Hip-Spine Classification Automation and Population Analysis Using Deep Learning

Robert Gabriel Ricotti, Seong Jang[1], Theofilos Karasavvidis, Cale Pagan, Jonathan M Vigdorchik[2], Peter Keyes Sculco, David Jacob Mayman, Seth A Jerabek[2]
[1]Weill Cornell, [2]Hospital For Special Surgery

INTRODUCTION: Concurrent spinal pathology often exists in the patient population undergoing total hip arthroplasty (THA). The hip-spine relationship influences lumbar lordosis (LL), acetabular anteversion, and pelvic rotation as one changes from a standing to a seated position. Spinal deformity, age, sex, and BMI can influence these parameters, and it is important to understand how they interact when planning for acetabular component placement in THA. The hip-spine classification was developed to better characterize spinopelvic mobility and guide operative management of THA patients. The objectives of this study were to 1) characterize the hip-spine classifications of 1,622 patients at a single institution, 2) utilize a novel deep learning (DL) method to provide validated measurements of spinopelvic parameters and mobility in the same patient cohort, and 3) analyze differences in hip-spine classification distributions based on age, sex, and BMI.

METHODS: A multi-modal DL workflow was developed to automatically measure spinopelvic tilt (SPT), sacral slope (SS), pelvic incidence (PI), and LL on preoperative standing and seated lateral spinopelvic radiographs in 1,622 patients who underwent THA at our institution. The final DL workflow's measurement accuracy was assessed using a hold-out set of 200 testing images (100 standing:100 sitting). These DL measurements were compared against two trained, blinded readers using inter-class correlation coefficients (two-way 147 mixed, single score, ICC3) and Bland-Altman plot analyses to assess bias. ICCs > 0.9 were defined as excellent, and the DL workflow was applied to the entire patient cohort to determine hip-spine classifications. The patients were then categorized into one of four hip-spine classifications: 1A) normal spinal alignment and mobility, 1B) normal spinal alignment and stiff spine, 2A) flatback deformity and normal mobility, and 2B) flatback deformity and stiff spine. Analyses were performed to investigate the relationship between age, sex, BMI, and preoperative hip-spine classification.

RESULTS: Mean age was 63.2 ± 11.5 years (range 13 to 83), mean BMI was 28.8 ± 5.7 (range 16.6 to 52.1), and 887 (54.9%) patients were female. The DL workflow demonstrated excellent accuracy in measuring SPT (ICC = 0.98) and PI-LL (ICC = 0.95) when compared to 2 different readers **(Figure 1)**. The population breakdown of hip-spine classifications in our cohort of 1,622 patients as measured by the DL workflow can be found in **Figure 2**. Hip-spine classifications differed based on age (over 65 vs. under 65, $p < 0.0001$) and BMI (obese vs. non-obese, $p = 0.01$) **(Figure 3)**. There were no differences in hip-spine classifications based on sex ($p = 0.12$). Additional univariate analyses did not indicate that age and BMI were confounders.

DISCUSSION AND CONCLUSION: The DL workflow in this study provided accurate and valid measurements of SPT and PI-LL in our cohort of 1,622 patients, which is the largest series to date. Automation of the hip-spine classification streamlines preoperative planning for THA and helps the surgeon plan for changes in spinopelvic parameters following THA. Future studies are encouraged to use novel DL methodology to investigate automation of postoperative hip-spine classification measurements in relation to clinical outcomes.

Figure 1. Inter-class correlations of measurements of hip-spine parameters between two human readers and the deep learning workflow

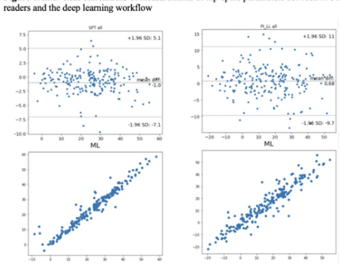
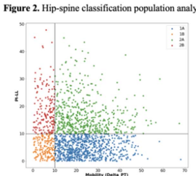


Figure 1. Inter-class correlation (ICC) between human readers and machine learning (ML). ICCs for measurement of spinopelvic tilt (SPT) = 0.98 (0.97-0.99); ICCs for measurement of pelvic incidence (PI) minus lumbar lordosis (LL) = 0.95 (0.93-0.97).

Figure 2. Hip-spine classification population analysis using deep learning workflow



Hip-Spine	% (n)
1A	46.1 (747)
1B	10.5 (170)
2A	32.6 (529)
2B	10.3 (167)

Figure 2. Hip-spine classifications based on measurements by DL workflow. The y-axis contains pelvic incidence (PI) minus lumbar lordosis (LL). The x-axis is the change in mobility, or spinopelvic tilt (delta PT).

Figure 3. Hip-spine classifications based on body mass index, sex, and age.

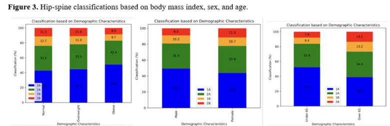


Figure 3. Hip-spine classifications based on demographics. Obese vs. non-obese groups, and those over 65 vs. under 65 years of age had statistically different hip-spine categorizations ($p = 0.01$ and $p < 0.001$, respectively). There were no significant differences in hip-spine classification between males and females ($p = 0.12$).