

Clinical Study

# Prediction of mechanical complications in adult spinal deformity surgery—the GAP score versus the Schwab classification

Eva Jacobs, MD<sup>a</sup>, Barend J. van Royen, MD, PhD<sup>b</sup>,  
Sander M.J. van Kuijk, MSc, PhD<sup>c</sup>, Johannes M.R. Merk, MSc<sup>d</sup>,  
Agnita Stadhouders, MD<sup>b</sup>, Lodewijk W. van Rhijn, MD, PhD<sup>a</sup>,  
Paul C. Willems, MD, PhD<sup>a,\*</sup>

<sup>a</sup> Department of Orthopedic Surgery, Maastricht University Medical Center, Maastricht, the Netherlands

<sup>b</sup> Department of Orthopedic Surgery, Amsterdam University Medical Center, Amsterdam, the Netherlands

<sup>c</sup> Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, Maastricht, the Netherlands

<sup>d</sup> Department of Finance, Maastricht University, Maastricht, the Netherlands

Received 29 August 2018; revised 27 November 2018; accepted 28 November 2018

## Abstract

**BACKGROUND CONTEXT:** Surgery for adult spinal deformity is a challenging and complex procedure with high reported complication (8.4%–42%) and revision rates (9%–17.6%). Failure to achieve or maintain adequate postoperative sagittal alignment has been reported to be the main cause of mechanical complications. In order to define appropriate surgical targets, the Scoliosis Research Society-Schwab classification and the Global Alignment and Proportion (GAP) score were established. In the literature, no study has yet compared these classification systems with respect to the risk of developing mechanical complications.

**PURPOSE:** To assess and compare the ability of the Schwab classification and the GAP score to predict mechanical complications following adult spinal deformity surgery.

**STUDY DESIGN:** Two-center, retrospective cohort study.

**PATIENT SAMPLE:** Thirty-nine patients suffering adult spinal deformity who underwent long segment spinal fusion ( $\geq 4$  levels), minimum follow-up of 2 years.

**OUTCOME MEASURES:** The ability of the Schwab classification and GAP score to predict mechanical failure was determined by computing the Area Under the receiver operating characteristic curve.

**METHODS:** Full-spine pre- and postoperative radiographs of all patients were analyzed for mechanical complications. Subsequently, the pre- and postoperative Schwab and GAP score were determined. Logistic regression analysis was used to assess the ability of both systems to determine which was the most appropriate for the prediction of mechanical failure. Correlations between the various factors constituting the GAP score and Schwab classification were estimated using the Spearman rank order correlation coefficient.

**RESULTS:** The results demonstrated that both classification systems are capable of predicting radiographic evidence of mechanical failure; however, the GAP score proved to be significantly better ( $p=.003$ ). The relative pelvic version of the GAP score serves a similar role as the pelvic tilt modifier from the Schwab classification ( $\rho=-0.84$ ,  $p<.01$ ). The relative lumbar lordosis from the GAP score functions much like the PI-LL modifier from the Schwab classification ( $\rho=-0.94$ ,  $p<.01$ ). The GAP score is most significantly dependent on relative spinopelvic alignment, relative lumbar lordosis, and

FDA device/drug status: Not applicable.

Author disclosures: *EJ*: Nothing to disclose. *BJVR*: Nothing to disclose. *SMJVK*: Nothing to disclose. *JMRM*: Nothing to disclose. *AS*: Nothing to disclose. *LWVR*: Nothing to disclose. *PCW*: Nothing to disclose.

\* Corresponding author. Maastricht University Medical Center, Department of Orthopedic Surgery, Research School CAPHRI, P. Debyelaan 25, 6229 HX, Maastricht, the Netherlands. Tel.: +31-43-3884152; fax: +31-43-3874893.

E-mail address: [p.willems@mumc.nl](mailto:p.willems@mumc.nl) (P.C. Willems).

relative pelvic version ( $\rho=0.85$ ,  $\rho=0.84$ , and  $\rho=0.84$ , respectively,  $p<.01$ ). Correlation with the lordosis distribution index was also significant but was not as strong ( $\rho=0.65$ ,  $p<.01$ ). Age, on the contrary, showed poor correlation with the GAP score ( $\rho=0.17$ ,  $p=.300$ ).

**CONCLUSIONS:** Both the Schwab classification and the GAP score are capable of predicting mechanical complications. The GAP score proved to be significantly more appropriate. This difference is probably attributed to the fact that in the GAP score all parameters are related to the patient's individual pelvic incidence. © 2018 Elsevier Inc. All rights reserved.

**Keywords:** Adult spinal deformity; Deformity surgery; GAP score; Global Alignment and Proportion Score; Sagittal alignment; Schwab classification

## Introduction

Degenerative changes have the potential to greatly disrupt the normal curvature of the spine, leading to sagittal malalignment [1]. The subsequent pain and decline in functional status constitute a concerning clinical picture [2]. As a response to positive sagittal (mal)alignment, the body progressively recruits compensatory mechanisms to counteract the anterior truncal shift [1]. A chain of compensation is initiated from the flexible parts of the spine, and extends to the hips, lower extremities, and cervical spine to preserve an erect posture and horizontal gaze. The interaction between deformity and compensatory mechanisms depicts the final presentation of patients with adult spinal deformity (ASD) [3]. In cases of severe deformity, surgical intervention has been shown to offer superior clinical and radiographic outcomes compared with nonoperative approaches [4–6].

Understanding of the sagittal alignment of the spine allows for comprehensive surgical planning which is key to achieving optimum postoperative alignment and improved outcomes [3]. Recent studies on outcomes following ASD surgeries have shown high rates of complications (8.4%–42%) [7,8] and revision rates (9%–17.6%) [3,9]. Increased junctional stress concentration causes soft-tissue and ligamentous failure, vertebral fracture, and bone implant interface failure. Proximal junctional kyphosis and proximal junctional failure (PJF) are recognized complications of ASD surgery with a reported prevalence of 20%–40% [3,10,11]. The development of junctional failure is frequently symptomatic and can lead to the need for revision surgery and can have potential catastrophic neurologic sequelae [12]. The major risk factors for junctional failure include older age (over 55 years of age), large abnormal preoperative sagittal parameters, osteoporosis, high body mass index, thoracoplasty procedures, and fusion to the lower lumbar vertebra and sacrum [11,13–15].

Adequate alignment following surgical treatment for ASD is not consistently achieved. Since revision rates following realignment procedures increase progressively with longer follow-up, it is possible that these revisions are driven by failure to achieve and/or maintain realignment [1,3]. Both under- and over-correction in the sagittal plane have been reported to be a main cause of mechanical

complications [12,16–18]. Nowadays, extensive research has converted theoretical concepts into clinically relevant guidelines on sagittal alignment. In treatment, planning for ASD accurate sagittal alignment analysis has become an essential tool. In order to understand appropriate targets for the correction of ASD, the Scoliosis Research Society (SRS)-Schwab classification and the Global Alignment and Proportion (GAP) score have been developed [19,20]. The SRS-Schwab classification uses three sagittal modifiers to quantify deformity [19]. The Schwab modifier thresholds were based on the established correlations between radiographic parameters and health-related quality-of-life measures [19,21]. However, despite achieving optimal Schwab values postoperatively, mechanical complications are not uncommon [22]. Yilgor et al. [20] noticed that it was not clear how the Schwab criteria influence the risk of mechanical complications. Therefore, they developed the GAP score in which thresholds were mainly determined on the risk of mechanical complications. In the GAP score, optimal sagittal alignment is based on four factors deviating from their ideal curves, and these factors are proportionally related to the pelvic incidence (PI) [20]. The overall goal of the GAP score is to achieve a more patient-tailored, or patient-specific, guide for spinopelvic alignment. To date, no study has yet compared the SRS-Schwab classification with the newly developed GAP score with respect to their ability to predict the risk of developing mechanical complications following ASD surgery. Therefore, the purpose of the current study is to assess whether the SRS-Schwab classification or the GAP score is the most appropriate for the prediction of radiographic failure in patients suffering from ASD.

## Materials and methods

The current study is a review of a two-center cohort of patients suffering ASD who underwent posterior spinal fusion and instrumentation. Medical records and radiographs of patients with ASD treated with posterior spinal fusion between 2005 and 2015 at two spinal deformity centers in the Netherlands (Maastricht University Medical Center and Amsterdam University Medical Center) were retrospectively reviewed. Institutional review board approval was obtained before study initiation (METC16-4-029) at each site. We

included adults suffering ASD (defined as a deformity of the spine in either the coronal ( $>10^\circ$  lateral deviation of the vertical axis) or sagittal plane (T1 pelvic angle  $>10^\circ$  or T1 spinopelvic inclination [T1-SPi] angle  $>0^\circ$ ), who received surgical treatment constituting posterior spinal fusion of at least four vertebrae, and of whom standing full-spine plain radiographs pre- and postoperatively were available. Patients were excluded if less than four vertebrae were fused or if there was incomplete radiographic follow-up (less than 2 years of follow-up, or no full-spine radiographs).

All radiographs were analyzed using validated software (Surgimap, Nemaris Inc, New York, NY, USA). Pelvic parameters that were measured were the PI, pelvic tilt (PT), and sacral slope. Regional spinal parameters included PI-LL mismatch, L1-S1 Lumbar Lordosis (L1-S1 LL), L4-S1 Lumbar Lordosis (L4-S1 LL), and Thoracic Kyphosis (TK, T4-T12). Sagittal alignment was assessed linearly by T1-Spi, T1 pelvic angle, global tilt, and global sagittal alignment [18, 23–25].

The GAP score comprises the relative pelvic version (RPV), relative lumbar lordosis (RLL), lordosis distribution index (LDI), relative spinopelvic alignment (RSA), and age [20]. The GAP score can range from 0 to 13 points. The cut-off points for the GAP score that were chosen were in accordance with the cut-off values as determined by Yilgor et al. [20]. Relative pelvic version (measured sacral slope minus ideal sacral slope) of less than  $-15^\circ$  was considered severe retroversion;  $-15^\circ$  to  $-7.1^\circ$ , moderate retroversion;  $-7^\circ$  to  $5^\circ$ , aligned; and more than  $5^\circ$ , anteversion. Relative lumbar lordosis (measured lumbar lordosis minus ideal lumbar lordosis) of less than  $-25^\circ$  was considered severe hypolordosis;  $-25^\circ$  to  $-14.1^\circ$ , moderate hypolordosis;  $-14^\circ$  to  $11^\circ$ , aligned; and more than  $11^\circ$ , hyperlordosis. Lordosis distribution index (L4-S1 lordosis divided by L1-S1 lordosis multiplied by 100) of less than 40% was considered severe hypolordotic maldistribution; 40%–49%, moderate hypolordotic maldistribution; 50%–80%, aligned; and more than 80%, hyperlordotic maldistribution. Relative spinopelvic alignment (measured global tilt minus ideal global tilt) of more than  $18^\circ$  was considered severe positive malalignment;  $10.1^\circ$ – $18^\circ$ , moderate positive malalignment;  $10^\circ$  to  $-7^\circ$ , aligned; and less than  $-7^\circ$ , negative malalignment. A GAP score of 0–2 was categorized as indicating a proportioned spinopelvic state; 3–6, as moderately disproportioned; more than 6, as severely disproportioned.

The original sagittal modifiers of the SRS-Schwab classification are sagittal vertical axis (SVA), PT, and PI-LL [19]. The SVA parameter is a distance and requires a calibrated image. Caused by the fact that in the current study not all images were calibrated we used the T1-SPi for the assessment of global spinal alignment, which strongly correlates with SVA [23]. Patients with a T1-SPi of less than  $1.35^\circ$  were classified with a T1-SPi modifier “0,” a T1-SPi between  $1.35^\circ$  and  $7.2^\circ$  was classified with a T1-SPi modifier “+” and more than  $7.2^\circ$  with a T1-SPi modifier “++.” Patients with a PI-LL value of less than  $10^\circ$  were classified

Table 1

Types and definitions of radiographic and implant-related mechanical complications used [20]

Type of mechanical complication	Definition*
Proximal junctional kyphosis (PJK)	$\geq 10^\circ$ increase in kyphosis between UIV and UIV+2 between early postoperative and follow-up radiographs
Proximal junctional failure (PJF)	Fracture of UIV or UIV+1, pullout of instrumentation at UIV, and/or sagittal subluxation
Distal junctional kyphosis/failure	$\geq 10^\circ$ postoperative increase in kyphosis angle between LIV and LIV-1 and/or pullout of instrumentation at LIV
Rod breakage	Single or double rod breakage
Implant-related complications	Other radiographic implant-related complications such as screw loosening, breakage, or pullout or interbody graft, hook, or set-screw dislodgement

\* UIV: upper instrumented vertebra, LIV: lowest instrumented vertebra, +2 and +1: 2 and 1 vertebrae above UIV, and -1: 1 vertebra below LIV.

with a PI-LL modifier “0,” a PI-LL value between  $10^\circ$  and  $20^\circ$  was classified with a PI-LL modifier “+” and greater than  $20^\circ$  with a PI-LL modifier “++.” Patients with a PT of less than  $20^\circ$  were classified with a PT modifier “0,” a PT between  $20^\circ$  and  $30^\circ$  was classified with a PT modifier “+” and greater than  $30^\circ$  with a PT modifier “++.” For statistical weight, the following values were defined for the Schwab modifiers: 1 for modifier “0,” 2 for modifier “+,” and 3 for modifier “++.”

The presence or absence of mechanical complications was defined as proximal junctional kyphosis or failure, distal junctional kyphosis or failure, rod breakage, and implant-related complications (Table 1) [20]. Revision surgery caused by mechanical complications was defined as “mechanical revision.” Based on these criteria, three groups were defined: (1) “Normal” group (without mechanical complications), (2) “Non-revised” group (with radiographic mechanical complications, but without clinical indication for revision), (3) “Revised” group (with radiographic mechanical complications, patients underwent revision surgery).

### Statistical analysis

We included all evaluable patients treated within the selected period of time. Baseline characteristics of the patients were described using mean and standard deviation (SD) for continuous variables, and as count and percentage for categorical variables. Patient characteristics in the three groups were compared using analysis of variance for continuous variables, and the chi-square test for categorical variables.

Table 2  
Comparison of baseline demographics between groups

Variable	Normal group (1)	Non-revised group (2)	Revised group (3)	Total	Chi-square or ANOVA (p value for difference)
No.	17 (44%)	12 (30%)	10 (26%)	39	
Age (y)	59.9±9.9	61.4±8.8	59.2±11.1	60.2±9.7	0.01 (.92)
Female gender (no.)	12 (71%)	11 (92%)	6 (60%)	29 (74%)	3.09 (.21)
Prior spine surgery (no.)	1 (6%)	0 (0%)	1 (10%)	2 (5%)	1.16 (.56)
Diagnosis (no.)					4.91 (.56)
Degenerative	13 (76%)	8 (67%)	9 (90%)	30 (77%)	
Idiopathic	1 (6%)	0 (0%)	0 (0%)	1 (3%)	
Post-traumatic	1 (6%)	3 (25%)	1 (10%)	5 (13%)	
Congenital	2 (12%)	1 (8%)	0 (0%)	3 (8%)	
Mean no. of vertebrae fused	7.7±2.8	7.7±3.7	7.3±2.0	7.6±2.9	0.11 (.74)

Analysis of variance was used to determine if differences existed in radiographic parameters between the three groups and to compare the pre- and postoperative outcomes. We used logistic regression analysis to assess whether the GAP score or the Schwab classification was most appropriate for prediction of mechanical failure after dichotomization of the outcome. The ability of both the GAP score and the Schwab classification to discriminate between those who experienced mechanical failure, and those who did not was subsequently determined by computing the area under the receiver operating characteristic curve, or area under the curve (AUC), resulting from the logistic regression analysis. For this analysis, the “Nonfailure” group was compared with the “Mechanical complication” group (comprising the “Non-revised” group and “Revised” group). In order to compare the receiver operating characteristic curves of the GAP score and the Schwab classification, a nonparametric approach by deLong et al. [26] was used.

Correlations between the various factors constituting the GAP score and Schwab classification were estimated using the Spearman rank order correlation coefficient.

All analyses were performed using R version 3.5.1.

## Results

A total of 39 ASD patients met the inclusion criteria and were included for analysis. The study population consisted of 10 men and 29 women, the average age was 60.18±9.67 years (range, 39–82 years; Table 2). The “Normal” group was formed by 17 (44%) patients, and radiographic mechanical complications were measured in 22 (56%) patients (12 [30%] patients in the “Non-revised” group, 10 [26%] patients in the “Revision” group). No statistically significant differences were found between the three groups for age, gender, prior spine surgery, diagnosis, or mean number of vertebrae fused (Table 2).

Preoperatively, there was no significant difference in GAP score and Schwab classification between the three groups ( $p>.05$  for all comparisons; Fig. 1A and B). For the “Normal” group, the average GAP score and Schwab values declined postoperatively ( $p=.045$  and  $p=.663$ , respectively), whereas for the “Mechanical complication” groups, these increased (Fig. 1A and B;  $p>.05$  for all comparisons).

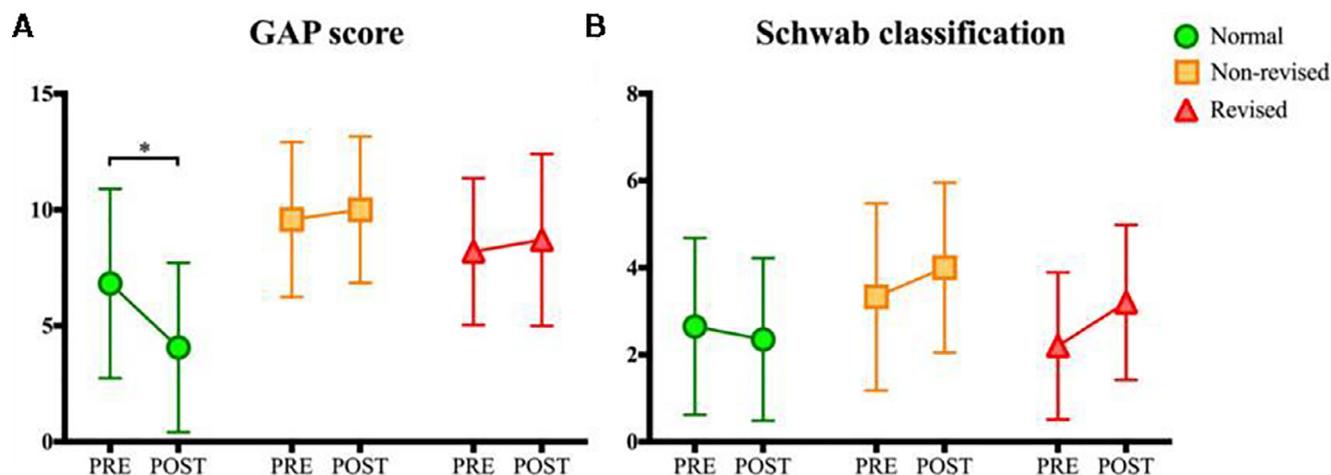


Fig. 1. A. The mean pre- and postoperative GAP scores per group (mean ± standard deviation, \* indicates  $p<0.05$ ). B. The mean pre- and postoperative values for the Schwab classification per group (mean ± standard deviation, \* indicates  $p<0.05$ ).

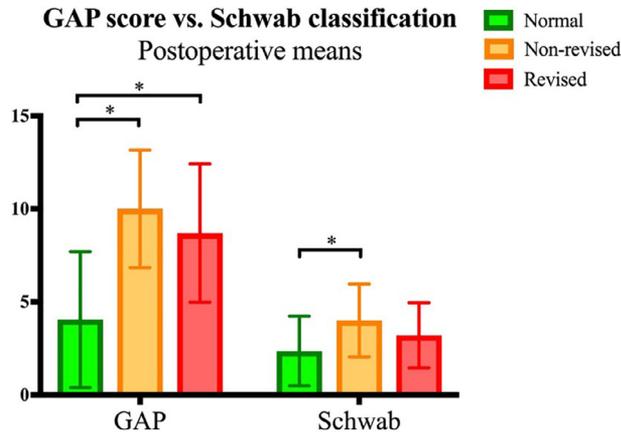


Fig. 2. Postoperative mean values for the GAP score and Schwab classification per group (mean ± standard deviation, \*indicates p<0.05).

Postoperatively, the “Normal” group demonstrated significantly lower mean GAP scores in comparison to the “Non-revised” group and the “Revised” group (p<.001 and p=.005, respectively; Fig. 2). The mean postoperative Schwab value for the “Normal” group was also significantly lower in comparison to the “Non-revised” group (p=.032), whereas the difference between the “Normal” group and the “Revised” group was not significant (p=.250; Fig. 2).

Logistic regression analysis for the prediction of mechanical failure revealed a significant association with both the GAP score (odds ratio [OR] 1.45, 95% confidence interval [CI]: 1.16–1.79, p=.001) and the Schwab classification (OR 1.44, 95% CI: 1.01–2.08, p=.046). The ability of both the GAP score and the Schwab classification to predict mechanical complications was expressed as the AUC. The AUC for the GAP score was 0.86 (95% CI: 0.75–0.97), whereas the AUC for the Schwab classification was 0.69 (95% CI: 0.52–0.86). The difference between AUC’s of the Schwab and the GAP score was statistically significant (p=.003), indicating that the GAP discriminates better between those patients who will develop mechanical failure and those who will not. Subsequently, a plot of the predicted probability for the GAP score is visualized in Fig. 3. For example, the risk of radiographic mechanical complications for patients with a postoperative GAP score

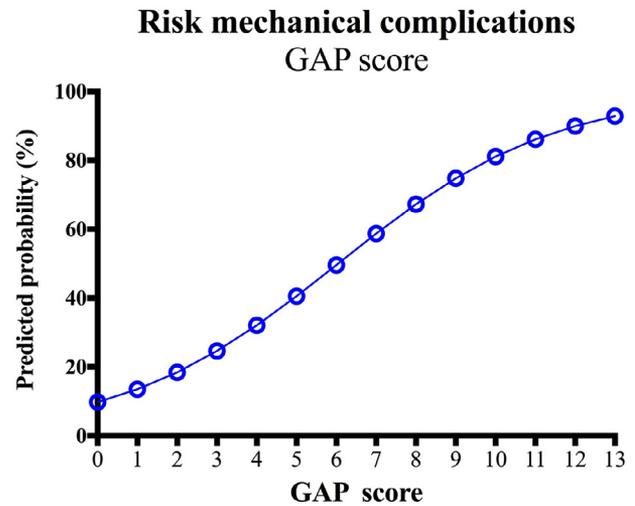


Fig. 3. Predicted probability curve for the GAP score.

of 2 is approximately 18%, whereas the risk is approximately 90% with a GAP score of 12.

Spearman rank order correlation analysis showed that the GAP score was most strongly correlated with RSA, RLL, and RPV ( $\rho=0.85$ ,  $\rho=0.84$ , and  $\rho=0.84$ , respectively, p<.01; Table 3). The correlation with the LDI was also significant but was as strong ( $\rho=0.65$ , p<.01). Age, on the contrary, showed poor correlation with the GAP score ( $\rho=.17$ , p=.300). For the Schwab classification, the Spearman rank order correlation analysis showed a significant correlation with all three sagittal modifiers (PI-LL  $\rho=0.90$ , Global alignment (T1-SPi)  $\rho=0.67$  and PT  $\rho=0.70$ , p<.01), of which PI-LL was the strongest.

The RPV of the GAP score serves a similar role as the PT modifier from the Schwab classification, demonstrating a significantly negative correlation coefficient ( $\rho=-0.84$ , p<.01). In Table 4, the distribution of patients over both subcategories (RPV and PT) is visualized, showing a few differences between the classification systems (eg, more patients are aligned according to the GAP score than according to the Schwab modifiers). In Fig. 4A, the sacral slope is plotted against the PI with respect to the “ideal sacral slope” [20]. The geometrical relationship between the pelvic angles and sacral slope should be kept in mind (PI equals the sum of the sacral slope and PT). Patients

Table 3

Spearman rank order correlation coefficient analysis for the various determinants of the GAP score and the Schwab classification. \* indicate p<.01, \*\* indicate p<.05

	Pelvic incidence minus lumbar lordosis (PI-LL)	Global alignment (T1-SPi)	Pelvic tilt (PT)
Relative pelvic version (RPV)	-0.69*	-0.02	-0.84*
Relative lumbar lordosis (RLL)	-0.94*	-0.54*	-0.58*
Lordosis distribution index (LDI)	0.04	0.12	-0.32**
Relative sagittal alignment (RSA)	0.85*	0.55*	0.67*

Table 4  
Distribution of patients over the various subcategories of the GAP score versus the Schwab classification (no. [%])

GAP	Normal	Non-revised	Revised	SCHWAB	Normal	Non-revised	Revised
<b>Relative pelvic version</b>				<b>Pelvic tilt</b>			
Anteversión	1 (6)	0 (0)	0 (0)	0	8 (47)	1 (8)	1 (10)
Aligned	10 (59)	1 (8)	2 (20)	+	5 (29)	5 (42)	7 (70)
Moderate retroversion	2 (12)	2 (17)	3 (30)	++	4 (24)	6 (50)	2 (20)
Severe retroversion	4 (24)	9 (75)	5 (50)				
<b>Relative lumbar lordosis</b>				<b>PI-LL</b>			
Hyperlordotic	0 (0)	0 (0)	0 (0)	0	10 (59)	3 (25)	4 (40)
Aligned	11 (65)	3 (25)	3 (30)	+	4 (24)	1 (8)	3 (30)
Moderate hypolordosis	3 (18)	2 (17)	2 (20)	++	3 (18)	8 (67)	3 (30)
Severe hypolordosis	3 (18)	7 (58)	5 (50)				
<b>Relative sagittal alignment</b>				<b>Alignment</b>			
Negative	0 (0)	0 (0)	0 (0)	0	7 (41)	4 (33)	3 (30)
Aligned	6 (35)	1 (8)	0 (0)	+	3 (18)	2 (17)	2 (20)
Moderately positive	7 (41)	0 (0)	2 (20)	++	7 (41)	6 (50)	5 (50)
Severely positive	4 (24)	11 (92)	8 (80)				
<b>Lordosis distribution index</b>				<b>Total</b>			
Severely hypolordotic	1 (6)	1 (8)	2 (20)	0	11 (65)	3 (25)	4 (40)
Moderate hypolordotic	1 (6)	2 (17)	0 (0)	+	3 (18)	2 (17)	3 (30)
Aligned	13 (76)	2 (17)	4 (40)	++	3 (18)	7 (58)	3 (30)
Hyperlordotic	2 (12)	7 (58)	4 (40)				
<b>Total</b>							
Proportioned	7 (41)	0 (0)	0 (0)	0	11 (65)	3 (25)	4 (40)
Moderately disproportioned	6 (35)	1 (9)	4 (40)	+	3 (18)	2 (17)	3 (30)
Severely disproportioned	4 (24)	11 (92)	6 (60)	++	3 (18)	7 (58)	3 (30)

with a high PI require a high PT (and sacral slope); however, according to the Schwab modifier, these values are classified as “too high” (“+” or “++”). On the other hand, patients with a small PI require an even smaller PT and should be classified as “++” instead of “+.”

The RLL from the GAP score functions much like the PI-LL modifier from the Schwab classification and,

therefore, also showed an extremely low correlation coefficient ( $\rho = -0.94$ ,  $p < .01$ ). The distribution of patients over both subcategories (RLL and PI-LL) is visualized in Table 4 and appeared to be highly comparable. In Fig. 4B, the lumbar lordosis is plotted against the PI with respect to the “ideal lumbar lordosis” [20]. Again, the three groups presented are the three Schwab modifiers. Only for patients

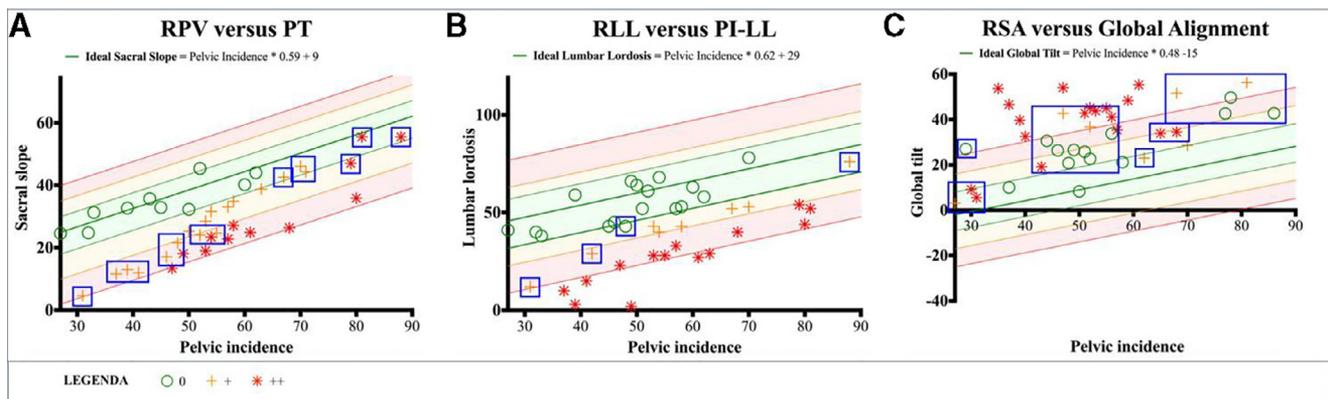


Fig. 4. A. Sacral slope (SS) per patient plotted against his/her pelvic incidence (PI) with respect to “ideal sacral slope” (green line). Green bold line represents alignment, upper orange and red areas represent anteversion, lower orange area represents moderate retroversion and lower red area represents severe retroversion. Symbols are coded according to the PT Schwab modifier (“0,” “+,” or “++,” displayed as “\*”). The blue boxes indicate differences between the GAP score and Schwab classification. Note:  $PI = PT + SS$ . B. C. Lumbar lordosis (LL) per patient plotted against his/her pelvic incidence (PI) with respect to “ideal lumbar lordosis” (green line). Green bold line represents alignment, upper orange and red areas represent hyperlordosis, lower orange area represents moderate hypolordosis and lower red area represents severe hypolordosis. Symbols are coded according to the PI-LL Schwab modifier (“0,” “+,” or “++,” displayed as “\*”). The blue boxes indicate differences between the GAP score and Schwab classification. C. Global tilt per patient plotted against his/her pelvic incidence (PI) with respect to “ideal global tilt” (green line). Green bold line represents alignment, upper orange area represents moderate positive malalignment, upper red area represents severe positive malalignment. Lower orange and red areas represent negative malalignment. Symbols are coded according to the PT Schwab modifier (“0,” “+,” or “++,” displayed as “\*”). The blue boxes indicate differences between the GAP score and Schwab classification.

with PI values near the upper-normal and lower-normal values the PI-LL show limitations. Between a PI of approximately 50–80, the PI-LL modifier is quite similar to the GAP score.

The relative spinal alignment demonstrated a correlation coefficient  $\rho=0.55$  ( $p<.01$ ) with the Global alignment (T1-SPI). The distribution of patients over both subcategories (RSA and Global alignment) is visualized in Table 4. According to the GAP score, approximately all patients show positive malalignment whereas according to the Schwab classification the number of aligned patients is twice as high (this might explain the relatively low, but still significant, correlation coefficient) between RSA (GAP) and Global alignment (Schwab). In Fig. 4C, the global tilt is plotted against the PI with respect to the “ideal global tilt” [20] providing more insight in the differences found. These results indicate that the Schwab Classification might underestimate the malalignment with regard to a patient’s individual PI.

The distribution of patients over the two “novel” subcategories of the GAP score is visualized in Table 4. Hyperlordotic maldistribution appeared to be an important cause for radiographic mechanical complications. Age did not show to be a contributing factor to ability of the GAP score to predict mechanical complications.

## Discussion

The current study aimed to assess which classification system, the SRS-Schwab classification or GAP score, is the most appropriate for the prediction of radiographic mechanical failure in patients operated for ASD. The results demonstrated that both classification systems can predict radiographic mechanical failure, but the GAP score proved to be better with an AUC of 0.86 compared with 0.69 ( $p=.003$ ).

The outcome of the current study might explain why it is not uncommon to observe complications even after ideal correction of all sagittal modifiers from the Schwab classification [20]. A multicenter study by Soroceano et al. [22] found a prevalence of radiographic and implant-related complications of 31.7% when using Schwab’s target values (revision surgery was indicated in 52.6% of these patients). When the surgical goals according to the GAP score are achieved postoperatively, the shape of the spine allows for equal distribution of loads across the spinal column, pelvis, and lower limbs. However, when the surgical goals according to the Schwab classification are achieved, this equilibrium might still be disrupted, which may lead to mechanical failure. There are several differences between the Schwab classification and the GAP score that might account for this [20]. For the Schwab classification, the PI minus lumbar lordosis, PT, and global alignment are independently used as numerical values [19,21]. On the contrary, for the GAP score, all parameters are evaluated in relation to the PI [27]. PI is a morphologic parameter that remains constant during adulthood and is not affected by patient position or spinal surgery [28]. A patient with an upper-limit PI will have a physiologically

high PT and sacral slope. According to the Schwab criteria the surgical goal in PT for these patients is  $<20^\circ$ ; however, this is too low in case of a high PI and may actually increase the risk of mechanical complications instead of preventing them (Fig. 4A: blue boxes on the right). For every patient the Schwab score advises a postoperative PT of  $<20^\circ$ ; however, for patients with a lower-limit PI this is not accurate enough because these patients require values  $<10^\circ$  (Fig. 4A: blue boxes on the left). Additionally, the PI-LL Schwab modifier does not apply for patients with extreme PI values either (Fig. 4B: blue boxes). Between a PI of approximately  $50^\circ$ – $80^\circ$  the PI-LL modifier is quite accurate. Furthermore, because the Global alignment (T1-SPI or SVA) is not regarded in relation to a patient’s PI, it is less accurate for the prediction of mechanical complications since it might underestimate the degree of malalignment (Fig. 4C: blue boxes). These drawbacks underline the need for more patient specific surgical thresholds as proposed by the GAP score.

A relatively novel subcategory of the GAP score is the LDI, which divides the lumbar lordosis in a lower arc (L4-S1) and an upper arc (L1-L3). As described by Rousouly et al. [28], maldistribution between these arcs might alter the distribution of loads on the spinal column and might subsequently cause mechanical failure. The importance of this subcategory for the prediction of mechanical complications via the GAP score in the current study was highlighted by the significant correlation coefficient ( $\rho=0.65$ ,  $p<.01$ ). Another addition to the GAP score was the subcategory “age.” However, in the current study age was not an important determinant for the GAP score, which is not in accordance with the literature [11]. Age, and especially age-related comorbidities (osteoporosis, sarcopenia, neurodegenerative factors), should always be considered during preoperative planning.

A good understanding of the principles of sagittal balance is vital to achieve optimum outcomes when treating ASD. The major difference between the Schwab classification and the GAP score is the fact the GAP score is patient tailored. Before the existence of the GAP score, the aim of surgical treatment was to restore the spinal column to “normal values” (that applied for every patient) rather than patient-specific values. This might also explain why Smith et al. [18] found that complete correction of sagittal imbalance, based on an SVA of less than 50 mm postoperatively, was associated with a higher incidence of acute PJF. And why Glattes et al. showed that changes in SVA did not contribute significantly to the development of abnormal proximal junctional kyphosis or PJF [15].

The current study was limited by a relatively small number of patients, although the differences that were observed were confirmed by statistical significance. Another limitation is that we had to substitute the SVA modifier of the Schwab classification by the T1-SPI. However, it was described by Lafage and Schwab [23] that the T1-SPI is strongly correlated with the SVA and therefore could be used when no calibrated radiographs are available.

It should be acknowledged that deformity surgery is a challenging and complex procedure with high reported complication rates. Surgical complications encompass a variety of entities which cannot all be predicted by means of radiographs (eg, neurologic deficits, incidental durotomy, and wound dehiscence). Furthermore, it should be taken into account that inadequate (pre- or postoperative) spinal alignment is not the only risk factor for complications. Older age, high body mass index, and low bone mineral density are important risk factors which should be taken into consideration when planning surgical treatment [12–15].

In conclusion, the ability to predict radiographic mechanical complications following ASD surgery by the Schwab classification and the GAP score was studied. Both classification systems were able to predict mechanical complications, however the GAP score proved to be significantly more appropriate. The difference found is probably attributed to the fact that in the GAP score all parameters are related to the patient's individual PI.

## References

- [1] Barrey C, Roussouly P, Perrin G, Le Huec JC. Sagittal balance disorders in severe degenerative spine. Can we identify the compensatory mechanisms? *Eur Spine J* 2011;20(Suppl 5):626–33.
- [2] Glassman SD, Bridwell K, Dimar JR, Horton W, Berven S, Schwab F, et al. The impact of positive sagittal balance in adult spinal deformity. *Spine (Phila Pa 1976)* 2005;30:2024–9.
- [3] Diebo BG, Henry J, Lafage V, Berjano P. Sagittal deformities of the spine: factors influencing the outcomes and complications. *Eur Spine J* 2015;24(Suppl 1):S3–S15.
- [4] Yamada K, Nakamae T, Shimbo T, Kanazawa T, Okuda T, Takata H, et al. Targeted therapy for low back pain in elderly degenerative lumbar scoliosis: a cohort study. *Spine (Phila Pa 1976)* 2016;41:872–9.
- [5] Takahashi S, Delecrist J, Passuti N. Surgical treatment of idiopathic scoliosis in adults: an age-related analysis of outcome. *Spine (Phila Pa 1976)* 2002;27:1742–8.
- [6] Shapiro GS, Taira G, Boachie-Adjei O. Results of surgical treatment of adult idiopathic scoliosis with low back pain and spinal stenosis: a study of long-term clinical radiographic outcomes. *Spine (Phila Pa 1976)* 2003;28:358–63.
- [7] Schwab FJ, Hawkinson N, Lafage V, Smith JS, Hart R, Mundis G, et al. Risk factors for major peri-operative complications in adult spinal deformity surgery: a multi-center review of 953 consecutive patients. *Eur Spine J* 2012;21:2603–10.
- [8] Bianco K, Norton R, Schwab F, Smith JS, Klineberg E, Obeid I, et al. Complications and intercenter variability of three-column osteotomies for spinal deformity surgery: a retrospective review of 423 patients. *Neurosurg Focus*, 2014;36:E18.
- [9] Pichelmann MA, Lenke LG, Bridwell KH, Good CR, O'Leary PT, Sides BA. Revision rates following primary adult spinal deformity surgery: six hundred forty-three consecutive patients followed-up to twenty-two years postoperative. *Spine (Phila Pa 1976)* 2010;35:219–26.
- [10] Hostin R, McCarthy I, O'Brien M, Bess S, Line B, Boachie-Adjei O, et al. Incidence, mode, and location of acute proximal junctional failures after surgical treatment of adult spinal deformity. *Spine (Phila Pa 1976)* 2013;38:1008–15.
- [11] Lau D, Clark AJ, Scheer JK, Daubs MD, Coe JD, Paonessa KJ, et al. Proximal junctional kyphosis and failure after spinal deformity surgery: a systematic review of the literature as a background to classification development. *Spine (Phila Pa 1976)*, 2014;39:2093–102.
- [12] Yagi M, Rahm M, Gaines R, Maziad A, Ross T, Kim HJ, et al. Characterization and surgical outcomes of proximal junctional failure in surgically treated patients with adult spinal deformity. *Spine (Phila Pa 1976)* 2014;39:E607–14.
- [13] O'Leary PT, Bridwell KH, Lenke LG, Good CR, Pichelmann MA, Buchowski JM, et al. Risk factors and outcomes for catastrophic failures at the top of long pedicle screw constructs: a matched cohort analysis performed at a single center. *Spine (Phila Pa 1976)*, 2009;34:2134–9.
- [14] Kim HJ, Bridwell KH, Lenke LG, Park MS, Song KS, Piyakulkaew C, et al. Patients with proximal junctional kyphosis requiring revision surgery have higher postoperative lumbar lordosis and larger sagittal balance corrections. *Spine (Phila Pa 1976)* 2014;39:E576–80.
- [15] Bridwell KH, Lenke LG, Cho SK, Pahys JM, Zebala LP, Dorward IG, et al. Proximal junctional kyphosis in primary adult deformity surgery: evaluation of 20 degrees as a critical angle. *Neurosurgery* 2013;72:899–906.
- [16] Bridwell KH, Baldus C, Berven S, Edwards C, Glassman S, Hamill C, et al. Changes in radiographic and clinical outcomes with primary treatment adult spinal deformity surgeries from two years to three- to five-years follow-up. *Spine (Phila Pa 1976)* 2010;35:1849–54.
- [17] Cho KJ, Suk SI, Park SR, Kim JH, Kang SB, Kim HS, et al. Risk factors of sagittal decompensation after long posterior instrumentation and fusion for degenerative lumbar scoliosis. *Spine (Phila Pa 1976)* 2010;35:1595–601.
- [18] Smith MW, Annis P, Lawrence BD, Daubs MD, Brodke DS. Acute proximal junctional failure in patients with preoperative sagittal imbalance. *Spine J* 2015;15:2142–8.
- [19] Schwab F, Ungar B, Blondel B, Buchowski J, Coe J, Deinlein D, et al. Scoliosis Research Society-Schwab adult spinal deformity classification: a validation study. *Spine (Phila Pa 1976)* 2012;37:1077–82.
- [20] Yilgor C, Sogunmez N, Boissiere L, Yavuz Y, Obeid I, Kleinstück F, et al. Global alignment and proportion (GAP) score: development and validation of a new method of analyzing spinopelvic alignment to predict mechanical complications after adult spinal deformity surgery. *J Bone Joint Surg Am* 2017;99:1661–72.
- [21] Schwab FJ, Smith VA, Biserni N, Gamez L, Farcy JP, Pagala M. Adult scoliosis: a quantitative radiographic and clinical analysis. *Spine (Phila Pa 1976)* 2002;27:387–92.
- [22] Soroceanu A, Diebo BG, Burton D, Smith JS, Deviren V, Shafrey C, et al. Radiographical and Implant-related complications in adult spinal deformity surgery: incidence, patient risk factors, and impact on health-related quality of life. *Spine (Phila Pa 1976)* 2015;40:1414–21.
- [23] Lafage V, Schwab F, Patel A, Hawkinson N, Farcy JP. Pelvic tilt and truncal inclination: two key radiographic parameters in the setting of adults with spinal deformity. *Spine (Phila Pa 1976)*, 2009;34: E599–606.
- [24] Protosaltis T, Schwab F, Bronsard N, Smith JS, Klineberg E, Mundis G, et al. The T1 pelvic angle, a novel radiographic measure of global sagittal deformity, accounts for both spinal inclination and pelvic tilt and correlates with health-related quality of life. *J Bone Joint Surg Am* 2014;96:1631–40.
- [25] Obeid I, Bourghli A, Larrieu D, Laouissat F, Challier V, Pointillart V, et al. The global tilt: evaluation of a parameter considering the global spinopelvic alignment. *J Med Liban*, 2016;64:146–51.
- [26] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 1988;44:837–45.
- [27] Obeid I, Boissiere L, Yilgor C, Larrieu D, Pellisé F, Alanay A, et al. Global tilt: a single parameter incorporating spinal and pelvic sagittal parameters and least affected by patient positioning. *Eur Spine J* 2016;25:3644–9.
- [28] Roussouly P, Nnadi C. Sagittal plane deformity: an overview of interpretation and management. *Eur Spine J* 2010;19:1824–36.