Deep Learning-Based Predictions of Polyethylene Insert Wear in Total Knee Replacements

Mattia Perrone[1], Scott Simmons[2], Philip Malloy[1, 3], Catherine Yuh[1], John Martin[1], Steven P. Mell[1]
[1]Rush University Medical Center, Chicago, IL, [2]Drury University, Springfield, MO, [3]Arcadia University, Glenside, PA
Email of Presenting Author: mattia_x_perrone@rush.edu

Author Disclosure Information: M.P., S.S., P.M., C.Y., J.M., S.P.M (N)

INTRODUCTION: Knee Osteoarthritis (OA) is a leading cause of disability[1]. End-stage treatment for knee OA is a total knee replacement (TKR)[2,3]. Polyethylene (PE) wear remains a leading cause of long-term failure in TKRs, despite significant advances in materials, such as crosslinked PE[4]. Previous studies leveraging finite element analysis (FEA) have shown that variability of gait patterns can lead to significant variability in PE wear rates[5,6]. Our group previously developed a multiple linear regression surrogate model that predicts wear volume from motion waveforms by combining design of experiments with FEA[6]. However, this existing framework does not consider the spatial wear distribution on the tibial liner, neglecting potentially crucial parameters associated with TKR failure such as wear depth or position. In addition, while FEA has proven to be a useful computational tool for predicting PE wear in TKR, the development of FEA models can be both resource-intensive and time-consuming. Therefore, the aim of this study was to leverage our existing FEA wear predictions to train and develop a deep learning framework capable of predicting PE wear distribution starting from input kinematic and kinetic patterns, that could significantly lower both computational costs and processing time compared to that of traditional FEA models.

METHODS: In brief, a previously published method was used to generate 314 unique anterior/posterior translation, internal/external rotation, flexion/extension, and axial loading time series waveforms based on ISO14243-3(2014) – these unique waveforms were then used as inputs to a well-validated FEA model[7] to predict output linear wear distribution on the PE liner[6]. For the deep learning model presented in this study, the 314 kinematic and kinetic time series waveforms were used as input training data, while wear predictions from our FEA model were used as output training data. The input time series data were normalized to 100 points, resulting in a total shape of the input data of 314x100x4. The output distribution of PE wear previously calculated using FEA (Abaqus) was used as the training dataset for the deep learning model. To do this, wear values at the FEA model nodes, which are non-uniformly organized, were interpolated to generate uniform 2D images. Various pixel grid sizes (100x100, 150x150, 200x200) were used on the 2D images to verify that results were not dependent on the size of the grid we interpolated to. A deep learning model was then developed to predict linear wear, as an image, using time series gait patterns as inputs (**Figure 1**). The proposed model consists of a U-net architecture incorporating 1) a dual-head transformer encoder to discern temporal dependencies within the multivariate time series and 2) deconvolutional layers in the decoder to ensure the reconstruction of image data. The dataset was split into a training/validation/test set (60%-20%-20%), and the model was trained using mean squared error as loss function. Model performance was evaluated by comparing the deep learning and FEA model predictions (ground truth) using metrics such as mean absolute percentage error (MAPE) for relevant geometric features of the wear scar, such as length, width, and area, in the medial and lateral compartments. Structural similarity index measure (SSIM) and normalized mutual information (NMI) were utilized for broader pattern identification across ground truth and predicted wear maps.

RESULTS: For the medial and lateral regions: MAPE values remained under 5% and 8% respectively for the wear scar's width and length measurements and below 5% for the area of the wear scar (**Table 1.A**). SSIM and NMI were consistently above 0.88 for all grid size configurations (**Table 1B**). A representative comparison of the FEA output (Ground Truth) versus the deep learning output (Prediction) is shown in **Figure 2**.

DISCUSSION: This study presents a novel approach to rapidly predict PE wear in TKR, using deep learning to generate a surrogate model from FEA simulations. All variables can be acquired in a non-invasive manner using existing motion analysis techniques such as wearable sensors or cameras and musculoskeletal models. Moreover, the deep learning model significantly reduced the computational time for generating wear predictions compared to FEA, with the former inferring in seconds, and the latter requiring hours. Future work will include applying and evaluating this methodology on patient data (as opposed to generated waveforms), and performing validation/verification on the deep learning framework against gold standard knee simulator testing.

SIGNIFICANCE: The rapid prediction time of this deep learning framework may allow clinical applications, such as early detection of gait patterns that indicate higher risks for implant failure, providing a chance for early intervention and potentially extending the lifespan of implants based on patient motion.

REFERENCES: 1) Cross, M. et al. (2014) 2) Shichman, I. et al. (2023) 3) AJRR, AAOS, (2020) 4) Asher, D. P. et al. (2024) 5) Brockett, C. L. et al. (2016) 6) Mell, S. P. et al. (2020) 7) Mell, S. P. (Springer, 2020)

ACKNOWLEDGEMENTS: Thanks to Zimmer for providing CAD models of the NexGen Cruciate Retaining TKR.

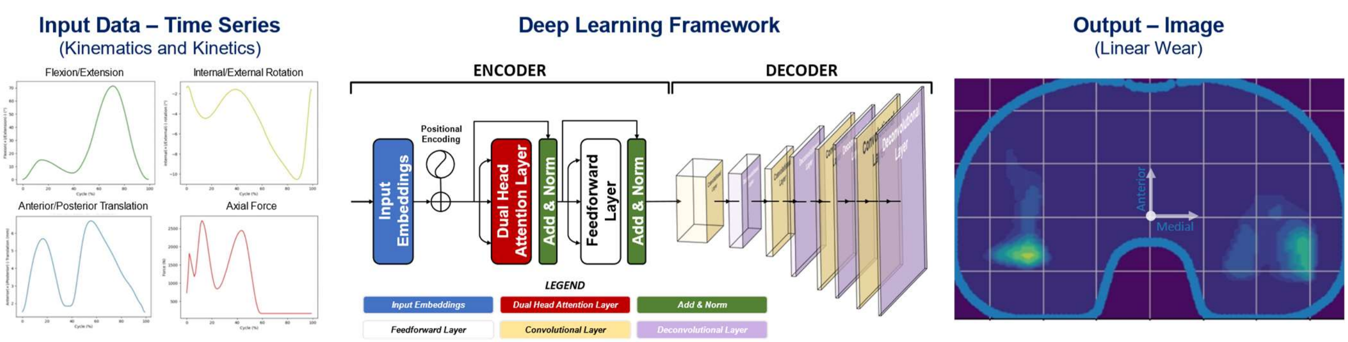

Figure 1. Input data, model architecture and output data

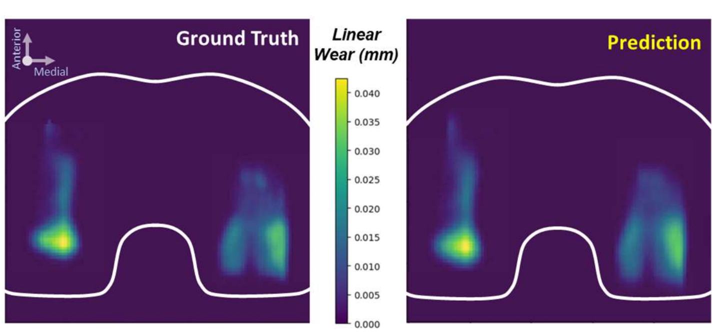

Figure 2. Comparison of ground truth/FEA output (left) and deep learning prediction (right) from a representative waveform/wear image pair.

Table 1. (A) MAPE values for wear scar width, length, and area **(B)** structural similarity index measure (SSIM) and normalized mutual information (NMI) across different grid. sizes.

A

Parameter	MAPE – 100x100	MAPE – 150x150	MAPE – 200x200
Width (Lateral)	7.69%	3.57%	5.58%
Length (Lateral)	5.28%	5.06%	4.69%
Area –(Lateral)	3.51%	4.05%	3.42%
Width (Medial)	1.93%	2.52%	1.43%
Length (Medial)	4.37%	4.31%	3.90%
Area – (Medial)	4.75%	4.61%	4.80%

B

Parameter	Grid size: 100x100	Grid size: 150x150	Grid size: 200x200
SSIM (lateral)	0.89 (0.88, 0.90)	0.90 (0.89, 0.91)	0.95 (0.94, 0.96)
SSIM (medial)	0.88 (0.87, 0.89)	0.94 (0.94, 0.95)	0.96 (0.95, 0.96)
NMI (lateral)	0.94 (0.94, 0.95)	0.95 (0.94, 0.95)	0.89 (0.88, 0.91)
NMI (medial)	0.96 (0.96, 0.97)	0.96 (0.95, 0.96)	0.96 (0.95, 0.96)