

Evaluation of Socioeconomic Status and Spine Health through a Deep Learning Based Image Analysis: Data from the UK Biobank

¹Mary A. Bucklin, ¹Ashrith Alavilli, ^{1,2}Rana A. Ahmad, ^{1,3}Scott J. Simmons, ¹John T. Martin

¹Rush University, Chicago, IL; ²University of Illinois Chicago, Chicago, IL; ³Drury University, Springfield, MO

Introduction: Chronic low back pain (cLBP) is a heterogeneous disease with biological, psychological, and social components. The interaction of these components is poorly understood¹. If we understand these complex interactions, we can improve individualized treatments. Social determinants of health play a primary role in many disease processes including spine disease. Individuals with lower socioeconomic status (SES) have increased LBP prevalence, severity, and interference with daily activities as well as longer periods of missed work due to LBP². Lower SES individuals also have worse access to care, and when they receive orthopedic care, experience more complications and worse outcomes³. In previous work, we screened hundreds of variables related to musculoskeletal health in the Osteoarthritis Initiative dataset and identified that income and education were the strongest predictors of LBP severity and chronicity⁴. However, we could not evaluate anatomical factors like intervertebral disc degeneration, which is frequently implicated as a driver of cLBP⁵, and has not been investigated in the context of SES. Our **hypothesis** is that low SES exacerbates disc degeneration and consequently CLBP. This hypothesis is supported by research in rheumatoid arthritis, where disease activity and tissue degeneration are more severe in individuals with low SES. Our long-term goal is to comprehensively evaluate which socioeconomic factors are most impactful in spine disease and LBP. Because there are many socioeconomic factors at the individual level (diet, exercise, health insurance status, medical mistrust, manual labor, etc.) and neighborhood level (housing quality, food swamp, personal safety, etc.) that contribute to health, a large-scale study is necessary to evaluate SES holistically^{6,7}. Here, we will utilize deep learning to automate the analysis of dual x-ray absorptiometry scans in the United Kingdom Biobank (UKB) dataset to enable a large-scale assessment of disc degeneration related to SES. We include a preliminary analysis here of a subset of the cohort (N=342) and performance of a deep learning model for image processing.

Methods: Study Population The UK Biobank (UKB) is a biomedical database that includes genetic and health information for 500,000 individuals from the UK. Demographic information includes standard age, sex, and other factors as well as detailed socioeconomic metrics such as income, employment, education, crime, barriers to housing and services, living environment as well as a composite measure of socioeconomic factors, the index of multiple deprivation (IMD). A subset of 50,000 participants additionally underwent lateral spine dual x-ray absorptiometry (DEXA) imaging. **Image Analysis** Lumbar vertebral bodies (T12 to L4) were manually segmented from DEXA scans to develop a training dataset for machine learning (N=342). Preliminary analysis demonstrated that a simple manual segmentation of the four corners of each vertebral body was more reliable and faster than other segmentation methods. A deep convolutional neural network was developed that receives a DEXA scan as input and outputs a quadrilateral that corresponds to the corners of 5 lumbar vertebral bodies. The model is a deep, fully convolutional, encoder-decoder network using DeepLabV3 with pre-trained backbone ResNet101. The model was trained via transfer learning on a subset of data (N=342 images) using an 90%/10% training/testing split on an NVIDIA A2000 GPU (learning rate = 0.5, epochs = 5, training time = 15:43, optimizer = Adadelta). **Statistical Analysis** To determine our preliminary model accuracy, we used the Intersection Over Union (IoU) metric, calculating an aggregate IoU for each spinal level (IoU = 1, perfect overlap; IoU = 0, no overlap). Following segmentation, we calculated disc height index⁸ (DHI), where DHI is equal to the disc height divided by the average height of the adjacent vertebral bodies. A mixed effects model of manually segmented data is presented to preliminary estimate a relationship between DHI and IMD, controlling for age.

Results: We confirmed a negative correlation between age and DHI ($\beta = -0.00114$, $p = 0.0004$) (Table 1), though no preliminary relationship between IMD and DHI could be detected. Our model predicts the quadrilaterals formed by vertebral corners in training and unseen test data (train IoU = 0.88 ± 0.02 , test IoU = 0.85 ± 0.01) where IoU is maximum at L1 (train 0.88, test 0.86), L2 (train 0.90, test 0.87), L3 (train 0.90, test 0.85), and L4 (train 0.92, test 0.86), and minimum at T12 (train 0.79, test 0.83) (Fig. 1).

Discussion: cLBP has been associated with socioeconomic status in many studies, however the mechanism is unclear. In this work, we seek to determine whether there is an anatomical driver of cLBP in lower SES populations by evaluating the relationship between SES and lumbar disc degeneration, an anatomical factor frequently implicated in cLBP. We developed a preliminary training dataset and deep learning platform to automatically measure DHI in lateral DEXA images towards segmenting 50,000 UK Biobank participants. Our early image processing results demonstrate reasonable success that could be improved by eliminating errors at T12 vertebrae. We confirm the reliability of our training dataset by demonstrating a well-known relationship between age and DHI. At this preliminary stage, however, we could not detect a relationship between IMD and DHI. Accordingly, we speculate that the psychosocial drivers of pain outweigh the anatomical drivers of pain in cLBP, which requires further confirmation.

Clinical Relevance: cLBP is the world's leading cause of disability and socioeconomic factors appear to play an important role. We preliminarily developed a computer vision framework to segment lumbar vertebral bodies in DEXA scans and confirmed age-related lumbar disc height narrowing indicative of degeneration. The preliminary results related to SES require further verification. Using a large-scale approach to analyzing SES, we can holistically evaluate socioeconomic factors and determine the most consequential in spine disease and back pain.

References: 1) Koes+ *BMJ* 2006; 2) Karan+ *Pain* 2020; 3) Cardinal+ *World NeuroSurg* 2022; 4) Huang *JOR Spine* 2022; 6) Brinjiki+ *AJNR* 2015; 7) Dey+ *RMD* 2022; 8) Martin+ *J Orthop Res* 2015; 9) Miller+ *Spine* 1988.

Acknowledgement: This work was funded by the National Institutes of Health (R00AR077685, T32AR073157).

Table 1 Linear mixed effects model. Fixed effects: index of multiple deprivation (IMD), age; random effect: subject; dependent variable: disc height index (DHI)

Independent Variable	Estimate	Standard Error	t Value	Pr > t
IMD	-0.00007	0.000195	-0.37	0.712
Age	-0.00114	0.00032	-3.55	0.0004

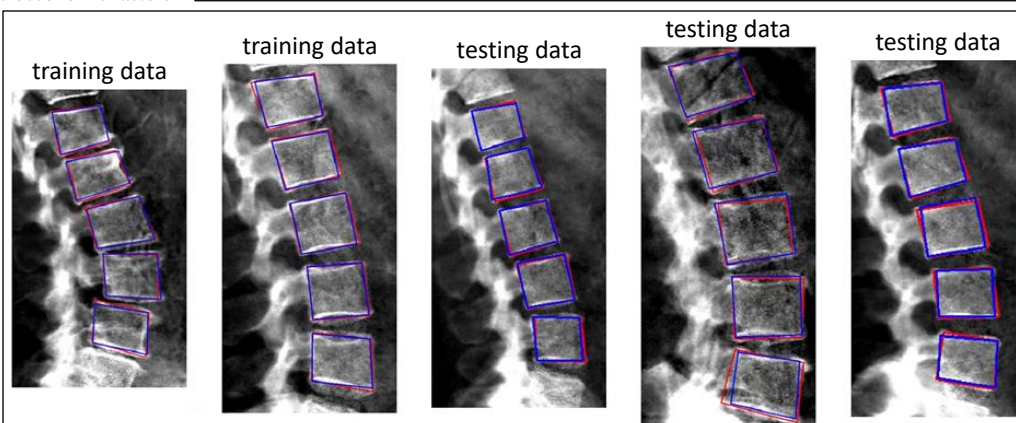


Figure 1 Examples of segmented T12 through L4 vertebral bodies in training and testing data (red quadrilaterals) and predictions using a deep convolutional neural network (blue quadrilaterals).