

Radiographic Clustering Enhances Subgroup Prediction of Mechanical Complications in Adult Spinal Deformity

Ibrahim Obeid (Bordeaux, France), Alice Baroncini (Milano, Italie), Cécile Roscop (Bordeaux, France), (Essg) European Spine Study Group (Barcelona, Espagne), Ferran Pellisé (Barcelona, Espagne), Javier Pizones (Madrid, Espagne), Frank Kleinstück (Zurich, Suisse), Ahmet Alanay (Istanbul, Turquie), Philippe Charles (Strasbourg, France), Anouar Bourghli (Riyadh, Arabie Saoudite), Daniel Larrieu (Bordeaux, France), Louis Boissiere (Pessac, France)

Introduction

Global risk models for mechanical complications after adult spinal deformity (ASD) correction may be limited by patient heterogeneity. Clustering may bridge this gap by enabling phenotype-specific prediction. However, it is unclear how the choice of variables used for clustering influences the performance of the prediction models.

Materials and Methods

Retrospective analysis of prospectively collected data. 723 ASD patients with complete preoperative data and ≥2-year follow-up were included. Predictors spanned demographics, surgical variables, PROMs, and radiographic parameters. Mechanical complications were predicted using supervised models with an 80/20 split and 5-fold cross-validation. Clustering strategies varied by feature emphasis and number of groups: clinical-heavy C16-3 (3 clusters), radiographic-heavy C12-3 and C12-7 (3 and 7 clusters, respectively), and SHAP-based supervised clustering.

Results

The best global model achieved AUC=0.77 (ElasticNet/random forest). Clustering did not improve overall performance, but radiographic-heavy C12-7 identified clinically relevant subgroups with equal or higher discrimination and/or higher sensitivity than the global model: younger clusters (AUC 0.76–0.82) and a severe sagittal and coronal imbalance cluster (AUC=0.75). Clinically-heavy C16-3 improved prediction in one subgroup (AUC=0.76), while most other clusters across approaches had AUC ≈0.70. SHAP-based supervised clustering showed comparable performance (AUC 0.70–0.72).

Table caption: Predictive performance for the best prediction model, for each clustering method.

Conclusion

This is the first study to evaluate the impact of clustering and models on the performance of a prediction. In ASD, the choice of clustering variables matters: radiographic-heavy clustering best supported the prediction of a radiographic-driven outcome and revealed transitions in risk across phenotypes. These findings provide a practical bridge from AI to subgroup-tailored complication risk stratification and follow-up planning.

Type of Clustering	Cluster	Split Method	Best Method	Metric Type	Accuracy	AUC	Sensitivity	Specificity
Complete dataset	-	80/20	Random Forest	-	0,73	0,776	0,369	0,893
C12_3	1	Random Forest	5-fold CV	AUC	0,709	0,674	0,283	0,893
	2	Random Forest	80/20	AUC	0,654	0,601	0,498	0,704
	3	ElasticNet	5-fold CV	AUC	0,701	0,924	0	0,98
C16_3	1	LDA	80/20	AUC	0,762	0,927	0,019	0,972
	2	Random Forest	80/20	AUC	0,693	0,631	0,499	0,752
	3	Random Forest	5-fold CV	AUC	0,694	0,668	0,231	0,913
C12_7	1	Random Forest	5-fold CV	AUC	0,754	0,617	0,615	0,678
	2	Random Forest	80/20	AUC	0,688	0,627	0,574	0,69
	3	Random Forest	80/20	AUC	0,717	0,646	0,489	0,781
	4	KNN	80/20	AUC	0,778	0,852	0,412	0,885
	5	Logistic Regression	80/20	AUC	0,754	0,706	0,294	0,897
	6	Random Forest	5-fold CV	AUC	0,817	0,734	0,263	0,962
	7	Survival Analysis	80/20	C-index	0,681	-	-	-