

# Inter- and intra-rater reproducibility of the Roussouly Classification System in patients with Adult Spinal Deformity

Tanvir J Bari<sup>1</sup>, Dennis W Hallager<sup>1</sup>, Niklas Tøndevold<sup>1</sup>, Ture Karbo<sup>1</sup>, Lars V Hansen<sup>1</sup>, Benny Dahl<sup>2</sup>, Martin Gehrchen<sup>1</sup>

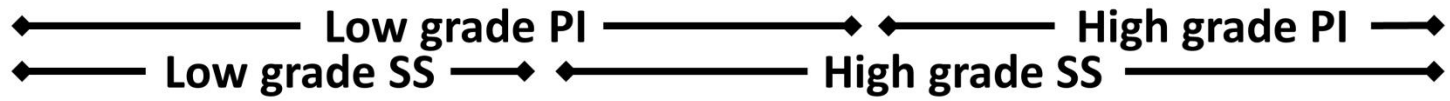
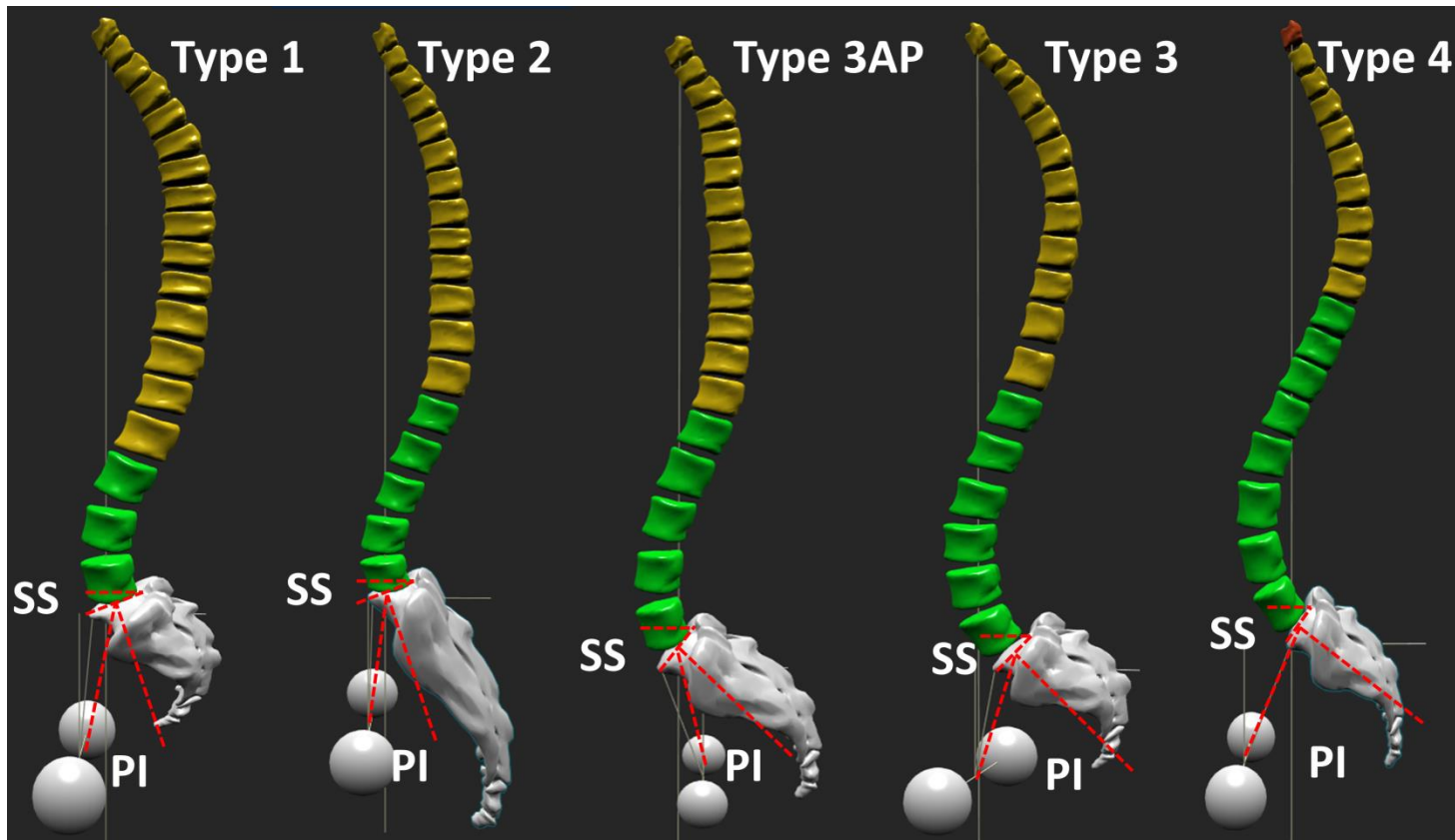
*<sup>1</sup>Spine Unit, Department of Orthopedic Surgery, Rigshospitalet, Copenhagen University Hospital, Denmark*

*<sup>2</sup>Department of Orthopedics and Scoliosis Surgery, Texas Children's Hospital and Baylor College of Medicine, TX, USA*

*Correspondence: [bari.tanvir@gmail.com](mailto:bari.tanvir@gmail.com)*

# Introduction

- The Roussouly Classification System was developed to describe the variation in sagittal spine shape in normal individuals.
- The sagittal profile of the spine was found to follow the orientation of the pelvis.
- As sacral slope increases so does the lower arch of lordosis as well as the global lordosis.



**Figure 1.** The 5 main spine types characterized by the degree of Pelvic Incidence (PI), Sacral Slope (SS), and number of lordotic vertebrae. *AP; Anteverted Pelvis*

# Table 1. Definition of Spine Types

Roussouly type	First criterion	Second criterion
Type 1	$SS < 35$	$\leq 3$ Lordotic vertebrates
Type 2	$SS < 35$	$> 3$ Lordotic vertebrates
Type 3	$35 \leq SS < 45$	
Type 3 AP	$35 \leq SS < 45$	$PI < 50^\circ$ or $PT < 5^\circ$
Type 4	$SS \geq 45$	
Kyphosis	No lumbar lordosis	No lumbar lordosis

SS: Sacral slope; PI: Pelvic Incidence, PT: Pelvic Tilt; AP: Anteverted Pelvis.

- To our knowledge, no previous study has investigated the reproducibility of this classification system.
- The aim of the current study was to provide estimates of the inter- and intra-rater reproducibility of the Roussouly Classification System in a population of patients with Adult Spinal Deformity (ASD).

## Methods

- We performed a reproducibility study in a consecutive, single-center cohort of patients referred for a condition of ASD.
- Patients with a history of previous spine instrumentation, spinal fracture or neuromuscular disease were excluded.
- A total of 64 patients were included.
- Two sessions of ratings were performed by four spine surgeons in a blinded fashion with a 14-day delay between ratings.

## Statistics

- A power calculation was performed prior to the study.
- Fleiss' Kappa coefficients were used to assess reliability as the number of raters were  $>2$ .
- Reliability were classified as slight ( $\leq 0.20$ ), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80) or almost perfect ( $\geq 0.81$ ).
- The Bhapkar test for marginal homogeneity was used to analyze differences between any two raters.

## Results

- The mean inter-rater reliability was moderate ( $\kappa = 0.60$ ) and on both readings all 4 raters agreed on the same Rousouly type in 47% cases.
- Mean intra-rater reliability was substantial,  $\kappa = 0.68$  (range: 0.57 to 0.78).
- Raters averagely assigned the same type on both readings in 76% of the cases (range: 67% to 83%).
- The two most experienced raters had the highest coefficients of inter-rater reliability ( $\kappa = 0.67$ ) and the highest degree of crude agreement (75%).



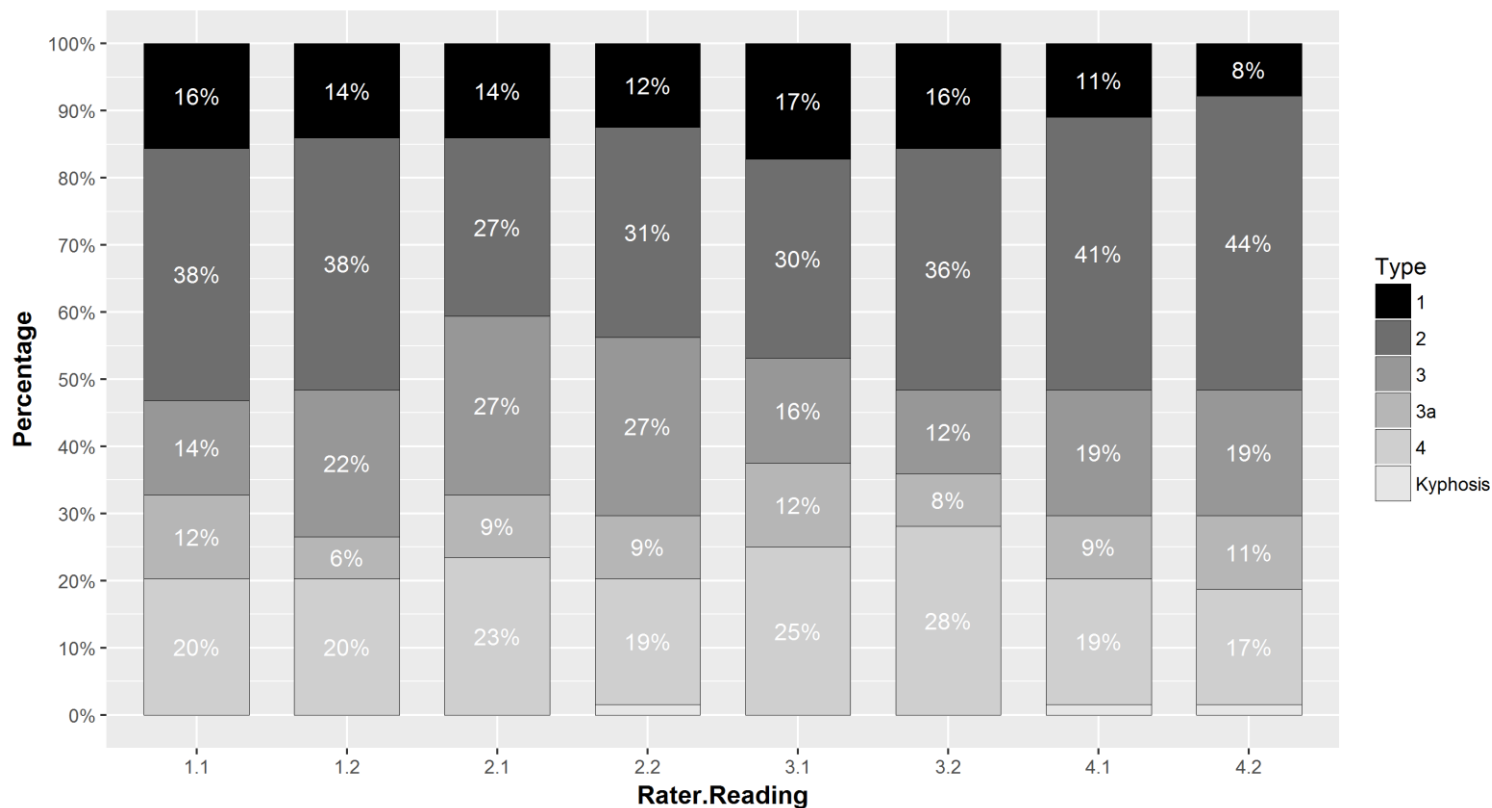
**Table 3. Roussouly Type inter- and intra-rater reliability**

	Fleiss kappa coefficients (95% CI)	*Interpretation of coefficient	Agreement
<b>Inter-rater</b>			
<b>Reading 1</b>	0.59 (0.49–0.70)	Moderate	47%
<b>Reading 2</b>	0.60 (0.50–0.69)	Moderate	47%
<b>Both readings</b>	0.60 (0.53–0.66)	Moderate	47%
<b>Intra-rater</b>			
<b>Rater 1</b>	0.57 (0.41–0.72)	Moderate	67%**
<b>Rater 2</b>	0.68 (0.54–0.82)	Substantial	75%**
<b>Rater 3</b>	0.78 (0.65–0.90)	Substantial	83%**
<b>Rater 4</b>	0.70 (0.56–0.85)	Substantial	78%**
<b>All raters</b>	0.68 (0.61–0.75)	Substantial	76%**

\*  $\leq 0.20$ : slight; 0.21-0.40: fair; 0.41-0.60: moderate; 0.61-0.80; substantial;  $\geq 0.81$ : almost perfect.

\*\* Percentage of cases where all 4 raters assigned the same Roussouly Type.

### Distribution of Roussouly type across 4 raters - Both readings



**Figure 2.** Distribution of Roussouly type by rater and reading. There was a significant difference in distribution of Roussouly types between Raters 3 and 4 ( $p=0.034$  in reading 1 and  $p=0.027$  in reading 2).

## Conclusion

- In this study of the Roussouly Classification System we present moderate inter-rater and substantial intra-rater reproducibility in a single-center cohort of 64 ASD patients.
- There was a trend for higher levels of reproducibility among more experienced raters.
- We conclude the reproducibility of this system to be acceptable and comparable to other ASD classification systems

# Conflicts of interest

## Funding sources:

*MG and BD received an institutional grant from K2M outside the submitted work.*

*MG and BD received an institutional grant from Medtronic outside the submitted work.*

*For the remaining authors, none were declared.*

## Conflicts of interest:

*No financial or competing interests in relation to this work.*