Introduction Unsupervised machine learning (ML) hierarchical clustering identifies 3 homogeneous groups of Adult Spinal Deformity (ASD) patients. Mean values for variables describing ASD patients and their outcomes might not reflect these known ASD subgroups. We aimed to analyze differences in treatment outcomes and PROM effect sizes (ES) amongst clusters and to identify how often overall group mean values were similar to cluster mean values. Materials and Methods Surgical patients with >2YFU were selected from a multicentre ASD database. ML applied to baseline demographic, radiological and PROMs variables grouped patients into 3 clusters: Young Coronal (YC) (age<47, coronal deformity, low disability); Old Moderate (OM) (age≥47, moderate coronal and/or sagittal deformity and disability); Old Severe (OS) (age≥47, severe coronal and/or sagittal deformity and disability). We compared surgical parameters (revision surgery, blood loss, pelvic fixation, 3-colum osteotomies [3CO]), 2YFU & 5YFU PROMs (ODI, SF36, SRS22) scores, change and effect sizes (ES), and rate of adverse events (major complications [MC] and unplanned reinterventions, [UR]) between the clusters using ANOVA and pairwise t-tests and z-tests. Results Of 429 patients operated before 2015, 372 had 2YFU data and were included in the analysis (70.2% had 5YFU data): 104 YC, 205 OM and 63 OS. All baseline values, surgical invasiveness parameters and rates of adverse events (MC, UR) differed ($p<0.05$) amongst clusters, increasing (or decreasing; Major Cobb) progressively from YC to OM to OS (Table). All PROMS differed ($p<0.05$) between the 3 clusters at 2YFU & 5YFU. The quality-of-life gain measured with SRS-22 self-image, satisfaction, mental health, sub-total and SF-36 MCS was similar in all clusters. SRS-22 self-image and subtotal scores showed the largest ES (0.81-1.28). For the other PROMs, 2YFU & 5YFU only showed similar gain in OM and OS for ODI, SRS-22 function and pain (Figure). Overall, total ASD group mean values did not reflect the more severe surgical invasiveness, adverse events and quality of life status of OS patients or the less severe status of YC adults. Conclusions ASD heterogeneity negates the use of overall mean values that do not respect clusters; they clearly do not represent a substantial proportion of ASD patients. Our results argue for PROM instruments that are more responsive in each cluster and cluster-specific benchmarking.

|  |  | ALL (372) | YC (104) | OM (205) | OS (63) | Pvalue (YCvsOMvsOS) |
|---|---|---|---|---|---|---|
| Baseline Characteristics | Age | 52.4 (SD 19.4) | 30.0 (SD 12.0) | 60.4 (SD 14.9) | 63.2 (SD 10.4) | <0.01 |
| | Major Cobb | 42.5 (SD 21.3) | 58.2 (SD 20.0) | 37.9 (SD 17.9) | 31.7 (SD 20.0) | <0.01 |
| | Global Tilt | 26.29 (SD 17.7) | 11.31 (SD 9.6) | 29.71 (SD 17.3) | 38.32 (SD 13.4) | <0.01 |
| | SVA | 41.67 (SD 63.0) | -4.12 (SD 31.1) | 48.71 (SD 60.9) | 93.82 (SD 59.0) | <0.01 |
| | ODI | 37.3 (SD 19.8) | 16.1 (SD 11.3) | 40.3 (SD 14.0) | 61.8 (SD 11.2) | <0.01 |
| Surgical Characteristics | Bleeding (ml) | 1447.6 (SD 1262.7) | 1152.4 (SD 966.8) | 1425.9 (SD 1326.0) | 1875.3 (SD 1275.2) | <0.01 |
| | Pelvic Fixation | 36.0% | 5.7% | 43.9% | 60.3% | <0.05 |
| | 3 column osteotomies | 14.5% | 4.8% | 15.1% | 28.6% | <0.05 |
| Adverse Events 5YFU | Patients with Major Complications | 30.4% | 14.4% | 32.7% | 49.2% | <0.05 |
| | Patients with Reinterventions | 28.2% | 12.5% | 30.7% | 46.0% | <0.05 |
| Quality of Life | | | | | | |
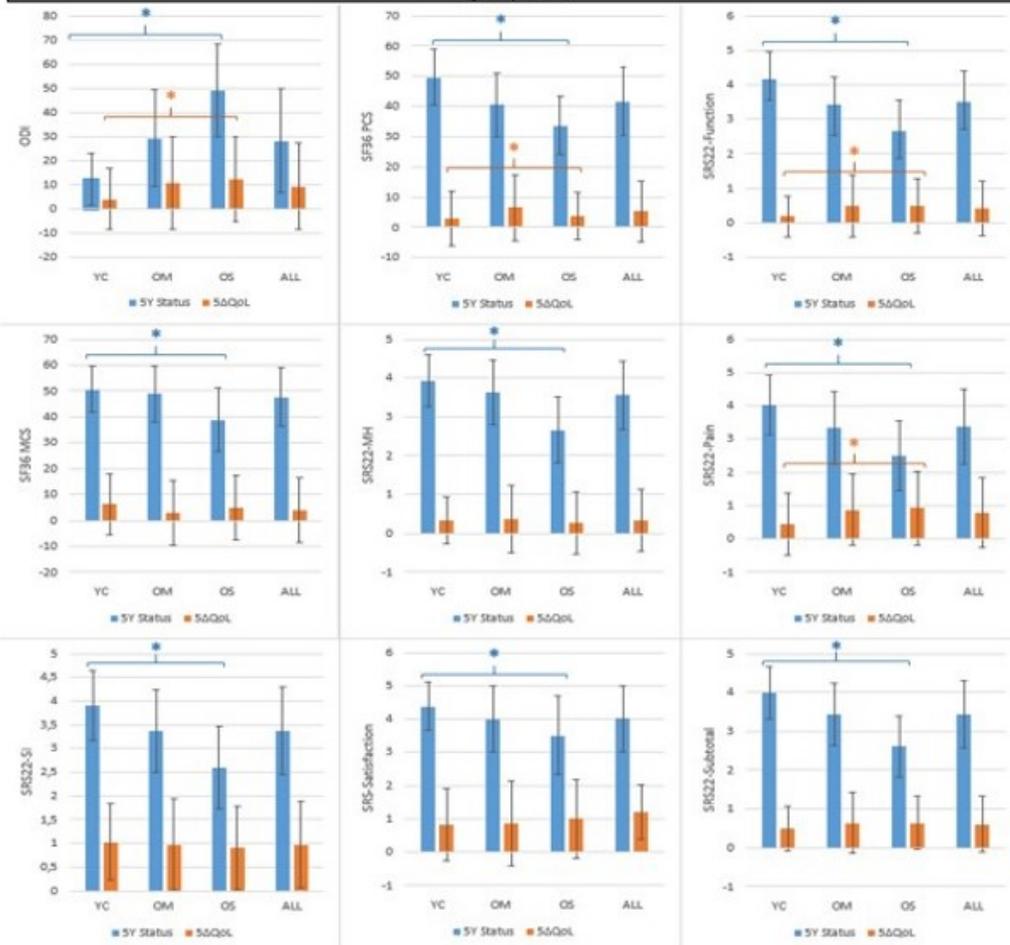


*<0.05 among clusters (YC, OM, OS) using ANOVA. YC, Young Coronal; OM, Old Moderate; OS, Old Severe; SD, standard deviation; SVA, Sagittal Vertical Axis.