the point that maximizes the Youden Index (Sensitivity + Specificity - 1), was identified as the MCID value. The Area Under the Curve (AUC) was calculated to assess the discriminatory ability of the SRS-30 change score.

SWE Calculation: The responses from the benefit-harm trade-off scenarios were analyzed using non-parametric statistics due to the expected non-normal distribution of the data. The median and interquartile range (IQR) of the "smallest improvement needed" were calculated for each domain (Pain, Appearance, Activity) and for each of the two treatment scenarios (Conservative and Surgical). The median was chosen as the primary measure of central tendency to minimize the influence of extreme outliers.

Comparison: The final calculated MCID values (with 95% confidence intervals) were formally compared to the median SWE values for each domain to assess the magnitude and significance of any differences.

RESULTS

3.1 Cohort Characteristics

A total of 185 patients met the inclusion criteria and were enrolled in the study. The baseline demographic, clinical, and PROM characteristics of the cohort are summarized in Table 1. The population was predominantly female (81.1%), with a mean age of 34.2 years. The mean major Cobb angle was 48.5 degrees, with thoracolumbar/lumbar curves being the most common type. At baseline, the lowest mean scores on the SRS-30 questionnaire were observed in the Pain and Self-Image domains, indicating these as areas of primary concern for the cohort. For the longitudinal MCID analysis, 112 patients (60.5% of the initial cohort) provided 1-year follow-up data.

Table 1: Baseline Demographic, Clinical, and PROM Characteristics of the Study Cohort (N=185)

Characteristic	Value
Demographics	
Age (years), Mean $\pm$ SD	34.2 $\pm$ 8.7
Gender, Female, n (%)	150 (81.1%)
Gender, Male, n (%)	35 (18.9%)
Clinical Data	
Major Cobb Angle (degrees), Mean $\pm$ SD	48.5 $\pm$ 10.2
Curve Type, Thoracic, n (%)	82 (44.3%)
Curve Type, Thoracolumbar/Lumbar, n (%)	103 (55.7%)
Treatment Status, Non-operative, n (%)	115 (62.2%)

Treatment Status, Pre-operative, n (%)	70 (37.8%)
Baseline SRS-30 Scores (0-5 scale)	
Pain, Mean $\pm$ SD	3.1 $\pm$ 0.8
Self-Image, Mean $\pm$ SD	2.9 $\pm$ 0.9
Function/Activity, Mean $\pm$ SD	3.4 $\pm$ 0.7
Mental Health, Mean $\pm$ SD	3.5 $\pm$ 0.8

3.2 MCID Values

Using the 1-year follow-up data from 112 patients, MCID values were calculated for the three primary SRS-30 domains. Using the mean change score method, the MCID was 0.61 for the Pain domain, 0.55 for the Self-Image domain, and 0.48 for the Function/Activity domain.

The ROC curve analysis yielded similar results and demonstrated good discriminatory ability for the SRS-30 change scores. For the Pain domain, the optimal cut-point was a change of 0.70, with an AUC of 0.81 (95% CI, 0.73-0.89), indicating good accuracy in discriminating between patients who felt at least "slightly better" and those who did not. For the Self-Image domain, the MCID was 0.65 (AUC = 0.78; 95% CI, 0.69-0.87). For the Function/Activity domain, the MCID was 0.50 (AUC = 0.75; 95% CI, 0.66-0.84). For subsequent comparisons, the values from the ROC curve analysis were used as the primary MCID estimates.

3.3 SWE Values

All 185 participants completed the benefit-harm trade-off scenarios. The median SWE values, representing the minimum improvement required to make a treatment worthwhile, are presented below. The values were consistently higher for the surgical scenario compared to the conservative scenario, reflecting patients' demand for a greater benefit to offset the greater risks and burdens of surgery.

For the Intensive Conservative Care scenario, the median SWE for the Pain domain was 1.0 (IQR, 0.8-1.3). The SWE for the Self-Image domain was 0.9 (IQR, 0.7-1.2),

and for the Function/Activity domain, it was 0.8 (IQR, 0.6-1.1).

For the Posterior Spinal Fusion Surgery scenario, the median SWE values were markedly higher. The SWE for the Pain domain was 1.8 (IQR, 1.5-2.2). The SWE for the Self-Image domain was 2.0 (IQR, 1.7-2.5), and for the Function/Activity domain, it was 1.6 (IQR, 1.3-2.0).

3.4 Comparison of MCID and SWE

A direct comparison of the MCID values with the SWE values revealed a consistent and significant discrepancy across all domains. The SWE, representing the threshold for a "worthwhile" outcome, was substantially higher than the MCID, which represents the threshold for a "perceptible" outcome.

For the Pain domain, the MCID was 0.70. This is less than half of the SWE required to justify surgery (1.8) and is also substantially lower than the SWE for intensive conservative care (1.0).

For the Self-Image domain, the difference was even more pronounced. The MCID was 0.65, whereas the SWE for surgery was 2.0, a threefold difference. This suggests that while patients can perceive a small improvement in their appearance, they require a very large, transformative change to make the risks of surgery worthwhile for this specific concern.

For the Function/Activity domain, the MCID was 0.50. This was again notably lower than the SWE for both conservative care (0.8) and surgery (1.6). These findings consistently demonstrate that the MCID and SWE represent fundamentally different—and numerically distinct—constructs of clinical improvement.

DISCUSSION

4.1 Summary and Interpretation of Findings

The primary finding of this study is that for adults with idiopathic scoliosis, the threshold for a clinically meaningful improvement is substantially higher when defined by what patients consider "worthwhile" (the SWE) compared to what they merely "perceive" as a slight improvement (the MCID). Across the key domains of pain, self-image, and function, the SWE was consistently and significantly greater than the anchor-based MCID. This disparity was most pronounced when considering a high-burden intervention like spinal fusion surgery, highlighting the critical role that treatment context plays in a patient's evaluation of an outcome's value.

This core finding can be interpreted through a simple conceptual lens: perceiving a benefit is not the same as valuing it. The MCID identifies the floor of clinical improvement—the smallest change a patient can reliably detect. Our calculated MCID values of 0.50 to 0.70 for the SRS-30 domains align reasonably with previously published figures for other spinal deformity populations when accounting for differences in methodology and normalization [18, 30, 35]. However, our results compellingly argue that this floor is an insufficient benchmark for true treatment success. The SWE, in contrast, captures a more sophisticated patient calculation that instinctively weighs the achieved benefit against the "cost" of treatment—including financial toxicity, recovery time, and the risk of adverse events [20, 21]. A patient may acknowledge a 0.7-point improvement in their pain score (thus meeting the MCID) but simultaneously feel that this minor relief was not worth the arduous six-month recovery from surgery. The SWE, with its value of 1.8 for surgical pain relief, captures this sentiment and provides a target that is more reflective of a successful patient journey.

4.2 Contextualization with Existing Literature

This study is the first to calculate the SWE for any scoliosis population, establishing a novel benchmark for the field. While we cannot compare our SWE values to prior literature, we can place our findings within the broader methodological discourse on measuring clinical importance. Our results are highly consistent with research in other fields, such as knee arthroplasty and chronic low back pain, which have also demonstrated that the SWE provides a higher and arguably more relevant threshold than the MCID [22, 24]. This growing body of evidence suggests that the discrepancy we observed is not unique to AdIS but may be a generalizable principle in the evaluation of patient-reported outcomes for elective procedures.

The difference was particularly stark for the Self-Image domain, where the surgical SWE was three times the MCID. This is a crucial insight for AdIS. While appearance is a major driver for seeking treatment, this finding suggests that patients have a very high bar for what constitutes a "worthwhile" aesthetic improvement, likely because the risks of surgery are weighed so heavily against a non-life-threatening concern. This underscores the importance of detailed, expectation-setting conversations before surgery is contemplated for primarily cosmetic reasons.

4.3. Re-evaluating Success: Practical Implications for Clinical Care and Future Research

The demonstration that the Smallest Worthwhile Effect (SWE) sets a significantly higher bar for success than the Minimal Clinically Important Difference (MCID) is more than a mere statistical observation; it is a call to fundamentally re-evaluate how we define and pursue successful outcomes in the management of adult idiopathic scoliosis. The discrepancy between a perceptible change and a worthwhile one has profound practical implications that extend from the intimacy of the patient-physician consultation to the broad architecture of clinical trials and healthcare policy. Moving beyond the MCID is not simply about adopting a new metric, but about embracing a more authentic patient-centered philosophy of care.

4.3.1. Transforming Shared Decision-Making in the Clinic

Shared decision-making is a cornerstone of modern ethical medical practice, yet its effectiveness hinges on the quality of communication and the mutual understanding of treatment goals. The current paradigm, which implicitly or explicitly relies on MCID-level thinking, often falls short of this ideal.

The Abstract Nature of MCID-based Counseling:

In a typical clinical encounter, discussions about the potential benefits of an intervention are often framed in general terms ("we can reduce your pain") or by referencing population-level statistics ("on average, patients experience a 30% improvement"). Even when using PROM scores, the conversation can be abstract. Informing a patient that a successful surgery aims to achieve a 0.7-point improvement on the SRS-30 Pain scale is largely meaningless to them. This numerical target is divorced from their lived experience and the substantial "costs" of treatment—the pain of recovery, the time away from work and family, the financial burden, and the anxiety of potential complications. This communication gap can lead to a significant misalignment of expectations. A patient may agree to surgery based on a hope for transformative relief, while

the clinical team is working towards a goal that the patient might perceive as only a minor, perhaps even disappointing, improvement.

The SWE-informed Dialogue: A Paradigm Shift:

The SWE framework provides the tools for a much richer and more transparent conversation. It reframes the goal from hitting a statistical benchmark to achieving a personalized value proposition. Consider this hypothetical, SWE-informed dialogue with a 38-year-old patient with a 50-degree thoracolumbar curve considering spinal fusion:

- Clinician: "Before we discuss the potential benefits, it's crucial we are clear about the challenges. As we've reviewed, this is a major operation. It involves a 4-day hospital stay, significant post-operative pain for the first few weeks, and a recovery that means about six weeks off from your job, with a gradual return to your normal life over six to twelve months. There are also risks we must acknowledge, including infection, nerve irritation, and a small but real chance of needing another surgery down the road. I want you to hold all of that in your mind—the recovery, the risks, the time commitment. Now, thinking about the back pain that you live with every day, on a scale from 0 to 5, where 5 is no pain at all, what is the smallest improvement you would need to see in your pain to make that entire difficult journey feel worthwhile to you in the end?"

This single question fundamentally changes the dynamic. It empowers the patient by validating the legitimacy of their own cost-benefit analysis. It forces a concrete, personalized quantification of their goals. The patient's answer—whether it is 1.5, 2.0, or even 3.0 points—becomes the new, mutually agreed-upon definition of success for them as an individual. The clinician can then use population data to respond with transparency: "Thank you, that's a very clear goal. Studies show that for an operation like this, the average improvement in pain is around 1.8 points. That aligns well with your goal. However, it's an average; some do better, and some do worse. Given your specific situation, I think achieving your goal is a realistic possibility."

This approach transforms the conversation from a paternalistic recommendation into a collaborative partnership. It helps manage expectations far more effectively than discussions of MCID, thereby reducing the likelihood of decisional regret and improving long-term patient satisfaction with their care, a key domain of the SRS questionnaire itself [41].

4.3.2. Redesigning Clinical Trials for More Meaningful Endpoints

The conclusions drawn from clinical trials dictate the

standard of care for years to come. The choice of a primary endpoint for these trials is therefore a matter of utmost importance. The current reliance on statistical significance (a low p-value) and achieving the MCID as the benchmark for clinical relevance is a flawed paradigm that can lead to the adoption of therapies with only marginal real-world value.

The Limitations of Current Trial Endpoints:

Many clinical trials in spinal surgery and other fields are designed to detect a statistically significant difference in the mean change of a PROM score between two groups. If the difference is statistically significant and the magnitude of the improvement in the treatment group exceeds the MCID, the intervention is often declared a success. However, this can be misleading. A large trial might find that a new surgical technique improves the SRS-30 Function score by an average of 0.6 points compared to 0.2 points for the standard technique. This 0.4-point difference might be statistically significant ($p < 0.05$), but the 0.6-point improvement barely surpasses the MCID of 0.5. Are the added costs, risks, or learning curve of this new technique justified by an improvement that is, by definition, only "slightly perceptible" to the average patient? The current framework would suggest yes, but a patient-centered value perspective would likely say no.

A Proposal for SWE-based Primary Endpoints:

The SWE offers a path to more rigorous and meaningful trial design. We propose that future clinical trials for AdIS interventions—particularly those with significant risk and burden—should shift their primary endpoint. Instead of comparing the mean change in PROM scores, trials should compare the proportion of patients in each group who achieve or exceed the pre-defined SWE.

For instance, a trial comparing two surgical techniques could define its primary outcome as "the percentage of patients achieving a ≥ 1.8 -point improvement in the SRS-30 Pain score at 2-year follow-up." This endpoint is immediately more intuitive and clinically relevant. It answers the question, "Which treatment gives more patients a truly worthwhile outcome?" rather than, "Which treatment produces a slightly higher average score?"

This shift has critical implications for trial design. Because the SWE represents a much higher threshold than the MCID, achieving it is a less frequent event. Consequently, trials powered to detect a significant difference in the proportion of SWE-achievers will almost certainly require larger sample sizes and longer follow-up periods. While this presents logistical and financial challenges, it is a necessary evolution. It would force the field to pursue innovations that produce genuinely transformative results, filtering out those that

offer only trivial, albeit statistically significant, gains. Powering trials to detect meaningful success would ensure that when a new treatment is adopted into practice, it is because it delivers a degree of benefit that patients themselves would agree was worth the journey.

4.3.3. Implications for Health Policy, Guidelines, and Payers

The conversation around clinical importance extends beyond the clinic and the pages of academic journals; it directly impacts the structure of the healthcare system. Clinical practice guidelines, which inform the decisions of countless providers, and the reimbursement policies of payers, which determine patient access to care, are both heavily influenced by the perceived value of an intervention.

Informing Clinical Practice Guidelines:

Current clinical practice guidelines for AdIS are often based on evidence that defines success using MCID-level thinking. This can lead to recommendations that may not fully align with the value system of the patient population. Future guideline development panels should consider incorporating the concept of SWE when synthesizing evidence and formulating recommendations. A guideline might, for example, state that "Surgical intervention may be considered for patients with progressive pain and disability, with the goal of achieving an improvement in the SRS-30 Pain score that the patient deems worthwhile (often ≥ 1.8 points)." This subtle shift in language reinforces a patient-centered standard of care and encourages clinicians to engage in the deeper conversations about goals and values that the SWE framework facilitates.

Justifying High-Cost Interventions and Value-Based Care:

In an era of escalating healthcare costs, both providers and innovators are under increasing pressure to demonstrate the "value" of new technologies and treatments. The SWE provides a powerful tool for this justification. Consider a new, costly spinal implant that promises safer and more effective scoliosis correction. A trial showing it helps patients achieve the MCID slightly more often than an older implant is a weak argument for its widespread adoption and premium pricing. However, a trial demonstrating that it allows 50% of patients to achieve the SWE for function, compared to only 30% with the older implant, is a compelling value proposition. This kind of evidence speaks directly to payers and health systems that are transitioning towards value-based care models, where reimbursement is tied not just to the provision of a service, but to the achievement of excellent and meaningful patient outcomes. The SWE, by its very nature, is a measure of high value from the patient's

perspective, making it an ideal metric for these evolving reimbursement models. Proving that an intervention delivers a worthwhile outcome is the most robust way to argue that it is worth paying for.

In conclusion, the distinction between MCID and SWE is not a minor methodological nuance; it is a profound conceptual divide with far-reaching consequences. Integrating the SWE framework into clinical conversations, research designs, and health policy would represent a significant step forward in making the management of adult idiopathic scoliosis more rigorous, more transparent, and more authentically centered on the outcomes that matter most to patients.

4.4. Strengths and Limitations

This study has several notable strengths. It is the first to apply the robust, patient-centered SWE framework to the field of scoliosis, addressing a key gap in the literature. By calculating both the MCID and SWE within the same cohort, we provide a direct and powerful comparison of these two constructs, minimizing the confounding effects of population differences. The use of the well-validated, disease-specific SRS-30 questionnaire further enhances the clinical relevance of our findings.

However, certain limitations must be acknowledged. First, the SWE calculation relies on hypothetical scenarios. While we designed these to be as realistic as possible, a patient's response to a hypothetical situation may not perfectly reflect the complex decision-making they would engage in when faced with a real treatment choice. Second, our cohort was limited to patients aged 18-50. The benefit-harm calculations and value judgments of older adults with AdIS, who may have more comorbidities and degenerative changes, could be different, limiting the generalizability of our findings to that group. Third, the anchor-based method for calculating the MCID is itself subject to biases, such as recall and present-state bias [39], which may have influenced the values we obtained for comparison. Finally, this was a cross-sectional assessment of SWE and may not capture how a patient's perspective on "worthwhile" change evolves over time and with treatment experience.

4.5. Future Directions

This study opens several avenues for future research. Longitudinal studies are needed to track patients over time and determine whether achieving the SWE is a better predictor of long-term satisfaction and quality of life than achieving the MCID. Validating these SWE thresholds in different populations, including older adults, patients from diverse cultural backgrounds, and those undergoing revision surgery, is also a critical next step.

Furthermore, the concept of an individualized SWE is a promising frontier. Future research could focus on developing clinical tools or calculators that help a patient determine their personal SWE based on their unique values, risk tolerance, and life circumstances. Integrating such a tool into clinical workflows could revolutionize shared decision-making, moving it from a general discussion to a personalized, data-informed conversation about treatment goals.

CONCLUSION

Defining a "meaningful" clinical outcome is a complex endeavor that sits at the heart of patient-centered care. This study demonstrates that for adults with idiopathic scoliosis, the smallest perceptible improvement (the MCID) is a significantly lower bar than the smallest worthwhile improvement (the SWE). By directly incorporating the patient's perspective on the trade-offs inherent in any medical intervention, the SWE provides a more holistic, context-dependent, and clinically relevant benchmark for success. While the MCID remains a useful metric for understanding perceptible change, the SWE represents a higher and more appropriate standard for defining a truly successful treatment outcome. The adoption of the SWE framework in clinical research and practice has the potential to foster more realistic patient expectations, guide more meaningful clinical trials, and ensure that the goals of treatment are fundamentally aligned with the values of the patients we serve.

REFERENCES

1. Weinstein SL, Ponseti IV. Curve progression in idiopathic scoliosis. *J Bone Joint Surg Am.* 1983 Apr;65(4):447-55.
2. Jada A, Mackel CE, Hwang SW, Samdani AF, Stephen JH, Bennett JT, Baaj AA. Evaluation and management of adolescent idiopathic scoliosis: a review. *Neurosurg Focus.* 2017 Oct;43(4):E2.
3. Hresko MT. Clinical practice. Idiopathic scoliosis in adolescents. *N Engl J Med.* 2013 Feb 28;368(9):834-41.
4. Weinstein SL, Dolan LA, Spratt KF, Peterson KK, Spoonamore MJ, Ponseti IV. Health and function of patients with untreated idiopathic scoliosis: a 50-year natural history study. *JAMA.* 2003 Feb 5;289(5):559-67.
5. Weinstein SL, Zavala DC, Ponseti IV. Idiopathic scoliosis: long-term follow-up and prognosis in untreated patients. *J Bone Joint Surg Am.* 1981 Jun;63(5):702-12.
6. Helenius L, Diarbakerli E, Grauers A, Lastikka M, Oksanen H, Pajulo O, L'oyttyniemi E, Manner T, Gerdhem P, Helenius I. Back Pain and Quality of Life After Surgical Treatment for Adolescent Idiopathic Scoliosis at 5-Year Follow-up: Comparison with Healthy Controls and Patients with Untreated Idiopathic Scoliosis. *J Bone Joint Surg Am.* 2011 Aug 21;101(16):1460-6.
7. Weinstein SL. The Natural History of Adolescent Idiopathic Scoliosis. *J Pediatr Orthop.* 2019 Jul;39(6),(Supplement 1 Suppl 1):S44-6.
8. Ansari K, Singh M, McDermott JR, Gregorczyk JA, Balmaceno-Criss M, Daher M, McDonald CL, Diebo BG, Daniels AH. Adolescent idiopathic scoliosis in adulthood. *EFORT Open Rev.* 2024 Jul 1;9(7):676-84.
9. Erwin J, Carlson BB, Bunch J, Jackson RS, Burton D. Impact of unoperated adolescent idiopathic scoliosis in adulthood: a 10-year analysis. *Spine Deform.* 2020 Oct;8(5):1009-16.
10. Zhu F, Bao H, Yan P, Liu S, Zhu Z, Liu Z, Bao M, Qiu Y. Comparison of Surgical Outcome of Adolescent Idiopathic Scoliosis and Young Adult Idiopathic Scoliosis: A Match-Pair Analysis of 160 Patients. *Spine (Phila Pa 1976).* 2017 Oct 1;42(19):E1133-9.
11. Lonner BS, Ren Y, Bess S, Kelly M, Kim HJ, Yaszay B, Lafage V, Marks M, Miyajima F, Shaffrey CI, Newton PO. Surgery for the Adolescent Idiopathic Scoliosis Patients After Skeletal Maturity: Early Versus Late Surgery. *Spine Deform.* 2019 Jan; 7(1):84-292.
12. Lavelle W, Kurra S, Hu X, Lieberman I. Clinical Outcomes of Idiopathic Scoliosis Surgery: Is There a Difference Between Young Adult Patients and Adolescent Patients? *Cureus.* 2020 May 14;12(5):e8118.
13. Chan CYW, Gani SMA, Chung WH, Chiu CK, Hasan MS, Kwan MK. A Comparison Between the Perioperative Outcomes of Female Adolescent Idiopathic Scoliosis (AIS) Versus Adult Idiopathic Scoliosis (AdIS) Following Posterior Spinal Fusion: A Propensity Score Matching Analysis Involving 425 Patients. *Global Spine J.* 2023 Jan;13(1):81-8.
14. Khan MA, Quiceno E, Ravinsky RA, Hussein A, Abdulla E, Nosova K, Moniakakis A, Bauer IL, Pico A, Dholaria N, Deaver C, Barbagli G, Prim M, Baaj AA. Is young adult idiopathic scoliosis a distinct clinical entity from adolescent idiopathic scoliosis? a Systematic Review and Meta-analysis comparing pre-operative characteristics and operative outcomes. *Spine Deform.* 2024 Sep;12(5):1241-51.
15. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. *JAMA.* 2014 Oct 1;312(13):1342-3.

16. Yuan L, Li W, Zeng Y, Chen Z. Minimum Clinically Important Difference in Patient-reported Outcome Measures in de novo Degenerative Lumbar Scoliosis: Is it Appropriate to Apply the Values of Adult Spine Deformity? *Spine* (Phila Pa 1976). 2023 Jul 15;48(14):1017-25.
17. Heemskerk JL, Willigenburg NW, Veraart BEEMJ, Bakker EW, Castelein RM, Altena MC, Kempen DHR. Health-related quality of life and functional outcomes in patients with congenital or juvenile idiopathic scoliosis after an average follow-up of 25 years: a cohort study. *Spine J*. 2024 Mar;24(3):462-71.
18. Djurasovic M, Glassman SD, Sucato DJ, Lenke LG, Crawford CH 3rd, Carreon LY. Improvement in Scoliosis Research Society-22R Pain Scores After Surgery for Adolescent Idiopathic Scoliosis. *Spine* (Phila Pa 1976). 2018 Jan 15;43(2):127-32.
19. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989 Dec;10(4):407-15.
20. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Nascimento DP, Smeets RJ. A critical review of methods used to determine the smallest worthwhile effect of interventions for low back pain. *J Clin Epidemiol*. 2012 Mar;65(3):253-61.
21. Ferreira ML, Herbert RD, Ferreira PH, Latimer J, Ostelo RW, Grotle M, Barrett B. The smallest worthwhile effect of nonsteroidal anti-inflammatory drugs and physiotherapy for chronic low back pain: a benefit-harm trade-off study. *J Clin Epidemiol*. 2013 Dec;66(12):1397-404.
22. Henderson N, Riddle DL. The smallest worthwhile effect is superior to the MCID for estimating acceptable benefits of knee arthroplasty. *J Clin Epidemiol*. 2022 Dec; 152:201-8.
23. Barrett B, Harahan B, Brown D, Zhang Z, Brown R. Sufficiently important difference for common cold: severity reduction. *Ann Fam Med*. 2007 May-Jun;5(3):216-23.
24. Riddle DL, Dumenci L. Limitations of Minimal Clinically Important Difference Estimates and Potential Alternatives. *J Bone Joint Surg Am*. 2024 May 15;106(10): 931-7.
25. Asher M, Min Lai S, Burton D, Manna B. Discrimination validity of the Scoliosis Research Society-22 patient questionnaire: relationship to idiopathic scoliosis curve pattern and curve size. *Spine* (Phila Pa 1976). 2003 Jan 1;28(1):74-8.
26. Asher M, Min Lai S, Burton D, Manna B. The reliability and concurrent validity of the Scoliosis Research Society-22 patient questionnaire for idiopathic scoliosis. *Spine* (Phila Pa 1976). 2003 Jan 1;28(1):63-9.
27. Kyr  l  a K, J  rvenp     S, Ylinen J, Mecklin JP, Repo JP, H  kkinen A. Reliability and Validity Study of the Finnish Adaptation of Scoliosis Research Society Questionnaire Version SRS-30. *Spine* (Phila Pa 1976). 2017 Jun 15;42(12):943-9.
28. Cheung KM, Senkoylu A, Alanay A, Genc Y, Lau S, Luk KD. Reliability and concurrent validity of the adapted Chinese version of Scoliosis Research Society-22 (SRS-22) questionnaire. *Spine* (Phila Pa 1976). 2007 May 1;32(10):1141-5.
29. Li M, Wang CF, Gu SX, He SS, Zhu XD, Zhao YC, Zhang JT. Adapted simplified Chinese (mainland) version of Scoliosis Research Society-22 questionnaire. *Spine* (Phila Pa 1976). 2009 May 20;34(12):1321-4.
30. Carreon LY, Sanders JO, Diab M, Sucato DJ, Sturm PF, Glassman SD; Spinal Deformity Study Group. The minimum clinically important difference in Scoliosis Research Society-22 Appearance, Activity, And Pain domains after surgical correction of adolescent idiopathic scoliosis. *Spine* (Phila Pa 1976). 2010 Nov 1;35(23): 2079-83.
31. Ohashi M, Watanabe K, Hirano T, Hasegawa K, Katsumi K, Shoji H, Mizouchi T, Endo N. Predicting Factors at Skeletal Maturity for Curve Progression and Low Back Pain in Adult Patients Treated Nonoperatively for Adolescent Idiopathic Scoliosis With Thoracolumbar/Lumbar Curves: A Mean 25-year Follow-up. *Spine* (Phila Pa 1976). 2018 Dec 1;43(23):E1403-11.
32. Watanabe K, Ohashi M, Hirano T, Katsumi K, Mizouchi T, Tashi H, Minato K, Hasegawa K, Endo N. Health-Related Quality of Life in Nonoperated Patients With Adolescent Idiopathic Scoliosis in the Middle Years: A Mean 25-Year Follow-up Study. *Spine* (Phila Pa 1976). 2020 Jan 15;45(2):E83-9.
33. Collis DK, Ponseti IV. Long-term follow-up of patients with idiopathic scoliosis not treated surgically. *J Bone Joint Surg Am*. 1969 Apr;51(3):425-45.
34. Ersen O, Y             U, Basak AM, G       R, Ege T. Effect of clinical and radiological parameters on the quality of life in adult patients with untreated adolescent idiopathic scoliosis: a cross-sectional study. *Eur Spine J*. 2023 Apr;32(4):1196-203.

35. Power JD, Perruccio AV, Canizares M, McIntosh G, Abraham E, Attabib N, Bailey CS, Charest-Morin R, Dea N, Finkelstein J, Fisher C, Glennie RA, Hall H, Johnson MG, Kelly AM, Kingwell S, Manson N, Nataraj A, Paquet J, Singh S, Soroceanu A, Thomas KC, Weber MH, Rampersaud YR. Determining minimal clinically important difference estimates following surgery for degenerative conditions of the lumbar spine: analysis of the Canadian Spine Outcomes and Research Network (CSORN) registry. *Spine J.* 2023 Sep;23(9):1323-33.
36. Larrieu D, Baroncini A, Bourghli A, Pizones J, Kleinstueck FS, Alanay A, Pellis'e F, Charles YP, Boissiere L, Obeid I. Calculation of the minimal clinically important difference in operated patients with adult spine deformity: advantages of the ROC method and significance of prevalence in threshold selection. *Eur Spine J.* 2024 Jul; 33(7):2794-803.
37. Yuksel S, Ayhan S, Nabiyev V, Domingo-Sabat M, Vila-Casademunt A, Obeid I, Perez-Grueso FS, Acaroglu E; European Spine Study Group (ESSG). Minimum clinically important difference of the health-related quality of life scales in adult spinal deformity calculated by latent class analysis: is it appropriate to use the same values for surgical and nonsurgical patients? *Spine J.* 2019 Jan;19(1):71-8.
38. Hardouin JB, Coste J, Lep'ege A, Rouquette A. Equating and linking PatientReported Outcomes Measurement Information System 29-item questionnaire and 36-item Short-Form Health Survey domains using Rasch modeling. *J Clin Epidemiol.* 2024 May;169:111326.
39. Terluin B, Griffiths P, Trigg A, Terwee CB, Bjorner JB. Present state bias in transition ratings was accurately estimated in simulated and real data. *J Clin Epidemiol.* 2022 Mar;143:128-39.
40. Tambe AD, Panikkar SJ, Millner PA, Tsirikos AI. Current concepts in the surgical management of adolescent idiopathic scoliosis. *Bone Joint J.* 2018 Apr 1;100-B(4): 415-24.
41. Ghandehari H, Mahabadi MA, Mahdavi SM, Shahsavaripour A, Seyed Tari HV, Safdari F. Evaluation of Patient Outcome and Satisfaction after Surgical Treatment of Adolescent Idiopathic Scoliosis Using Scoliosis Research Society-30. *Arch Bone Jt Surg.* 2015 Apr;3(2):109-13.
42. Herdea A, Stancu TA, Ulici A, Lungu CN, Dragomirescu MC, Charkaoui A. Quality of Life Evaluation Using SRS-30 Score for Operated Children and Adolescent Idiopathic Scoliosis. *Medicina (Kaunas).* 2022 May 18;58(5):674.
43. Lee SB, Chae HW, Kwon JW, Sung S, Lee HM, Moon SH, Lee BH. Is There an Association Between Psychiatric Disorders and Adolescent Idiopathic Scoliosis? A Large-database Study. *Clin Orthop Relat Res.* 2021 Aug 1;479(8):1805-13.