



Evaluation of an Artificial Intelligence (AI) Based Scoliosis Measurement Program

Miao Wang¹, Ming Sun¹, Haisheng Li¹, Yukang Yang², Peter Helmig¹, Kestutis Valancius¹, Tianyu Liu², Kresten W. Richers¹, Kristian Høy¹, Thomas Bender¹, Cody Bunger¹, Yu Wang³, Ebbe S. Hansen¹

¹Dept. Orthopaedic surgery, Aarhus University Hospital, Denmark

²Department of Automation, Tsinghua University, China

³Dept. Orthopaedic surgery, Beijing University Hospital, China

Background

Scoliosis refers to the three-dimensional deformity of the spine with segments of the spine bent laterally with vertebral rotation. Reliable measurement of spinal curve is crucial for determining therapeutic decision for scoliosis patients. Cobb Angle is the gold standard, but it is an objective measurement with variant from surgeon to surgeon. A solid and reliable measurement tool is needed. Artificial Intelligence has showed great potential in image measurement. Aim. To compare the performance of an AI based scoliosis measurement tool with senior scoliosis surgeons in Denmark

Methods

Trained the AI algorithm with 650 scoliosis X-ray images by using Convolutional Neural Network (CNN). Another 100 scoliosis X-ray have been assigned into two groups randomly. Each group has been measured by AI and senior spine surgeons separately. Surgeons measured Cobb angles twice with minimal 3 weeks interval. Intraclass correlation coefficients (ICC) were used to determine the interobserver and intraobserver reliabilities. The correlation of scoliosis curve angle measurements has been tested with Pearson correlation coefficient and the mean absolute error.

Results

ICC is 0.96 in group 1 and 0.90 in group 2, which means excellent reliability. Pearson Correlation coefficient was 0.956 in group 1 and 0.930 in group 2. Spearman rank-order correlation was 0.960 ($p < 0.001$) in group 1 and 0.900 ($p < 0.001$) in group 2. The absolute error between AI

and surgeons are $3.5^{\circ} \pm 3.1^{\circ}$ in group 1 and $5.0^{\circ} \pm 3.8^{\circ}$ in group 2. In total the absolute error is $4.2^{\circ} \pm 3.3^{\circ}$. In 67% of all cases, there were only 0° - 5° different between AI program and spine surgeons.

Conclusions

There is statistic correlation of Cobb angle measurement between our new developed AI program and senior spine surgeons. The reliability is statistic excellent in both patients' groups. Our new AI program can provide reliable Cobb angle measurement as good as senior spine surgeons.