Spine Deformity (2024) 12:1127–1136
https://doi.org/10.1007/s43390-024-00864-5

CASE SERIES

Which sagittal plane assessment method is most predictive of complications after adult spinal deformity surgery?

Javier Pizones[1] · **Jeffrey Hills**[2] · **Michael Kelly**[3] · **Caglar Yilgor**[4] · **Lucía Moreno-Manzanaro**[1] · **Francisco Javier Sánchez Perez-Grueso**[1] · **Frank Kleinstück**[5] · **Ibrahim Obeid**[6] · **Ahmet Alanay**[4] · **Ferran Pellisé**[7] · **ESSG European Spine Study Group**[8]

Received: 14 November 2023 / Accepted: 16 March 2024 / Published online: 12 April 2024

Abstract

Purpose Different methods of sagittal alignment assessment compete for predicting adverse events after adult spinal deformity (ASD) surgery. We wanted to study which method provides greater benefit.

Methods Retrospective study of 391 patients operated for ASD, with > 6 instrumented levels, fused to the pelvis, and 2 years of follow-up. Three alignment methods were analyzed 6-week postoperatively: (1) Roussouly mismatch; (2) GAP score/GAP categories; (3) T4-L1-Hip axis. Binary logistic regression generated models that best predict the following adverse events: mechanical complications (MC): in general and isolated (PJK, PJF, rod breakage); reinterventions (in general and after MC); and readmissions. ROC/AUC analysis was also implemented. In a second regression round, we added different variables that were selected on univariate analysis—demographic, surgical, and radiographic—to complete the models.

Results The best predictor parameters in most models were T4-L1PA mismatch and GAP score; we could not prove a predictive ability of the Roussouly mismatch. The T4-L1PA mismatch best predicted general MC, PJK, PJK + PJF, and readmission, while the GAP score best predicted PJF and reinterventions (for MC and for any complication). However, the variance explained by these models was limited (Nagelkerke's $R2 = 0.031$–0.113), with odds ratios ranging from 1.070 to 1.456. ROC curves plotted an AUC between 0.57 and 0.70. Introducing additional variables (demographic, surgical, and radiographic) improved prediction in all the models (Nagelkerke's $R2 = 0.082$–0.329) and allowed predicting rod breakage.

Conclusion The T4-L1-Hip axis and GAP score show potential in predicting adverse events, surpassing the Roussouly method. Despite partial efficacy in complication anticipation, recognizing postoperative sagittal alignment as a key modifiable risk factor, the crucial need arises to integrate diverse variables, both modifiable and non-modifiable, for enhanced predictive accuracy.

Level of evidence Level IV.

Keywords Adult spinal deformity · Sagittal alignment · Predictive models · Mechanical complications · Reinterventions · Fused to pelvis · Spinal surgery

✉ Javier Pizones
javierpizones@gmail.com

1 Spine Unit, Department of Orthopedic Surgery, Hospital Universitario La Paz, Paseo de la Castellana 261, 28046 Madrid, Spain

2 University of Texas, San Antonio, TX, USA

3 Rady Children Hospital San Diego, University of California, San Diego, CA, USA

4 Department of Orthopedics and Traumatology, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

5 Department of Orthopedics, Schulthess Klinik, Zurich, Switzerland

6 Spine Surgery Unit, Pellegrin University Hospital, Bordeaux, France

7 Spine Surgery Unit, Hospital Universitario Vall d'Hebron, Barcelona, Spain

8 Spine Research Unit, Vall d'Hebron Institute of Research, Barcelona, Spain

Introduction

Sagittal alignment has been deeply researched since the transition of this century. The anatomical basis to understand spinal alignment was originated in the French school [1, 2], and the clinical application in adult spinal deformity (ASD) was first highlighted by Glassman and Schwab [3–5]. Sagittal alignment is now widely accepted to be a critical factor in patient assessment and surgical decision making. Postoperative global sagittal alignment and the residual body compensation (mainly pelvic retroversion) present in fused patients have a large impact on clinical results and potential complications [6], stressing the importance of an adequate sagittal plane restoration when performing surgery.

Traditionally, two classifications have been used to assess the sagittal plane: the Roussouly classification [7], and the SRS-Schwab classification [5, 8]; and more recently two methods were proposed to evaluate long-term implications of realignment: the GAP score [9] and the T4-L1-Hip axis [10]. All four have been investigated for their association with clinical outcomes and mechanical complications. Sometimes different methods are used in combination [11]. However, there is no consensus on which of these methods is more precise to predict complications following ASD surgery. Prior investigations show inconsistent results depending on the study cohort characteristics, the methodology used to assess complications, and the analyzed adverse events [12–17]. Importantly, the primary outcome frequently varies between studies (mechanical complications vs only proximal junctional kyphosis vs readmissions, etc.), leading to difficult interpretations when the methods are compared [18–21].

The aim of our study was to analyze a cohort of ASD operated panlumbar fusions and compare three different alignment assessment methods based on normative data (Roussouly, GAP score, and T4-L1-Hip axis) to determine the method most predictive of different adverse events.

Methods

We conducted a retrospective study with patients taken from a prospective multicenter database (ESSG: European Spine Study Group) comprising adult spinal deformity (ASD) patients. General inclusion criteria for the database were patients older than 18 years with at least one of the following requirements: coronal Cobb ≥ 20°, sagittal vertical axis (SVA) ≥ 5 cm, pelvic tilt (PT) ≥ 25°, or thoracic kyphosis ≥ 60°. Institutional review board approval was obtained at each participating institution prior to patient enrollment. Informed consent was obtained from each patient.

For this specific study we selected patients fulfilling the following criteria: ASD operated patients, more than 6 instrumented levels (UIV at or above L1), fused distally to the pelvis, and complete data at the immediate postoperative (6 weeks) period and 2 years after surgery.

We analyzed three different alignment methods based on normative data in the 6-week postoperative radiograph: (1) mismatch according to the Roussouly classification (categorical variable: matched or mismatched): the postoperative type was calculated based on the 6-week sacral slope magnitude [7] compared to the ideal type based on PI [22]; (2) GAP score (continuous variable) and GAP categories (categorical variable: proportioned, moderately disproportioned, severely disproportioned) as described by Yilgor et al. [9]; (3) T4-L1-Hip axis method as termed by Hills et al. [10]: composed of two different continuous variables: the relative L1PA, which was calculated at 6 weeks as (L1PA minus ideal L1PA) (ideal L1PA was calculated based on normative data using the formula 0.5xPI-21); (4) and T4-L1PA mismatch (difference between 6-week T4PA and 6-week L1PA (mismatch considered as > 4º)).

We performed binary logistic regressions with stepwise likelihood ratio method, generating models that best predicted each of the following adverse events:

- mechanical complications in isolation:
    - proximal junctional kyphosis—PJK: a difference of more than 25° in the proximal junctional angle (PJA) measured from the immediate postoperative sagittal radiograph to a given successive follow-up sagittal radiograph.
    - proximal junctional failure—PJF: PJA failure due to fracture of UIV or UIV + 1, implant failure—pull-out/dislodgment, or proximal spondylolisthesis, usually causing symptoms. All required revision surgery
    - PJK and PJF together: this variable groups all PJK (as per definition, only angular PJA kyphosis > 25º mechanical complications in isolation) and PJF (proximal failure needing revision surgery).
    - rod breakage: breakage of one or more rods detected in the follow-up radiographs.
- mechanical complications (MC) in general: any mechanical complication; this variable groups proximal junctional complications and rod breakages and instrumentation failure (screw dislodgment, pull-out, osteolysis).
- reinterventions in general: any unplanned reoperation after the index surgery for any complication that includes MC and non-MC reinterventions such as infections, postoperative radicular pain, and pseudomeningoceles.

- reinterventions due to MC: any unplanned reoperation after the index surgery due to a mechanical complication (PJK, PJF, rod breakage, screw pull-out, etc.)
- readmissions: any unplanned readmission due to a complication secondary to the index surgery.

Receiver operating characteristics (ROC) analysis was performed to calculate the predictive ability of the models in the occurrence of each adverse event. The area under the curve (AUC) was used to evaluate the accuracy of each evaluation system together with the Odds Ratio (OR) and 95% confidence intervals.

Univariate analysis was carried out with additional variables (baseline demographic and radiographic, surgical, and 6-week radiographic) and tested for each of the adverse events previously listed. Those variables showing statistical difference -set at $p < 0.05$- (as risk factors for a specific adverse event) were added to the previous models, and binary logistic regression analysis were run again to create a second round of predictive models that combined the previous assessment aligning methods with the additional selected variables (listed below).

- demographic variables: age, gender, ASA score, and body mass index (BMI)
- baseline PROMS: ODI score, SRS-22 subtotal, SF-36 PCS, and SF-36 MCS
- surgical variables: surgical approach, surgical time, estimated blood loss, upper instrumented vertebra (UIV), and the use of spinal osteotomies.
- radiographic variables: preoperative and at 6 weeks: pelvic incidence (PI), lumbar lordosis (LL), T2–T12 kyphosis, T10–L2 kyphosis, relative spinopelvic alignment (RSA) based on global tilt (GT), relative lumbar lordosis (RLL) based on L1–S1, lumbar distribution index (LDI) (L4–S1/L1–S1 × 100), relative pelvic version (RPV) based on sacral slope (SS), T4 pelvic angle (T4PA), and L1 pelvic angle (L1PA).

Statistical analysis was carried out using the SPSS software (version 20, SAS Institute Inc., Cary, NC). Normality of the variables was tested using Kolmogorov–Smirnov test. The distribution of quantitative variables is given as mean and standard deviation or median and quartiles ($Q1$; $Q3$) as required; qualitative values are expressed in percentages. Univariate analysis was performed when comparing qualitative variables using the chi-square test and the Fisher exact test, and quantitative variables using the Student *t*-test or Mann–Whitney *U*. The significance threshold was set at 5% ($p < 0.05$), and those that demonstrated statistical significance were used in the multivariable analysis. Binary logistic regression was performed with the stepwise likelihood ratio method and the variance explained by Nagelkerke's $R2$. ROC curves were used to plot AUC, and OR was calculated to assess alignment methods' predictive accuracy.

Results

391 patients met inclusion criteria and were analyzed in this study. Table 1 summarizes the baseline demographic, radiographic and surgical variables of the cohort. We can highlight a 45.5% rate (178) of MC in this series of pan-lumbar fusions (rod breakage 18.7% (73), PJK 11.8% (46), PJF 5.4% (21), PJK + PJF = 17.3% (67), other (dislodgment, screw pull-out, osteolysis) 9.6% (38)), with 36.1% (141) rate of general reinterventions and 28.4% (111) of reinterventions due to MC, and 31.2% rate (122) of readmissions.

In the first round of binary logistic regression models performed to competitively test the three alignment assessment methods measured at 6 weeks postoperatively, we found that the best overall predictor parameters of adverse events in most models were T4-L1PA mismatch and GAP score. We could not prove a predictive ability of the Roussouly mismatch. The T4-L1PA mismatch best predicted general MC, PJK, PJK + PJF, and readmissions, while the

Table 1 Sample description

Number of patients	391
Age (years)	65 ± 10.2; 66 (59;72)
BMI	26.1 ± 4.5; 25.6(23;29.1)
ASA score	2.1 ± 0.5; 2(2;2)
Gender	82.1% female
Diagnosis	Degenerative 61.5% (243) Idiopathic 23% (91) Failed-back 7.6% (30) Posttraumatic 4% (12) Congenital 1.8% (7) Neuromuscular 1.3% (5) Ankylosing spondylitis 0.8% (3)
Prior spine surgery	45.9% (181)
Follow-up (years)	4.8 ± 1.8
Baseline ODI	49.3 ± 17.6
Baseline SRS-22Total	2.53 ± 0.6
Baseline SF-36PCS	32.6 ± 7.7
Baseline SF-36MCS	40.6 ± 12.8
Pelvic Incidence (º)	57.9 ± 13.7
Baseline GAP score	9.1 ± 3.8
Baseline T4PA	26.6 ± 12.5
Baseline L1PA	16.5 ± 11.1
Surgical approach	89% only posterior; 11% anterior + posterior approach
Upper instrumented vertebra	T1–T6 24.7%; T8–L1 75.3%
Number of levels fused	11.2 ± 3.4

Data expressed as mean ± standard deviation or median and quartiles

GAP score best predicted PJF and reinterventions due to MC and to any complication (Table 2). We were not able to predict rod breakage in this first round by any of the methods ($p > 0.05$). The variance (predictive ability, Nagelkerke's $R2$) of the adverse events explained by these models based ranged between 0.031 and 0.113, and the odds ratios (expressed as exponential b) ranged between 1.070 and 1.456 (Table 2).

To measure the predictive accuracy of the different sagittal alignment methods, ROC curves were plotted for each adverse event, obtaining a range of AUC between 0.57 and 0.70 (Table 3, Fig. 1). When this analysis was performed, the GAP score performed almost as good as the T4-L1PA mismatch to predict MC, PJK + PJF, and readmissions. The GAP score best predicted PJF and reintervention due to any complication, while T4-L1PA mismatch best predicted PJK.

Table 2 First round of binary logistic regression

Adverse event	Evaluation method selected by the model	*B*	T.E.	Wald	*df*	Sig.	Exp(*B*)	*R*2 Nagelkerke
Mechanical complication	T4-L1PA 6w mismatch	0.067	0.027	6.261	1	0.012	1.070	0.035
	Constant	– 0.543	0.171	10.014	1	0.002	0.581	
PJK	T4-L1PA 6w mismatch	0.088	0.037	5.598	1	0.018	1.092	0.041
	Constant	– 2.34	0.283	69.014	1	0.000	0.096	
PJF	6w GAP score	0.375	0.120	9.755	1	0.002	1.456	0.113
	Relative L1PA 6w	– 0.113	0.58	3.802	1	0.51	0.893	
	Constant	– 4.24	0.663	41.005	1	0.000	0.014	
PJK + PJF	T4-L1PA 6w mismatch	0.099	0.033	8.956	1	0.003	1.104	0.05
	Constant	– 1.94	0.244	63.304	1	0.000	0.143	
Rod breakage	Roussouly 6w mismatch	0.407	0.337	1.458	1	0.227	1.502	0.000
	Constant	– 1.78	0.248	51.705	1	0.000	0.168	
Reintervention due to mechanical complication	6w GAP score	0.113	0.060	3.554	1	0.059	1.120	0.031
	Constant	– 0.63	0.306	4.238	1	0.040	0.533	
Reintervention due to any complication	6w GAP score	0.155	0.064	5.929	1	0.015	1.167	0.05
	Constant	– 0.31	0.306	1.043	1	0.307	0.732	
Readmission	T4-L1PA 6w mismatch	0.087	0.029	8.987	1	0.003	1.091	0.05
	Constant	– 1.20	0.196	37.542	1	0.000	0.302	

For each adverse event, the best predictive model is calculated with the different sagittal alignment evaluation methods

Variables(s) introduced in phase 1: Postoperative Roussouly mismatch, postoperative GAP score, postoperative GAP category, T4-L1PA postoperative mismatch, relative L1PA postoperative mismatch

T.E. technical estimation, *df* degrees of freedom

Table 3 ROC curves calculated for each adverse event with each created model

Adverse event	Parameter	Area	95% CI	Sig
Mechanical complication	T4-L1PA 6w mismatch	0.58	0.51–0.65	0.025
	6w GAP score	0.57	0.50–0.64	0.049
PJK	T4-L1PAA 6w mismatch	0.659	0.56–0.75	0.004
PJF	6w GAP score	0.701	0.57–0.83	0.009
PJK + PJF	T4-L1PA 6w mismatch	0.646	0.56–0.73	0.002
	6w GAP score	0.604	0.52–0.69	0.027
Rod breakage	No parameter showed significance			
Reintervention due to mechanical complication	No parameter showed significance			
Reintervention due to any complication	6w GAP score	0.616	0.53–0.71	0.14
	6w GAP category	0.594	0.50–0.69	0.047
Readmission	T4-L1PA 6w mismatch	0.597	0.52–0.67	0.014
	6w GAP Score	0.598	0.52–0.67	0.013

PJK proximal junctional kyphosis, *PJF* proximal junctional failure, *CI* confidence interval

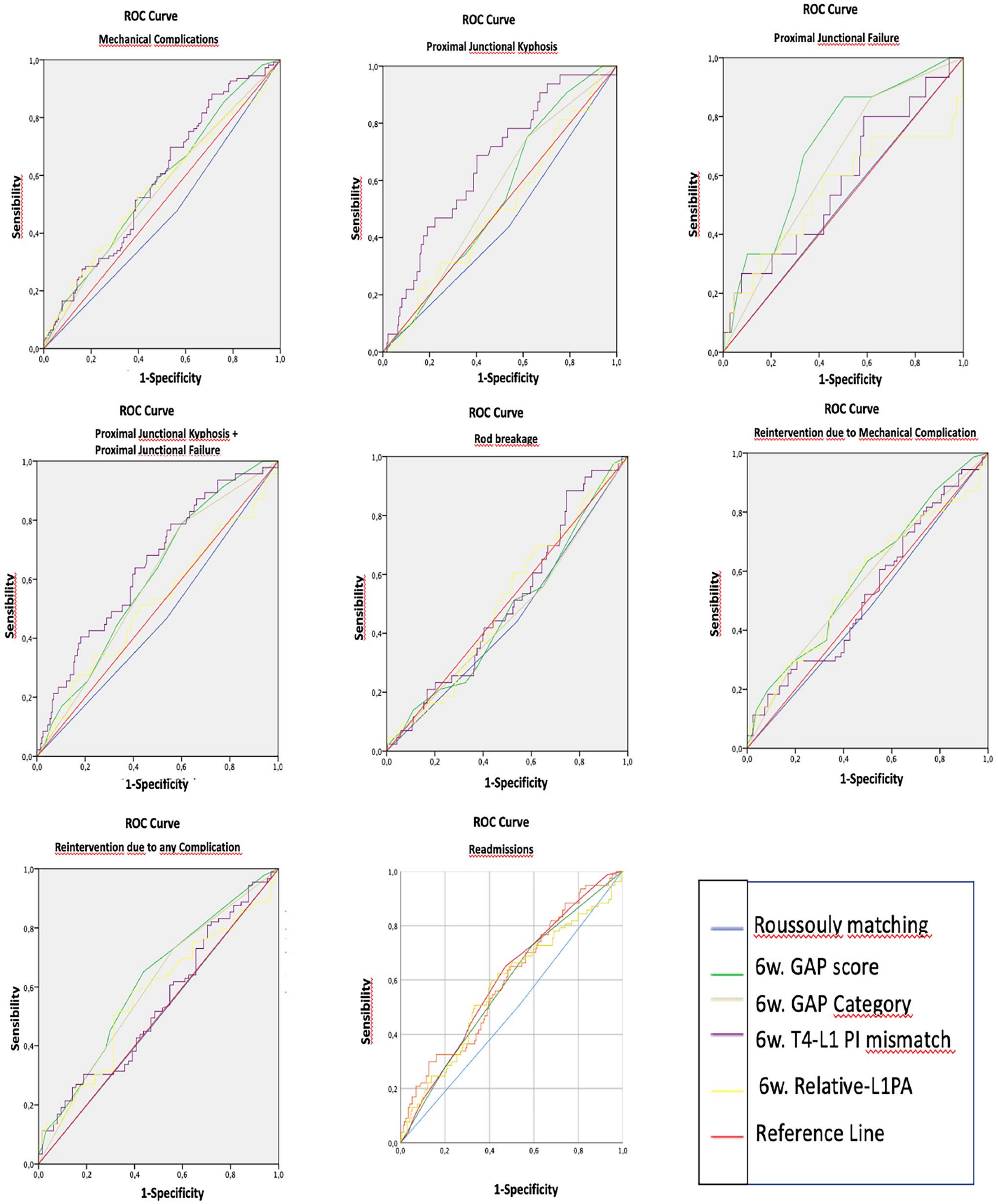

Fig. 1 Area under the curve (AUC) showing the predictive ability of the different sagittal alignment methods for each adverse event

Introducing additional variables in a second round of binary regressions improved predictions in all models (Nagelkerke's $R2$ between 0.082 and 0.329). And only by adding them to the models was it possible to predict rod breakage together with the T4-L1-Hip axis method. Depending on the model, the following variables were significant (Table 4): baseline demographic (ODI, SF-36PCS, SF-36-MCS, ASA score); surgical (blood loss, UIV level); 6w postoperative radiographic (T2–T12 kyphosis, RSA, RLL, T4PA, L1PA, and LDI).

Discussion

Different methods for sagittal alignment assessment compete for predicting adverse events after ASD surgery [7–10]. These methods help surgical planning, allowing surgeons to personalize alignment goals for each patient using sagittal preoperative measures. Published results are contradictory depending on the investigated adverse event, the characteristics of the selected cohort, and the analyzed alignment method [18–21]. Evaluating the predictive ability of three of these methods in a cohort of panlumbar fused ASD patients, we have seen in our study that the best predictor parameters in most models were T4-L1PA mismatch and GAP score; we could not prove a predictive ability of the Roussouly mismatch.

The T4-L1-Hip axis method [10] was developed after the analysis of a disease-free multinational volunteer cohort of individuals with normal sagittal balance and spinopelvic alignment. The T4-L1-Hip axis method defines ideal alignment using parameters that are either fixed or directly modifiable in surgery, including pelvic incidence (PI), the L1PA (defined as a function of PI), and the T4-L1PA mismatch (which describes a thoracic alignment relative to the lumbar). Sagittal alignment targets should restore patients to an ideal shape: ideal L1PA based on PI with the formula L1PA = 0.5 × PI-21; and T4PA should be nearly equivalent to the L1PA (within 4º), aligning the T4-L1-Hip axis. While this method proposing the correct position of T4 over L1 in space appears to be accurate for describing normal thoracic and lumbar alignment, no studies have yet confirmed that realignment based on these targets improves outcomes. The method has been described in a set of long thoracolumbar fusions, surgeries in which both T4 and L1 are captured and positioned with instrumentation. The utility of this alignment scheme is unclear in patients with an upper instrumented vertebra in the lower thoracic spine.

The GAP score [9] includes 4 sagittal parameters (L1–S1 lumbar lordosis, L4–S1 lumbar lordosis, sacral slope and global tilt), and targets are calculated based on the ideal situation of each one (respectively, relative lumbar lordosis—RLL, lumbar distribution index—LDI, relative pelvic version—RPV, and relative spinopelvic alignment—RSA) calculated with formulae based on PI. An age modifier adds a biological component to the prediction. The score stratifies patients into three categories originally associated with an increasing rate of mechanical complications. In a recent systematic review analyzing GAP score capacity in predicting MC occurrence, authors studied eleven retrospective articles plotting a global AUC of 0.68 ± 0.2, showing a moderate predictive accuracy [23].

The Roussouly classification [7] was created from a cohort of healthy population, and groups patients depending on their sacral slope in 4 different types. For each type there is a characteristic location of the lumbar apex and the inflection point, as well as a specific number of vertebrae included in the lordosis and lordosis arches distribution. These characteristics have later been proposed as ideal targets to set when surgically restoring the sagittal profile [14, 24].

The SRS-Schwab classification [5, 8] was not used in our analysis as it is the only one not based on normative data from a cohort of asymptomatic subjects, but on a series of adult deformity patients. This classification uses three sagittal modifiers to quantify deformity (sagittal vertical axis-SVA, pelvic tilt-PT, and PI-LL mismatch). It sets thresholds for each parameter based on the correlation between radiographic parameters and HRQOL measures (mainly disability using ODI over or under 40). This classification was not conceived as a predictor of complications, and despite achieving optimal values postoperatively, mechanical complications are not uncommon [25]. We believe this is because it is not patient tailored and might underestimate the malalignment with regard to a patient's individual PI [9, 18]. For example, the threshold PT > 20º is considered pathological in the classification, but this value should be considered normal in patients with high PI [26].

A few papers have compared these different methods of alignment assessment. Jacobs et al. [18] highlighted GAP's ability (AUC for GAP score was 0.86) over the SRS-Schwab classification (AUC = 0.69) in predicting mechanical complications and attributed it to the fact that in the GAP score all parameters are related to patient`s individual PI (it is patient tailored), making RPV a better parameter than pelvic tilt—PT, and RLL a better parameter than PI-LL mismatch. The GAP score was also found to be more effective in predicting PJK (AUC = 0.863) and PJF (AUC = 0.724) than the Roussouly classification by Sun et al. [19]. They further showed GAP categories (AUC = 0.561) to be equally effective than GAP score (AUC = 0.555) in predicting implant-related complications. On the other hand, Teles et al. [21] showed the SRS-Schwab to be more predictive of mechanical complications (AUC 0.67) than the GAP score (AUC 0.53).

In our study, the predictive accuracy of the tested methods varied depending on the adverse event to predict and the

Table 4 Binary logistic regression crating models to predict adverse events with all the variables selected by univariate analysis

Adverse event	B	E.T.	Wald	gl	Sig.	Exp (B)	R2 Nagelkerke
Mechanical complication							0.135
ASA	−0.485	0.256	3.579	1	0.059	0.616	
Blood loss	0.000	0.000	5.162	1	0.023	1.000	
ODI	−0.023	0.010	5.380	1	0.020	0.977	
SF-36PCS	−0.051	0.022	5.531	1	0.019	0.950	
T4PA 6w	0.099	0.035	8.160	1	0.004	1.104	
L1PA 6w	−0.087	0.043	4.109	1	0.043	0.917	
LDI 6w	−1.710	0.843	4.115	1	0.043	0.181	
Constant	3.593	1.40	6.551	1	0.010	36.33	
PJK							0.144
GAP score 6w	−0.241	0.116	4.294	1	0.038	0.786	
Age	0.051	0.027	3.490	1	0.062	1.052	
SF-36PCS	−0.058	0.029	3.978	1	0.046	0.944	
T4PA 6w	0.164	0.055	8.820	1	0.003	1.178	
L1PA 6w	−0.138	0.057	5.949	1	0.015	0.871	
Constant	−3.639	1.95	3.464	1	0.063	0.026	
PJF							0.329
GAP score 6w	1.200	0.345	12.085	1	0.001	3.320	
GAP category 6w			4.972	2	0.083		
GAP category (1)	5.395	2.58	4.356	1	0.037	220.300	
GAP category (2)	3.436	1.55	4.909	1	0.027	31.070	
ODI	−0.058	0.033	3.069	1	0.080	0.944	
Srs-22Total	−1.986	1.07	3.445	1	0.063	0.137	
SF-36MCS	0.062	0.037	2.865	1	0.091	1.064	
RLL 6w	0.187	0.062	9.198	1	0.002	1.206	
Constant	−5.026	4.19	1.437	1	0.231	0.007	
PJK + PJF							0.129
ODI	−0.028	0.013	4.483	1	0.034	0.972	
SF-36PCS	−0.076	0.028	7.328	1	0.007	0.927	
T2–T12 6w	0.029	0.012	6.308	1	0.012	1.030	
T4PA6w	0.042	0.020	4.443	1	0.035	1.043	
Constant	0.148	1.47	0.010	1	0.920	1.160	
Rod breakage							0.144
T4-L1PA 6w mismatch	0.226	0.102	4.912	1	0.027	1.254	
Relative L1PA 6w mismatch	0.132	0.071	3.426	1	0.064	1.141	
Blood loss	0.001	0.000	9.943	1	0.002	1.001	
RSA 6w	−0.198	0.076	6.856	1	0.009	0.820	
LDI6w	−1.733	0.915	3.589	1	0.058	0.177	
Constant	−0.785	0.860	0.833	1	0.361	0.456	
Reintervention due to mechanical complication							0.254
GAP score 6w	0.484	0.213	5.156	1	0.023	1.622	
GAP category			4.627	2	0.099		
GAP category (1)	2.998	1.39	4.608	1	0.032	20.04	
GAP category (2)	1.600	0.867	3.406	1	0.065	4.955	
T4-L1PA 6w mismatch	−0.462	0.304	2.320	1	0.128	0.630	
ASA	−0.911	0.349	6.806	1	0.009	0.402	
UIV(1)	0.912	0.489	3.479	1	0.062	2.489	
T4PA 6w	0.542	0.308	3.101	1	0.078	1.720	
L1PA 6w	−0.450	0.312	2.086	1	0.149	0.638	
RLL 6w	0.048	0.023	4.140	1	0.042	1.049	

Table 4 (continued)

Adverse event	B	E.T.	Wald	gl	Sig.	Exp (B)	R2 Nagelkerke
Constant	−2.924	1.72	2.889	1	0.089	0.054	
Reintervention due to any complication							0.235
GAP score 6w	0.184	0.104	3.138	1	0.076	1.202	
ASA	−1.230	0.361	11.63	1	0.001	0.292	
UIV(1)	1.055	0.503	4.406	1	0.036	2.873	
Srs-22Total	1.007	0.517	3.800	1	0.051	2.738	
SF-36PCS	−0.062	0.029	4.560	1	0.033	0.939	
SF-36MCS	−0.056	0.024	5.583	1	0.018	0.946	
T4PA 6w	0.073	0.033	4.950	1	0.026	1.075	
RLL 6w	0.044	0.023	3.641	1	0.056	1.045	
Constant	3.171	1.35	5.479	1	0.019	23.83	
Readmission							0.082
UIV(1)	0.834	0.369	5.106	1	0.024	2.302	
T4PA 6w	0.056	0.017	10.20	1	0.001	1.057	
Constant	−1.850	0.335	30.43	1	0.000	0.157	

Variable(s) introduced in step 1: 6w Roussouly mismatch. 6w GAP score. 6w GAP category 6w T4-L1PA mismatch. 6w L1PA relative post. age. ASA score. BMI. gender. Surgical time. Blood loss. Upper instrumented vertebra. Osteotomy. ODI. SRS-22Total. SF-36PCS. SF-36MCS. T2–T12 postop. T10–L2 postop. RSA postop. T4PA postop. L1PA postop. PI-LL postop. RPV postop. RLL postop. LDI post. RSA post

statistics applied. Readmissions, mechanical complications as a whole, PJK, and the combination of PJK and PJF were better predicted with the T4-L1PA mismatch using regression analysis. It is logical to think that if the ideal lumbar distribution is not achieved (L1PA) and thoracic anterior forces lie in front of L1 (mismatch T4PA), the fate of the construct is to fail over time. However, when ROC analysis was performed we discovered a similar predictive ability with the GAP score.

On the other hand, reintervention for mechanical complications and PJF were better predicted with the GAP score when logistic regression statistics were run. But the T4-L1-Hip axis seemed to predict similarly when ROC analysis was performed. Many authors have also reported GAP's ability to predict complications [18, 27, 28]. Even with the existing controversy with the use of this method, it is still a good guide to assess ideal alignment. Predictions are aided in this method by a biological marker (age), which we find necessary because, as we have seen in the current study, the accuracy in prediction when only sagittal alignment assessment is used has a lot of room for improvement.

We could not prove in our current study the impact of the Roussouly classification in predicting adverse events as seen in other publications [14, 29], and this is shared by other authors [19, 21]. This classification is less analytical than the other available methods. This might have been our fault as classifying patients only based on three SS categories, as we did in this study, might be misleading if the rest of the characteristics are not considered. First, because patients having SS values under 35º can be classified as either type 1 or type 2. Second, because SS is subjected to pelvic compensation, which may lead to misclassifying the patient having spinal pathology. Moreover, this classification used as a predictive method may be too simplistic to reflect the continuum.

The variance explained for adverse events with the current sagittal alignment methods was low (between 3 and 11%) using Nagelkerke's statistics, testing the overall performance of the models. This statistical tool is based on the log-likelihood as a type of scoring rule [30]. It can be interpreted as a measure of how close the prediction of the model is to the observed 0 and 1 outcomes. The AUC values also vary from 0 to 1; values of 0.5 indicate that the model performance is not better than random; 0.5–0.7 indicates poor performance. However, a word of caution should be taken when interpreting pseudo-$R2$ statistics as they do not represent the proportionate prediction in error as the $R2$ in linear regression does. We cannot conclude that only 3–11% of the adverse events are answered by the models, it just shows part of the quality of the models. This is due to its heteroscedasticity (the error of variance differs for each value of the predicted score).

While sagittal alignment has a clear and demonstrated impact on patient outcomes, sagittal malalignment represents one of many risk factors associated with postoperative complications. In our cohort predictions improved with the addition of variables such as PROMs (ODI, SF-36PCS, SF-36MCS), ASA score, surgical variables (blood loss, UIV level), and some radiographic postoperative parameters (T2–T12 kyphosis, RSA, RLL, T4PA, L1PA, and LDI).

We were not able to predict rod breakage by any of these alignment methods alone (consistent with Sun et al. [19]).

The addition of a surgical variable (blood loss) and several independent sagittal regional parameters (RSA, LDI) resulted in some improvement, but the utility and external validity of such a model are likely low. Therefore, other unknown factors apart from sagittal alignment, or the variables we currently collect in our databases, must be involved in adverse events and outcomes. The biomechanical loads supported by a fused spine [31] and patient's biological state likely play a considerable role in the risk of complications, such as bone quality [27], muscle status [32], patient comorbidities or frailty [33], and serum markers, telomere length, or omics [34]. All of these are future potential variables to add to these equations to improve predictability.

Limitations of the study comprise the retrospective design which may have introduced selection or information biases. As other articles, the selected cohort may not be comparable to other series. The use of postural radiographic parameters (pelvic tilt, sacral slope) and the use of postoperative radiographic parameters of unfused segments of the spine may bias the results. Additionally, there may be considerable difference in risk profiles for those patients with an upper thoracic UIV vs those with a lower thoracic UIV. Another limitation is that we did not account for rod type, number, or diameter in our study. These factors have the potential to impact the occurrence of specific mechanical complications. Finally, statistics need to be carefully interpreted as pseudo-$R2$ methods do not exactly represents biological processes. Furthermore, comparison between methods is difficult to interpret as their value measures were conceived in different scales. A prospective study comparing surgeons using different alignment schemes is needed to determine which planning scheme results in the least error and which planning scheme results in least complications.

In conclusion, our examination of the T4-L1-Hip axis and the GAP score suggests their potential superiority in predicting adverse events compared to the Roussouly method. Despite this, it is essential to acknowledge that the three analyzed methods exhibit only partial efficacy in anticipating complications. While postoperative sagittal alignment emerges as a pivotal modifiable risk factor, it is imperative to underscore the significance of incorporating a diverse range of variables—both modifiable and non-modifiable—for an enhanced predictive accuracy. Therefore, a comprehensive approach, encompassing a broader spectrum of factors, is indispensable for refining our predictive models and advancing the understanding of adverse events in this context.

Acknowledgements This study was presented at the GEER meeting 2023.

Author contributions All authors have done the following: Substantial contributions to the conception and design of the work, acquisition, analysis, or interpretation of data. Drafting the work and revising it critically for important intellectual content. Final approval of the version to be published. And all authors agree to be accountable for the author's own contributions and for ensuring that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and documented in the literature.

Funding A DePuy Synthes Spine and Medtronic research grants were received in partial support of this work. The device(s)/drug(s) is/are FDA approved or approved by the corresponding national agency for this indication. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Data availability All data were from a multicenter European database.

Declarations

Conflict of interest Javier Pizones, Jeffrey Hills, Lucía Moreno-Manzanaro, and Francisco Javier Sánchez Perez-Grueso have nothing to disclose. Michael Kelly received honorarium from Wolters Kluwer. Caglar Yilgor is a Consultant for Medtronic. Frank Kleinstück did teaching and speaking for DePuy Synthes. Ibrahim Obeid received grants from DePuy synthes and did consulting for DePuy synthes, Medtronic, Clariance, Spineart, Alphatec. Ahmet Alanay received grants from Medtronic, DePuy Synthes is a Consultant for Globus medical and ZimVie. Ferran Pellisé is a Consultant for Medtronic and Depuy Synthes.

Institutional review board Ethical approval was obtained before patient enrollment and data collection protocols. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed consent All the participating patients gave prior informed consent to their inclusion in this study. Patients signed informed consent regarding publishing their data and photographs.

References

1. Stagnara P, De Mauroy JC, Dran G et al (1982) Reciprocal angulation of vertebral bodies in a sagittal plane: approach to references for the evaluation of kyphosis and lordosis. Spine 7:335–342. https://doi.org/10.1097/00007632-198207000-00003
2. Duval-Beaupère G, Schmidt C, Cosson P (1992) A Barycentremetric study of the sagittal shape of spine and pelvis: the conditions required for an economic standing position. Ann Biomed Eng 20:451–462. https://doi.org/10.1007/BF02368136
3. Glassman SD, Berven S, Bridwell K et al (2005) Correlation of radiographic parameters and clinical symptoms in adult scoliosis. Spine 30:682–688. https://doi.org/10.1097/01.brs.0000155425.04536.f7
4. Glassman SD, Bridwell K, Dimar JR et al (2005) The impact of positive sagittal balance in adult spinal deformity. Spine 30:2024–2029
5. Schwab F, Farcy J-P, Bridwell K et al (2006) A clinical impact classification of scoliosis in the adult. Spine 31:2109–2114. https://doi.org/10.1097/01.brs.0000231725.38943.ab
6. Takemoto M, Boissière L, Vital J-M et al (2017) Are sagittal spinopelvic radiographic parameters significantly associated with quality of life of adult spinal deformity patients? Multivariate linear regression analyses for pre-operative and short-term post-operative health-related quality of life. Eur Spine J 26:2176–2186. https://doi.org/10.1007/s00586-016-4872-y

7. Roussouly P, Gollogly S, Berthonnaud E, Dimnet J (2005) Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position. Spine 30:346–353
8. Schwab F, Ungar B, Blondel B et al (2012) Scoliosis Research Society-Schwab adult spinal deformity classification: a validation study. Spine 37:1077–1082. https://doi.org/10.1097/BRS.0b013e31823e15e2
9. Yilgor C, Sogunmez N, Boissière L et al (2017) Global alignment and proportion (GAP) score. J Bone Joint Surg 99:1661–1672. https://doi.org/10.2106/JBJS.16.01594
10. Hills J, Lenke LG, Sardar ZM et al (2022) The T4–L1-Hip axis: defining a normal sagittal spinal alignment. Spine 47:1399–1406. https://doi.org/10.1097/BRS.0000000000004414
11. Passias PG, Pierce KE, Raman T et al (2021) Does matching roussouly spinal shape and improvement in SRS-Schwab modifier contribute to improved patient-reported outcomes? Spine 46:1258–1263. https://doi.org/10.1097/BRS.0000000000003999
12. Liu Y, Liu Z, Zhu F et al (2013) Validation and reliability analysis of the new SRS-Schwab classification for adult spinal deformity. Spine 38:902–908. https://doi.org/10.1097/BRS.0b013e318280c478
13. Terran J, Schwab F, Shaffrey CI et al (2013) The SRS-Schwab adult spinal deformity classification: assessment and clinical correlations based on a prospective operative and nonoperative cohort. Neurosurgery 73:559–568. https://doi.org/10.1227/NEU.0000000000000012
14. Sebaaly A, Gehrchen M, Silvestre C et al (2019) Mechanical complications in adult spinal deformity and the effect of restoring the spinal shapes according to the Roussouly classification: a multicentric study. Eur Spine J 37:1763–1810. https://doi.org/10.1007/s00586-019-06253-1
15. Bari TJ, Ohrt-Nissen S, Hansen LV et al (2019) Ability of the global alignment and proportion score to predict mechanical failure following adult spinal deformity surgery-validation in 149 patients with two-year follow-up. Spine Deformity 7:331–337. https://doi.org/10.1016/j.jspd.2018.08.002
16. Kwan KYH, Lenke LG, Shaffrey CI et al (2021) Are higher global alignment and proportion scores associated with increased risks of mechanical complications after adult spinal deformity surgery? An external validation. Clin Orthop Relat Res 479:312–320. https://doi.org/10.1097/CORR.0000000000001521
17. Gupta MC, Yilgor C, Moon HJ et al (2021) Evaluation of global alignment and proportion score in an independent database. Spine J 21:1549–1558. https://doi.org/10.1016/j.spinee.2021.04.004
18. Jacobs E, van Royen BJ, van Kuijk SMJ et al (2019) Prediction of mechanical complications in adult spinal deformity surgery-the GAP score versus the Schwab classification. Spine J 19:781–788. https://doi.org/10.1016/j.spinee.2018.11.013
19. Sun X, Sun W, Sun S et al (2021) Which sagittal evaluation system can effectively predict mechanical complications in the treatment of elderly patients with adult degenerative scoliosis? Roussouly classification or Global Alignment and Proportion (GAP) Score. J Orthop Surg Res 16:641–649. https://doi.org/10.1186/s13018-021-02786-8
20. Passias PG, Bortz C, Pierce KE et al (2022) Comparing and contrasting the clinical utility of sagittal spine alignment classification frameworks: Roussouly versus SRS-Schwab. Spine 47:455–462. https://doi.org/10.1097/BRS.0000000000004300
21. Teles AR, Aldebeyan S, Aoude A et al (2022) Mechanical complications in adult spinal deformity surgery: can spinal alignment explain everything? Spine 47:E1–E9. https://doi.org/10.1097/BRS.0000000000004217
22. Laouissat F, Sebaaly A, Gehrchen M, Roussouly P (2017) Classification of normal sagittal spine alignment: refounding the Roussouly classification. Eur Spine J 18:40–10. https://doi.org/10.1007/s00586-017-5111-x
23. Quarto E, Zanirato A, Pellegrini M et al (2022) GAP score potential in predicting post-operative spinal mechanical complications: a systematic review of the literature. Eur Spine J 31:3286–3295. https://doi.org/10.1007/s00586-022-07386-6
24. Pizones J, Perez-Grueso FJS, Moreno-Manzanaro L et al (2020) Ideal sagittal profile restoration and ideal lumbar apex positioning play an important role in postoperative mechanical complications after a lumbar PSO. Spine Deformity 8:491–498. https://doi.org/10.1007/s43390-019-00005-3
25. Soroceanu A, Diebo BG, Burton D et al (2015) Radiographical and implant-related complications in adult spinal deformity surgery: incidence, patient risk factors, and impact on health-related quality of life. Spine 40:1414–1421. https://doi.org/10.1097/BRS.0000000000001020
26. Yilgor C, Sogunmez N, Yavuz Y et al (2017) Relative lumbar lordosis and lordosis distribution index: individualized pelvic incidence-based proportional parameters that quantify lumbar lordosis more precisely than the concept of pelvic incidence minus lumbar lordosis. Neurosurg Focus 43:E5-9. https://doi.org/10.3171/2017.8.FOCUS17498
27. Noh SH, Ha Y, Park JY et al (2021) Modified global alignment and proportion scoring with body mass index and bone mineral density analysis in global alignment and proportion score of each 3 categories for predicting mechanical complications after adult spinal deformity surgery. Neurospine 18:484–491. https://doi.org/10.14245/ns.2142470.235
28. Ohba T, Ebata S, Oba H et al (2019) Predictors of poor global alignment and proportion score after surgery for adult spinal deformity. Spine 44:E1136–E1143. https://doi.org/10.1097/BRS.0000000000003086
29. Pizones J, Moreno-Manzanaro L, Sánchez Pérez-Grueso FJ et al (2019) Restoring the ideal Roussouly sagittal profile in adult scoliosis surgery decreases the risk of mechanical complications. Eur Spine J 14:925. https://doi.org/10.1007/s00586-019-06176-x
30. Steyerberg EW, Vickers AJ, Cook NR et al (2010) Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21:128–138. https://doi.org/10.1097/EDE.0b013e3181c30fb2
31. Ignasiak D, Behm P, Mannion AF et al (2023) Association between sagittal alignment and loads at the adjacent segment in the fused spine: a combined clinical and musculoskeletal modeling study of 205 patients with adult spinal deformity. Eur Spine J 32:571–583. https://doi.org/10.1007/s00586-022-07477-4
32. Katsu M, Ohba T, Ebata S et al (2020) Potential role of paraspinal musculature in the maintenance of spinopelvic alignment in patients with adult spinal deformities. Clin Spine Surg 33:E76–E80. https://doi.org/10.1097/BSD.0000000000000862
33. Miller EK, Neuman BJ, Jain A et al (2017) An assessment of frailty as a tool for risk stratification in adult spinal deformity surgery. Neurosurg Focus 43:E3. https://doi.org/10.3171/2017.10.FOCUS17472
34. Safaee MM, Ames CP (2023) Biological age and spine surgery-genomics and the future of risk assessment. World Neurosurg 172:47. https://doi.org/10.1016/j.wneu.2023.02.011