Utility Of Radiomics Features For Automated Clinical Grading Of Ultrasound Images Of The Thoracolumbar Fascia For Evaluation Of Nonspecific Low Back Pain

Georgina Flynn-Smith[1], Alicia Roldan, DO[2], Julianne Wood[1], Per Gunnar Brolinson, DO[3], David Redden, PhD[4], Connor Barrand[3], Valerie Daniels[3], Kevin White[3], Albert Kozar, DO[3], Vincent Wang, PhD[1]
[1]Virginia Polytechnic Institute and State University, Blacksburg, VA, [2]Montefiore Medical Center, Bronx, NY, [3]Edward Via College of Osteopathic Medicine, Blacksburg, VA, [4]Edward Via College of Osteopathic Medicine, Auburn, AL
gflynnsmith@vt.edu

Disclosures: G. Flynn-Smith: None. A. Roldan: None. J. Wood: None. P. Brolinson: 9; American Osteopathic Academy of Sports Medicine. D. Redden: None. C. Barrand: None. V. Daniels: None. K. White: None. A. Kozar: 9; ACGME Osteopathic Recognition Committee. V. Wang: 8; Connective Tissue Research.

INTRODUCTION: The Thoracolumbar Fascia (TLF), a key mechanical stabilizer of the trunk, has recently emerged as a primary source of low back pain previously categorized as nonspecific (*1*). Ultrasound (US) imaging is commonly used for point-of-care clinical assessment of the TLF. Clinical evaluation criteria for assessment of TLF images include morphological assessments (e.g. thickness measurements), neovascularity, and more recently, stiffness measured by shear wave elastography (SWE). While these evaluation criteria typically have moderate-to-high interrater reliability, the process of manually measuring and grading images is very time consuming. The emerging field of radiomics, which uses mathematical algorithms to extract quantitative features from medical images, presents a promising approach to increase the efficiency of clinical image interpretation. The purpose of this study was to evaluate the agreement between radiomics features and clinical grading/measurements of the TLF in US images to assess the potential suitability of these features for future development of an automated clinical grading method.

METHODS: The current study utilized 502 grayscale (GS), 1185 SWE, and 501 power Doppler (PD) de-identified US images of the TLF from men & women aged 18-88, with and without low back pain, obtained as part of a previous, IRB-approved study (*2*). A single trained clinician performed post-hoc measurement of fascial thickness (GS images) and shear wave speed (SWE images) using the US machine's (Supersonic Aixplorer MACH 30) on-board measurement tools. The same clinician graded neovascularity in PD images using the modified Öhberg scale (*3*). To extract radiomics features, images were first manually segmented by a single trained researcher using a custom MATLAB code. Blob analysis was performed on the resulting PD colormaps and pixelwise texture analysis was performed on the SWE maps, both using a custom MATLAB code. Shape-based features were extracted from GS images using pyRadiomics. Agreement between each of the 3 clinical grading criteria/measurements and its corresponding radiomics feature(s) was evaluated using Pearson correlation for continuous variables and Spearman rank correlation for ordinal variables. To evaluate the utility of these features for automated grading, root mean squared error (RMSE) of predicted clinical measurement and misclassification rate of neovascularity grade were calculated from linear and ordinal logistic regression models, respectively.

RESULTS SECTION: Mean shear wave speed (SWS) calculated by texture analysis and measured by the US machine demonstrated very high agreement (**Figure 1**), with a Pearson correlation coefficient (r) of 0.93 ($p<0.001$) and RMSE of 0.242 m/s. The shape-based feature MinorAxisLength demonstrated a high agreement with fascial thickness measures ($r = 0.79$; $p = p<0.001$; RMSE = 1.24mm). PD colormap total blob area and number of blobs both demonstrated high agreement with neovascularity grade, with Spearman's ρ values of 0.783 ($p<0.001$) and 0.737 ($p<0.001$), respectively. Logistic regression miscalculation rate for neovascularity grade based on total blob area (**Figure 2**) and number of blobs was 0.186 and 0.182, respectively.

DISCUSSION: We investigated the agreement between radiomics features and clinical grading/measurement of the TLF in GS, SWE and PD US images. All parameters exhibited a high correlation with the corresponding clinical grading/measurement. However, while the shear wave linear regression model had a prediction error of less than 10%, the prediction error for the neovascularity logistic regression models and the thickness linear regression models were substantially higher, at over 18% and 25%, respectively. Further research is needed to investigate additional image features or combinations of features that may result in more accurate predictions.

SIGNIFICANCE/CLINICAL RELEVANCE: This study is among the first to investigate the correlation between quantitative image features and clinical interpretation of US images of the TLF. Our results suggest that these radiomics features correlate strongly with the clinical image interpretation and constitute a promising approach for development of an automated grading technique to improve the efficiency of clinical workflow.

REFERENCES: 1. G. Casato, *et al.*, *Eur. J. Transl. Myol.* **29**, (2019). 2. A. Roldan *et al.*, *6th Int. Fascia Res. Congr,.* (2022). 3. L. Ohberg, H. Alfredson, *Br. J. Sports Med.* **36**, (2002).

ACKNOWLEDGMENTS: We would like to thank Zach Kozar and Gabriel Chester for their contributions to the code used in this study.

IMAGES AND TABLES:

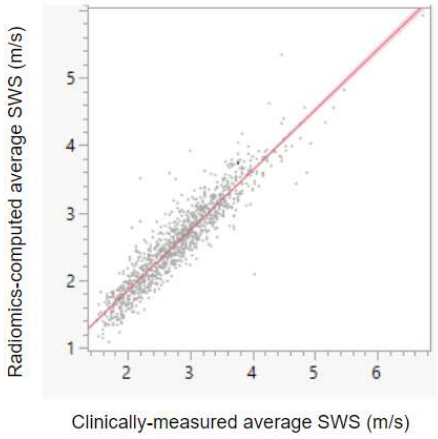

Figure 1: Correlation between clinically-measured mean SWS and radiomics-computed mean SWS

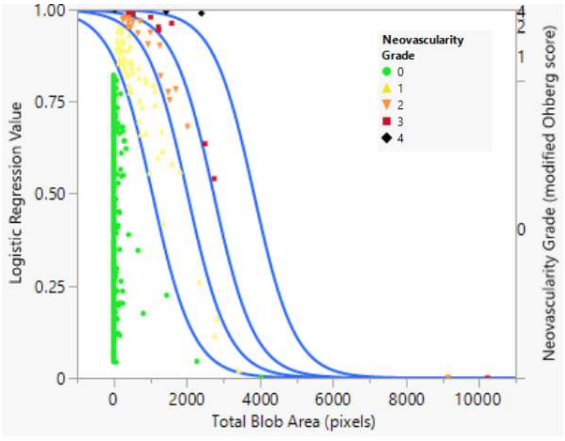

Figure 2: Ordinal logistic regression of neovascularity grade vs. total blob area of power Doppler colormap