

# Reliability of Pelvic Obliquity Measurements in Non-Ambulatory Neuromuscular Scoliosis: A Comparison of the Osebold and Maloney Methods

**Daiki Urayama**

The University of Tokyo

**Yuki Taniguchi**

`taniguchi-tky@umin.ac.jp`

The University of Tokyo

**Teppei Suzuki**

Kobe Medical Center

**Masaaki Ito**

Kobe Medical Center

**Koki Uno**

Kobe Medical Center

**Hiroyuki Nakarai**

The University of Tokyo

**Hideki Nakamoto**

The University of Tokyo

**So Kato**

The University of Tokyo

**Sakae Tanaka**

The University of Tokyo

**Yasushi Oshima**

The University of Tokyo

---

## Research Article

**Keywords:** Maloney, Osebold, pelvic obliquity, postural variation, assisted seated radiography

**Posted Date:** June 15th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-9022555/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.  
[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Abstract

## Purpose

Postural variation during assisted seated radiography may affect the accuracy of pelvic obliquity (PO) measurements in non-ambulatory patients with neuromuscular scoliosis. This study compared the Osebold and Maloney methods.

## Methods

Postoperative radiographs obtained in patients aged  $\geq 15$  years were retrospectively evaluated in 35 non-ambulatory patients with neuromuscular scoliosis who had undergone posterior spinal fusion from the upper thoracic spine to the pelvis. PO was measured on three postoperative seated whole-spine radiographs from distinct follow-up visits by three independent raters using the Osebold and Maloney methods. Intersession variability was defined as the difference between the maximum and minimum PO values ( $\Delta PO$ ). Test-retest reliability was assessed using intraclass correlation coefficients (ICC) and minimal detectable change at the 95% confidence level ( $MDC_{95}$ ). Between-method agreement and systematic bias were evaluated using paired analyses and Bland-Altman plots.

## Results

Median  $\Delta PO$  was significantly smaller with the Maloney method than with the Osebold method ( $2.5^\circ$  vs  $8.0^\circ$ ,  $p < 0.001$ ). The Maloney method exhibited excellent intersession reliability (ICC: 0.991 vs 0.885) and a lower  $MDC_{95}$  ( $4.9^\circ$  vs  $13.3^\circ$ ). Using a predefined clinical threshold of  $\Delta PO < 5^\circ$ , the Maloney method yielded a higher clinical success rate (91% vs 32%). Comparison analyses revealed a consistent directional bias, with the Osebold method yielding smaller PO values with wide limits of agreement.

## Conclusion

The Maloney method yielded more stable, readily detectable results, whereas the Osebold method tended to systematically underestimate PO. These findings support the use of the Maloney method for longitudinal PO assessment.

## Introduction

Neuromuscular scoliosis (NMS) is often characterized by progressive spinal deformity, marked coronal imbalance, and pelvic obliquity (PO), particularly in non-ambulatory patients [1]. Progressive PO reduces trunk stability and contributes to impaired sitting tolerance, increased pain, pressure injuries, and caregiver burden [2–4]. As deformity advances, surgical correction is often required. However, whether fusion should be extended to the pelvis remains debated [5]. Substantial preoperative PO is widely considered an indication for pelvic fixation, and accurate PO assessment is essential for selecting appropriate distal fusion levels [6]. Furthermore, postoperative progression of residual PO, particularly after floating fusion in which the construct does not include pelvic fixation, may result in pelvic

imbalance and necessitate revision surgery [7]. Therefore, precise PO evaluation is required throughout perioperative management.

Obtaining standardized seated anteroposterior radiographs in non-ambulatory patients with NMS is technically challenging. These patients often require assistance to maintain an upright posture, and the degree of support may vary from one radiographic session to another. Even minor differences in seated posture can alter measured PO angles. Therefore, clarifying the impact of intersession posture differences on PO measurements is clinically relevant for accurate preoperative assessment and reliable postoperative monitoring. Among the various methods reported for assessing PO, the Osebold and Maloney methods are most widely used in clinical practice [8–10]. The Osebold method defines PO as the angle between the superior iliac crest line and a horizontal baseline [11] (Fig. 1), while the Maloney method is defined as the angle between a vertical line perpendicular to the line connecting both iliac crests and the line connecting the centers of T1 and S1 [12] (Fig. 1). Both methods have demonstrated comparable inter- and intraobserver reliability in previous studies, with no clear superiority of either method [8, 13]. Nonetheless, the extent to which posture-related variation affects measurement consistency has not been fully investigated. In the Osebold method, the superior iliac crest line is referenced to a fixed horizontal baseline, making the measurement sensitive to small differences in trunk support (Fig. 1). Conversely, in the Maloney method, both the T1-S1 axis and the iliac crest line shift in a coordinated manner with posture changes, which may yield smaller intersession angular differences compared to that via the Osebold method (Fig. 1). Accordingly, it was hypothesized that the Maloney method is less influenced by posture than the Osebold method and may be more suitable for capturing longitudinal PO change. To test this hypothesis, this study aimed to compare the intersession variability of the Osebold and Maloney methods in non-ambulatory patients with NMS treated with spinal fusion with pelvic fixation and to evaluate the stability of each method across repeated radiographic sessions.

Clarifying the intersession test–retest reliability of these methods is essential not only for accurate preoperative assessment of PO and appropriate selection of distal fusion levels, but also for determining whether postoperative changes in PO reflect true deformity progression rather than postural variation.

## Materials and Methods

### Patient Selection

This retrospective observational study was conducted at two high-volume spine centers and included non-ambulatory patients with NMS who underwent posterior spinal fusion extending from the upper thoracic spine to the pelvis. Institutional review board approval was obtained at both participating institutions, and informed consent was obtained from patients or their legal guardians, as required. In accordance with institutional ethics committee regulations, an opt-out consent procedure was implemented for earlier cases. Patients were eligible for inclusion if they met the following criteria: (1) previous posterior spinal fusion with pelvic fixation; (2) at least three seated anteroposterior whole-spine radiographs obtained at distinct postoperative follow-up visits; (3) 15 years of age or older at the time of

imaging; and (4) no radiographic evidence of postoperative instrumentation failure, including screw loosening or implant breakage. This cohort was selected to support the primary study objective, which was to evaluate the stability of different PO measurement methods. Restricting the analysis to patients with intact pelvic fixation minimized the likelihood of true longitudinal changes in PO due to crankshaft phenomenon, distal adding-on, or other mechanical failures, thereby allowing observed differences to predominantly reflect measurement variability rather than true anatomical change. Patients were excluded if they had an insufficient number of postoperative radiographs or radiographic data that were incomplete or inadequate for longitudinal analysis.

A power analysis was performed to determine the minimum sample size required for reliability assessment. An expected intraclass correlation coefficient (ICC) of 0.90 was assumed based on prior reliability studies of PO measurements. Using a minimum acceptable ICC, significance level, and power of 0.75, 0.05, and 0.80, respectively, the required sample size for a two-way random-effects ICC model was estimated to be at least 24 patients evaluated by each rater [14]. Thus, radiographic images from 35 patients meeting the predefined inclusion criteria were extracted and analyzed, exceeding this threshold and providing adequate statistical precision. Baseline demographic and clinical characteristics of the study cohort are summarized in Table 1.

## **Radiographic Measurements**

All radiographs were acquired according to standardized institutional clinical protocols, with efforts to maintain neutral pelvic alignment and symmetric weight-bearing, as tolerated. In patients unable to maintain an independent seated position, the radiology technician provided support during radiographic imaging. PO was measured by a team of three board-certified spine surgeons (D.U., H.N., and Y.T.). All evaluators were blinded to the clinical data and each other's measurements. Each surgeon independently measured PO at three postoperative time points for each patient using a digital workstation for Digital Imaging and Communications in Medicine format. The established methods were applied, namely the Osebold and Maloney methods (Figure 1), as described above.

## **Outcomes Measures**

The primary outcome was intersession variability in PO across the three postoperative time points. This was quantified as the difference between the maximum and minimum PO values ( $\Delta$ PO) for each patient–rater-measurement method combination. Secondary outcomes included the clinical success rate (proportion of measurements with  $\Delta$ PO < 5°, a threshold selected based on prior studies demonstrating that the intrinsic measurement error of spinal angular parameters such as the Cobb angle is approximately 5° [15]), the intersession test-retest reliability of each measurement method, and agreement between the two methods for absolute PO measurements.

## **Statistical Analysis**

Intersession PO variability ( $\Delta$ PO values) for the Osebold and Maloney methods was compared within each patient–rater combination using a paired Wilcoxon signed-rank test. Values were expressed as median [interquartile range], and between-method differences were evaluated using the Hodges–Lehmann median difference with corresponding 95% confidence intervals. Clinical success rates were compared between methods using McNemar’s test.

Intersession test–retest reliability was assessed using a two-way random-effects, absolute-agreement ICC, with subjects defined as patient–rater units. Interpretation of ICC values followed established criteria, in which ICC < 0.50, 0.50–0.75, 0.75–0.90, and > 0.90 is considered poor, moderate, good, and excellent reliability, respectively [16]. The standard error of measurement (SEM) was calculated as the square root of the mean square error obtained from the corresponding analysis of variance model, and the minimal detectable change at the 95% confidence level ( $MDC_{95}$ ) was determined as  $1.96 \times \sqrt{2} \times SEM$  [17]. Moreover, interobserver reliability was evaluated using ICC after pooling all time points across raters. Agreement between the Maloney and Osebold methods was further assessed using Bland–Altman analysis, reporting the mean difference and 95% limits of agreement. Proportional bias was examined by regressing the between-method difference (Maloney – Osebold) on the average PO for each measurement. A p-value < 0.05 was considered statistically significant. All statistical analyses were performed using IBM SPSS Statistics for Windows, version 29 (IBM Japan Business Services Co., Tokyo, Japan).

## Results

### Measurement Variability and Clinical Acceptability

Three independent raters measured PO in 35 patients using both the Maloney and Osebold methods, resulting in 105 measurements per method. Median [IQR]  $\Delta$ PO was 8.0° [4.7–11.0] with the Osebold method and 2.5° [1.4–3.7] with the Maloney method. The median between-method difference was 5.1° (Hodges–Lehmann estimate, 95% CI, 4.4–6.0,  $p < 0.001$ , Wilcoxon signed-rank test). Using the predefined  $\Delta$ PO threshold of <5°, the clinical success rate was significantly higher with the Maloney method (91.4%, 96/105) than with the Osebold method (32.4%, 34/105) ( $p < 0.001$ , McNemar test) (Table 2).

### Reliability of Pelvic Obliquity Measurements

Because the number of radiographs included exceeded the minimum required by the power analysis, reliability measurements were performed. Both methods showed high reproducibility across sessions and raters. Intersession test-retest reliability was excellent for the Maloney method (ICC = 0.991; 95% CI, 0.987–0.993), whereas the Osebold method demonstrated lower, but still excellent, reliability (ICC = 0.885; 95% CI, 0.845–0.916). Corresponding SEM values were 1.8° and 4.8°, yielding  $MDC_{95}$  values of 4.9° and 13.3°, respectively (Table 3). When all raters and time points were pooled, interobserver reliability remained excellent for both methods, with ICC values of 0.945 (95% CI, 0.923–0.961) and 0.960 (95% CI, 0.945–0.971) for the Maloney and Osebold methods, respectively. Therefore, PO

measurements were highly consistent, with the Maloney method exhibiting particularly low measurement error and greater sensitivity to change.

### **Between-method systematic difference**

Across all measurements ( $n=315$ ), the Maloney method yielded consistently larger PO values than did the Osebold method. The median difference (Maloney – Osebold) was  $3.3^\circ$  (Hodges–Lehmann:  $3.6^\circ$ , 95% CI  $2.9^\circ$ – $4.3^\circ$ ,  $p<0.001$  by Wilcoxon test). Bland–Altman analysis confirmed a systematic bias, with a mean difference of  $3.9^\circ$  and 95% limits of agreement of  $-9.5^\circ$  to  $17.3^\circ$  (Figure 2). Linear regression further demonstrated a significant proportional bias, with between-method differences increasing with larger PO ( $\beta = 0.20$ ,  $p < 0.001$ ). Collectively, these analyses indicated that the Maloney method consistently measured larger PO than did the Osebold method.

## **Discussion**

This study demonstrated that the Maloney method had smaller intersession variability, higher intersession test-retest reliability, and a substantially lower  $MDC_{95}$  than did the Osebold method, suggesting greater robustness to posture-related variation. Additionally, method-comparison analyses revealed a consistent directional bias, with the Osebold method systematically yielding smaller PO values than the Maloney method. The wide between-method dispersion indicates that measurements cannot be reliably substituted; therefore, the two methods should not be considered interchangeable for surgical decision-making.

Accurate PO assessment is central to surgical decision-making in non-ambulatory patients with NMS, particularly for selecting the distal fusion level, determining the need for pelvic fixation, and evaluating postoperative alignment [4, 5, 18]. In clinical practice, however, standardized seated radiographs are often difficult to obtain in this population because subtle between-session variations in sitting balance and trunk support are common and may directly influence pelvic alignment [19]. Under these conditions, even small changes in seated alignment can alter the orientation of the superior iliac crest line, resulting in angular shifts with the Osebold method, which relies on a fixed horizontal reference. Conversely, the Maloney method incorporates both the iliac crest line and the T1–S1 axis, which tend to shift in the same direction with changes in trunk posture, thereby offsetting angular deviation. This mechanistic difference was reflected in the reproducibility analyses: the Maloney method demonstrated markedly smaller  $\Delta PO$  and higher intersession test–retest reliability than the Osebold method. Thus, the Maloney method may be more resilient to posture-related perturbations during seated radiography and, therefore, may provide a more stable basis for repeated PO assessment.

The large  $MDC_{95}$  observed with the Osebold method has implications for clinical decision-making because it limits the ability to detect clinically meaningful change. Previous reports have often used a threshold of approximately  $10^\circ$  to define clinically relevant PO progression [5, 20]. Therefore, an  $MDC_{95}$  of  $13^\circ$  suggests limited sensitivity of the Osebold method for detecting changes commonly regarded as clinically meaningful. Further, this threshold substantially exceeds the variability generally reported for

Cobb angle assessment, for which meaningful change is typically defined as greater than 5° [15, 21]. Conversely, the smaller MDC<sub>95</sub> and greater posture-resilience of the Maloney method provide a more reliable basis for detecting clinically relevant change over time. This interpretation is supported by our clinical success analysis, in which 91% of Maloney measurements, compared with only 32.4% of Osebold measurements, fell within the clinically acceptable threshold of  $\Delta PO < 5^\circ$ , indicating greater stability of the Maloney method for postoperative monitoring.

In the present study, a consistent directional bias was observed, with the Osebold method systematically yielding smaller PO values than the Maloney method. This finding is consistent with those of a previous study demonstrating that the Osebold method tended to produce smaller PO values compared to the Maloney method in coronally unbalanced spines [13]. Notably, this bias likely reflects inherent differences between the two methods. If a patient assumed a seated posture in which the T1–S1 axis was perfectly perpendicular to the horizontal line, the PO measured using the Osebold method would theoretically coincide with that obtained using the Maloney method. In routine clinical settings, nonetheless, patients tend to adopt a seated posture that balances the spine and pelvis by inclining the spinal axis, thereby partially compensating for pelvic tilt. Consequently, PO values derived from the Osebold method would be expected to be smaller than those obtained using the Maloney method. Additionally, this mechanism plausibly explains the finding that the measurement discrepancy between the Maloney and Osebold methods increased with the magnitude of PO. Accordingly, the methods are not interchangeable, and the systematic underestimation associated with the Osebold method suggests that the Maloney method may be more suitable for detecting subtle postoperative changes. However, each method appears to capture distinct aspects of pelvic alignment. The Osebold method may be more intuitive for reflecting sitting balance and functional posture. Therefore, clarifying the complementary roles of these two approaches and establishing appropriate indications for their use in clinical practice remain important areas for future investigation.

This study has some limitations. First, this analysis was limited to postoperative patients with radiographically stable alignment, and it remains uncertain whether similar posture-related effects would be observed preoperatively or in patients undergoing different fusion strategies. Second, although all included spines were considered radiographically stable, NMS may progress even after skeletal maturity. Therefore, subtle structural changes during follow-up cannot be fully excluded and may have partly contributed to the observed intersession variability [22].

## Conclusion

In non-ambulatory patients with NMS, the Maloney method demonstrated greater stability across repeated seated radiographs, with smaller between-session variability and lower measurement error, than did the Osebold method. Conversely, the Osebold method was more susceptible to systematic, posture-dependent PO underestimation. Collectively, these findings support the Maloney method as a more reliable basis for postoperative PO monitoring. Additionally, our results suggest that PO

measurements obtained using the Osebold method exhibit limited reproducibility and warrant cautious interpretation when used to assess PO.

## Declarations

## Author Contribution

All authors contributed to the conception and design of this study. Material preparation, data collection, and analyses were performed by D.U., Y.T., T.S., M.I., H.N., H.N. and S.K. The first draft of the manuscript was written by D.U., and all authors commented on previous versions of the manuscript. Y.O., S.T., and K.U. supervised the study. All authors have read and approved the final version of the manuscript.

## Acknowledgement

We would like to thank Editage ([www.editage.jp](http://www.editage.jp)) for English-language editing.

## Data Availability

Because this study targeted a rare disease, the dataset used in the present study is not publicly available due to concerns regarding the protection of personal information. Data are available from the corresponding author upon reasonable request.

## References

1. Cloake T, Gardner A (2016) The management of scoliosis in children with cerebral palsy: a review. *J Spine Surg* 2(4):29–309. <https://doi.org/10.21037/jss.2016.09.05>
2. Sewell MD, Wallace C, Malagelada F, Gibson A, Noordeen H, Tucker S et al (2015) Does spinal fusion and scoliosis correction improve activity and participation for children with GMFCS level 4 and 5 cerebral palsy? *Med (Baltim)* 94(49):e1907. <https://doi.org/10.1097/MD.0000000000001907>
3. Kalen V, Conklin MM, Sherman FC (1992) Untreated scoliosis in severe cerebral palsy. *J Pediatr Orthop* 12(3):337–340. <https://doi.org/10.1097/01241398-199205000-00010>
4. Kim HS, Kwon JW, Park KB (2022) Clinical issues in indication, correction, and outcomes of the surgery for neuromuscular scoliosis: narrative review in pedicle screw era. *Neurospine* 19(1):177–187. <https://doi.org/10.14245/ns.2143246.623>
5. Tøndevold N, Lastikka M, Andersen T, Gehrchen M, Helenius I (2020) Should instrumented spinal fusion in nonambulatory children with neuromuscular scoliosis be extended to L5 or the pelvis? *Bone Joint J* 102–B 2261–267. <https://doi.org/10.1302/0301-620X.102B2.BJJ-2019-0772.R2>
6. Modi HN, Suh SW, Song HR, Yang JH, Jajodia N (2010) Evaluation of pelvic fixation in neuromuscular scoliosis: a retrospective study in 55 patients. *Int Orthop* 34(1):89–96.

<https://doi.org/10.1007/s00264-008-0703-z>

7. Saito W, Inoue G, Shirasawa E, Imura T, Nakazawa T, Miyagi M et al (2021) Limitations of posterior spinal fusion to L5 for flaccid neuromuscular scoliosis focusing on pelvic obliquity. *Spine Deform* 9(2):559–565. <https://doi.org/10.1007/s43390-020-00214-1>
8. Shrader MW, Andrisevic EM, Belthur MV, White GR, Boan C, Wood W (2018) Inter- and intraobserver reliability of pelvic obliquity measurement methods in patients with cerebral palsy. *Spine Deform* 6(3):257–262. <https://doi.org/10.1016/j.jspd.2017.10.001>
9. Gupta MC, Wijesekera S, Sossan A, Martin L, Vogel LC, Boakes JL et al (2007) Reliability of radiographic parameters in neuromuscular scoliosis. *Spine* 32(6):691–695. <https://doi.org/10.1097/01.brs.0000257524.23074.ed>
10. Rouissi J, Arvieu R, Dubory A, Vergari C, Bachy M, Vialle R (2017) Intra and inter-observer reliability of determining degree of pelvic obliquity in neuromuscular scoliosis using the EOS-CHAIR® protocol. *Childs Nerv Syst* 33(2):337–341. <https://doi.org/10.1007/s00381-016-3326-5>
11. Osebold WR, Mayfield JK, Winter RB, Moe JH (1982) Surgical treatment of paralytic scoliosis associated with myelomeningocele. *J Bone Joint Surg Am* 64(6):841–856. <https://doi.org/10.2106/00004623-198264060-00006>
12. Maloney WJ, Rinsky LA, Gamble JG (1990) Simultaneous correction of pelvic obliquity, frontal plane, and sagittal plane deformities in neuromuscular scoliosis using a unit rod with segmental sublaminar wires: a preliminary report. *J Pediatr Orthop* 10(6):742–749. <https://doi.org/10.1097/01241398-199011000-00007>
13. Karkenny AJ, Magee LC, Landrum MR, Anari JB, Spiegel D, Baldwin K (2021) The variability of pelvic obliquity measurements in patients with neuromuscular scoliosis. *JB JS Open Access* 6(1). <https://doi.org/10.2106/JBJS.OA.20.00143>. e20.00143
14. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. *Stat Med* 17(1):101–110. [https://doi.org/10.1002/\(sici\)1097-0258\(19980115\)17:1%3C101::aid-sim727%3E3.0.co;2-e](https://doi.org/10.1002/(sici)1097-0258(19980115)17:1%3C101::aid-sim727%3E3.0.co;2-e)
15. Morrissy RT, Goldsmith GS, Hall EC, Kehl D, Cowie GH (1990) Measurement of the Cobb angle on radiographs of patients who have scoliosis: evaluation of intrinsic error. *J Bone Joint Surg Am* 72(3):320–327. <https://doi.org/10.2106/00004623-199072030-00002>
16. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
17. de Vet HCW, Terwee CB, Knol DL, Bouter LM (2006) When to use agreement versus reliability measures. *J Clin Epidemiol* 59(10):1033–1039. <https://doi.org/10.1016/j.jclinepi.2005.10.015>
18. Soini V, Karkkola S, Raitio A, Syvänen J, Helenius I (2025) Pelvic vs lumbar fusion in neuromuscular scoliosis—a systematic review. *Glob Spine J* 15(7):3395–3404. <https://doi.org/10.1177/21925682251328620>
19. Loughenbury PR, Tsirikos AI (2022) Current concepts in the treatment of neuromuscular scoliosis: clinical assessment, treatment options, and surgical outcomes. *Bone Jt Open* 3(1):85–92.

<https://doi.org/10.1302/2633-1462.31.BJO-2021-0178.R1>

20. Sussman MD, Little DM, Alley RM, McCoig JA (1996) Posterior instrumentation and fusion of the thoracolumbar spine for treatment of neuromuscular scoliosis. *J Pediatr Orthop* 16(3):304–313. <https://doi.org/10.1097/00004694-199605000-00004>
21. Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. *J Bone Joint Surg Am* 72(3):328–333. <https://doi.org/10.2106/00004623-199072030-00003>
22. Thometz JG, Simon SR (1988) Progression of scoliosis after skeletal maturity in institutionalized adults who have cerebral palsy. *J Bone Joint Surg Am* 70(9):1290–1296. <https://doi.org/10.2106/00004623-198870090-00002>

## Tables

Table 1. Demographic and Clinical Characteristics of the Study Cohort

| Characteristic                                            | n=35                   |
|-----------------------------------------------------------|------------------------|
| Sex                                                       |                        |
| Male                                                      | 20                     |
| Female                                                    | 15                     |
| Age at surgery, years; mean [SD]                          | 15.2 [2.2]             |
| Preoperative Cobb angle of main curve, degrees; mean [SD] | 109.2 [22.9]           |
| Upper Instrumented Vertebrae                              | T1:4<br>T2:17<br>T3:14 |

SD, standard deviation.

Table 2. Intersession Variability of Pelvic Obliquity Measurements Using the Osebold and Maloney Method

| Parameter                              | Osebold        | Maloney       | Between-Method Comparison <sup>1</sup> | p-value <sup>2</sup> |
|----------------------------------------|----------------|---------------|----------------------------------------|----------------------|
| ΔPO, mean ± SD (°)                     | 8.1 ± 4.4      | 2.8 ± 1.9     | ☒                                      | <0.001               |
| ΔPO, median [IQR] (°)                  | 8.0 [4.7–11.0] | 2.5 [1.4–3.7] | 5.1 (95% CI 4.4 to 6.0)                | <0.001               |
| Clinical success rate <sup>3</sup> (%) | 32.4           | 91.4          | ☒                                      | <0.001               |

1. Hodges–Lehmann median difference (Maloney – Osebold).
2. Wilcoxon signed-rank test for within-patient comparison.
3.  $\Delta PO < 5^\circ$  defined as radiographic stability

CI, confidence interval; IQR, interquartile

Table 3. Intersession Test–Retest Reliability of Pelvic Obliquity Measurements Using the Osebold and Maloney Methods

| Parameter                                              | Osebold             | Maloney             |
|--------------------------------------------------------|---------------------|---------------------|
| Intersession test–retest ICC(A,1)[95% CI] <sup>1</sup> | 0.885 [0.845–0.916] | 0.991 [0.987–0.993] |
| SEM(°) <sup>2</sup>                                    | 4.78                | 1.77                |
| MDC <sub>95</sub> (°) <sup>3</sup>                     | 13.25               | 4.91                |
| Interobserver ICC(A,1)                                 | 0.960 [0.945–0.971] | 0.945 [0.923–0.961] |

1. ICC(A,1): Two-way random-effects model, absolute agreement; assesses test–retest reliability across sessions for each method.
2. SEM = standard error of measurement, calculated as the square root of the mean square error from the ANOVA model.
3. MDC<sub>95</sub> = minimal detectable change at the 95% confidence level, calculated as  $1.96 \times \sqrt{2} \times \text{SEM}$ .

ICC, intraclass correlation coefficient.

## Figures

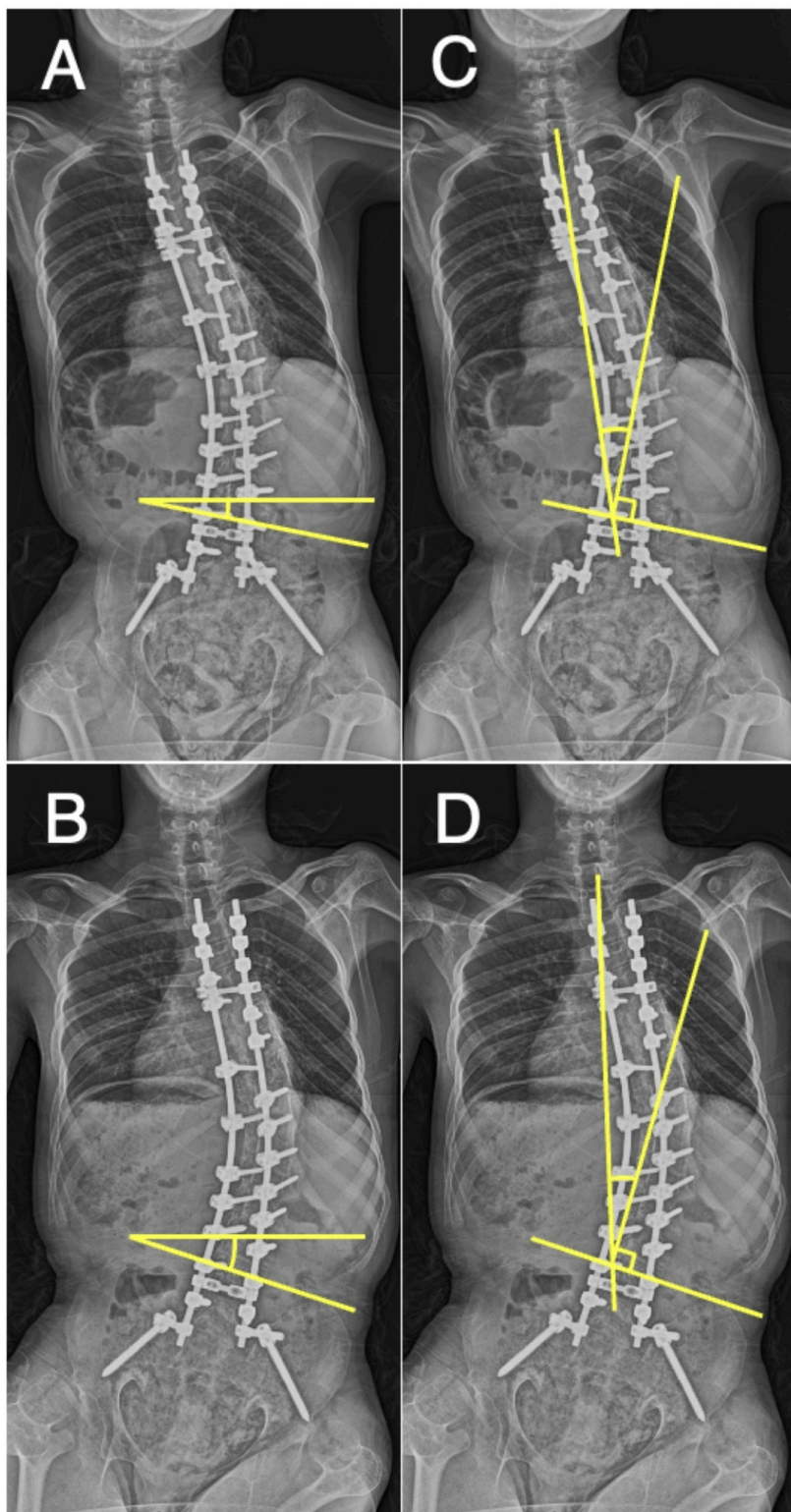


**Figure 1**

Representative seated anteroposterior radiographs illustrating intersession pelvic obliquity (PO) measurements in a nonambulatory patient with neuromuscular scoliosis using the Osebold method

(A, B) and the Maloney method (C, D). (A) and (C) depict the same radiograph, whereas (B) and (D) depict the same radiograph obtained at a different session, illustrating intersession test-retest reliability for

each measurement method.

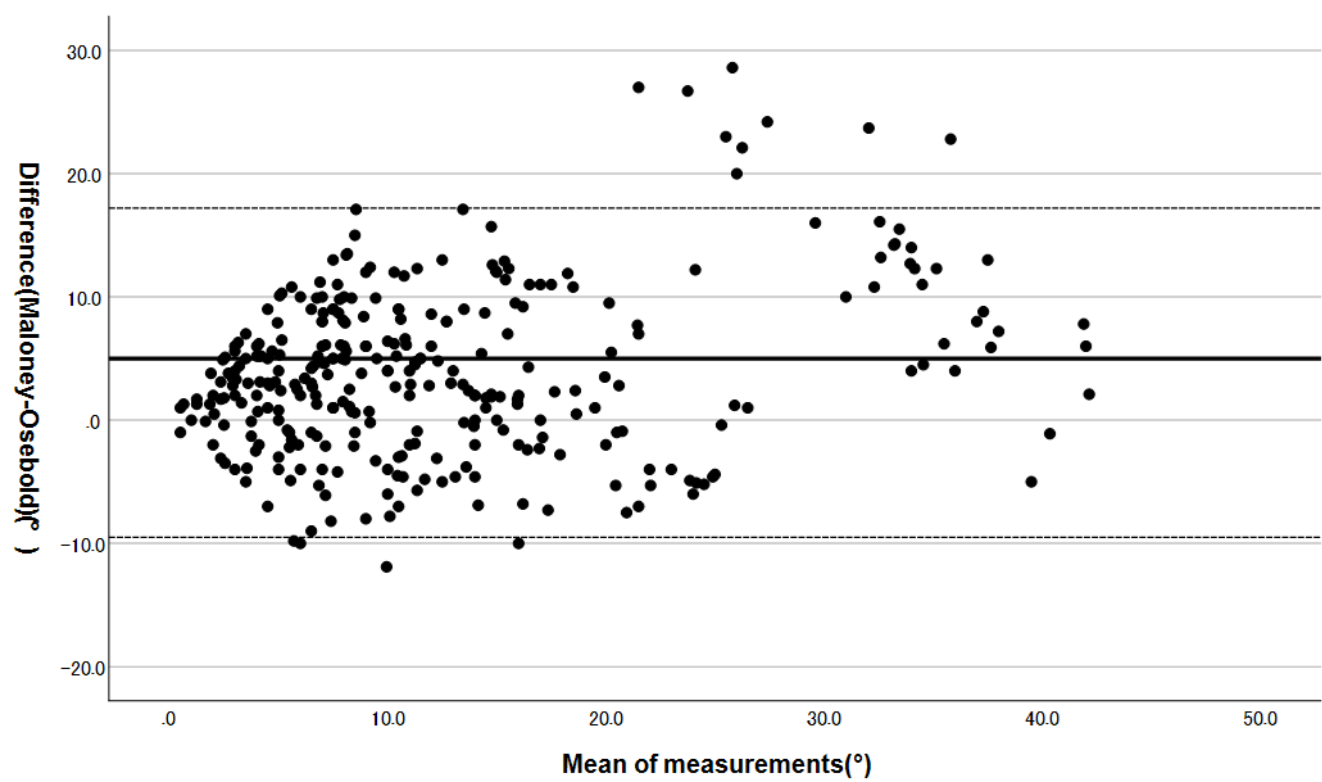


**Figure 2**

Bland–Altman plot illustrating agreement between the Maloney and Osebold methods

The solid line represents the mean difference (3.9°), and the dashed lines indicate the 95% limits of agreement (−9.5° to 17.3°). A significant proportional bias was observed ( $\beta = 0.20$ ,  $p < 0.001$ ).