Predicting Functional Recovery Following Primary Total Knee Arthroplasty Using Generative Machine Learning

Thomas Harte, Huy Phan, Alison K Klika, Angelica Parra, David Joseph Jacofsky, Wael K Barsoum

INTRODUCTION:

Total Knee Arthroplasty (TKA) is one of the most effective elective surgeries available, offering significant pain relief and improved function for patients with severe knee osteoarthritis. However, as patient expectations evolve and enabling technologies advance, the adequacy of current outcome assessment tools, particularly Patient-Reported Outcome Measures (PROMs), is increasingly questioned. While PROMs are widely endorsed by regulatory bodies and, increasingly, payers, to measure a successful outcome, they exhibit substantial limitations such as ceiling effects among high-functioning patients.

Furthermore, in pre-operative scenarios, PROMs provide only a subjective snap shot in time, providing limited insight into a patient's disease progression or likely outcome following intervention. With the exponential growth of novel datasets from wearables and smartphones, paired with advances in computer science and artificial intelligence, we propose that such technologies can provide clinicians with better insights into a patient's presenting complaint as well as a predictive tool for their eventual outcome, providing a valuable tool in shared decision making and setting accurate patient expectations for TKA outcomes.

The purpose of this study was to use longitudinal (i.e. 365 d) preoperative daily step count data to predict postoperative step recovery curves for individual patients, a generative machine learning algorithm called Gaussian Process Latent Variable Model (GPLVM).

METHODS:

We passively collected both pre- and postoperative daily step data from 1,132 TKA patients' mobile and wearable devices via Apple Health and Google Fit using a patient engagement app (myrecovery, London, UK). From the 1,132 patients, we randomly selected activity records from 911 patients to train the GPLVM, using 365 days of preoperative data and 180 days of postoperative data to train the model. The dataset was first pre-processed with 13-day moving average.

We then tested the trained GPLVM using the preoperative steps (365 d prior to TKA) to predict postoperative steps (180 d post-TKA) for the remaining 228 patients to generate a predicted recovery curve. Accuracy was reported as the mean absolute error (MAE) between ground-truth step counts and predictive step counts.

RESULTS:

The 1,132 TKA patients had their surgery in the United States, and had an average age of 69.2 years (±9.2 SD). 75% of preoperative (365 d) and postoperative (180 d) daily step counts were available, and the remaining step data were imputed using mean values from 5 nearest neighbours.

The GPLVM was able to predict a patient's steps within an average accuracy (MAE) of 178 steps (± 90), or 5.6% of a patient's mean daily steps (Figures 1-3). As the size and scope of our training data set increases, the accuracy will significantly improve.

DISCUSSION AND CONCLUSION:

This study demonstrates the capabilities of generative AI in modelling patient recovery across a cohort of TKA patients. We believe that this is a valuable tool for the preoperative assessment of a patient and supports shared decision making between the surgeon and patient. These data would facilitate conversations around making a decision on if and/or when to treat, and setting realistic data driven expectations regarding a post operative recovery.

Furthermore, novel data sets such as activity paired with advanced tools such as generative ML offer the potential complement PROMs in a meaningful way, by addressing the limitations inherent to those measures.

FIG 1: Step prediction on a 81 Year old Female

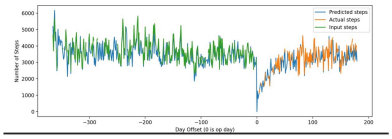

FIG 2: Step prediction for 68 year old male. In this case, 60 days of pre-op steps were omitted from the input to demonstrate the predictive capability of the model.

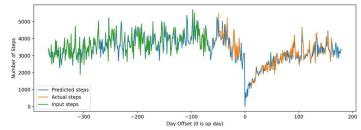

FIG 3: Step prediction for 57 year old male