# Generalized Osteoarthritis Prediction Model For Healthy Knees: Data From The Osteoarthritis Initiative

Mikael J. Turunen[1], Mika E. Mononen[1]
[1]University of Eastern Finland, Kuopio, Finland
mikael.turunen@uef.fi

**Disclosures:** M. J. Turunen (4-Aikoa Technologies Oy), M. E. Mononen (4-Aikoa Technologies Oy)

**INTRODUCTION**: In recent years, finite element analysis (FEA) has emerged as a promising tool for quantitative assessment of the individual risk of knee osteoarthritis (KOA) onset and progression [1]. Concurrently, a multitude of machine learning algorithms have been devised for predicting KOA [2]. The main challenge of FEA are the requirements of computational power and time for predictions of the personalized KOA risk. This is not practical to be usable as a clinical tool to assess the effects of different interventions such as gait alteration, weight loss, or surgical operation. The downside of ML models is the wide range of predictor variables that are difficult to quantify and their sensitivity to the target variable: a small change in the variables can lead to radically different risk classification results, while the small change in FEA results in a moderate change in the risk prognosis. In this work, we combine the best of both approaches by training an ML model to predict FEA model results for the risk of KOA. The aim of this study was to develop a Gaussian process regression (GPR) model for a generalized KOA risk in a population without need of clinical imaging data. Therefore, to exclude the need for the clinical imaging data, another GPR model was developed to estimate knee geometry and cartilage thickness based on general subject characteristics.

**METHODS:** To obtain quantitative predictions for personalized risk for the onset and progression of KOA, we conducted atlas-based FEA simulations [1,3] on 2092 radiographically healthy knees at baseline from the Osteoarthritis Initiative database (OAI, http://nda.nih.gov/oai). The final FEA simulation output was the volumetric cartilage degeneration, i.e., the development of KOA [1,3], which we used as the target variable in training the GPR model for KOA risk ($GPR_{KOA}$, Fig 1). We generated the atlas-based FEA models based on age, weight, and 5 anatomical dimensions of the knee joint [1,3], and used these variables also as input (predictor values) in training the $GPR_{KOA}$. The dimensions (measured from MRI scans) included medial and lateral joint spaces ($JS_{MED}$, $JS_{LAT}$), medial and lateral maximum condyle anterior posterior dimensions ($AP_{MED}$, $AP_{LAT}$), and distal femur medial-lateral size. To omit a need for data from clinical imaging, we trained additional GPR models to predict knee joint dimensions ($GPR_{DIM}$) measurable from anterior-posterior X-rays: lateral and medial joint spaces ($JS_{LAT}$, $JS_{MED}$) and the full medial-lateral width of the distal femur ($W$, Fig 1). $GPR_{DIM}$ was trained using 2024 healthy knees (age 45-67 years) at baseline (1669 for training, 355 for validation, stratified split) from the OAI database with using subject age, weight, height, and gender as a set of predictor variables. All dimensions were measured using an in-house Matlab GUIs (v. R2022b, MathWorks Inc.). As the $GPR_{KOA}$ model requires more input parameters than the output of the $GPR_{DIM}$ models provide, in the workflow we implemented the calculation of $AP_{MED}$ and $AP_{LAT}$ by assuming linear associations (Fig 1) based on average fraction of these dimensions (N>2000 knees). To note, we used X-ray-based (not MRI-based) measurements in the $GPR_{DIM}$ training since we aimed to evaluate the prediction performance based on dimensions measurable from single radiographs (no anterior-posterior dimensional data) as radiographs are taken much more frequently than MRI in clinical evaluation. Additionally, this also maximized the amount training data for $GPR_{KOA}$ models.

*Statistics:* We grouped the knees based on 8-year follow-up Kellgren-Lawrence (KL) grades [4] into 4 groups: KL0, KL1, KL2, and KL3&4. We evaluated the accuracy of the $GPR_{DIM}$ models using Pearson correlation and Root Mean Squared Error (RMSE). We used the non-parametric Mann-Whitney U-test to evaluate KL-group-wise differences in the predicted degenerations in the validation data (N=355) based the $GPR_{KOA}$ and atlas-based FEA workflows and the non-parametric Wilcoxon signed rank test to evaluate differences within each KL group between the $GPR_{KOA}$ and FEA workflows predictions. For all statistical tests, we considered $p < 0.05$ as the level of significance. We used ROC (receiver operating characteristics) curve and AUC (area under the curve) to evaluate the classification performance between healthy (by combining KL0 and KL1 groups) and osteoarthritic (KL3&4).

**RESULTS:** The correlation of the $GPR_{DIM}$ models was weakest for $JS_{MED}$ and highest for $W$ (Table 1). The $GPR_{KOA}$ and FEA-based workflows provided similar predictions of cartilage degeneration (Fig 2) in the KL0, KL1, and KL3&4 groups. A significant difference was observed in the KL2 group. Both workflows showed similar differences between the groups (Fig 2): KL0 vs KL2, KL0 vs KL3&4, and KL1 vs KL3&4. The classification performance between healthy (KL0&1) and osteoarthritic (KL3&4) was practically the same with both workflows (FEA: AUC=0.74, $GPR_{KOA}$: AUC = 0.73).

**DISCUSSION:** The unexpectedly high prediction accuracy achieved by the $GPR_{DIM}$ models in estimating knee dimensions is noteworthy. The findings suggest that in addition to knee size, cartilage thickness (in a healthy knee) can be estimated with good accuracy based on only age, weight, height, and gender. Similar relations have also been published earlier [5,6]. When comparing the prediction accuracies for KOA progression between the atlas-based FEA workflow and the $GPR_{KOA}$ model, strikingly similar results were observed, as indicated by the analogous AUC values. This parity in accuracy is intriguing, given that the $GPR_{KOA}$ model achieves such predictive power without necessitating the inclusion of clinical image data from subjects, unlike the FEA workflow which heavily relies on subject-specific information. In parallel, existing studies have reported comparable accuracies (with an AUC value of 0.75) using statistical mining and machine learning models, where the input dataset included an extensive array of 31 variables obtained through questionnaires [7]. In contrast, our model demonstrates its efficacy with only four input variables. Given that AUC values in the range of 0.7-0.8 are generally deemed acceptable for discrimination accuracy, this study underscores the capacity of a well-trained generalized KOA prediction model to offer insights into the likelihood of KOA development, particularly at a population level and for providing generalized risk assessments for individuals. Furthermore, the versatility of GPR models extends to the evaluation of the impact of changes in input parameters (e.g., weight) on OA risk. Nevertheless, for predictions tailored to individual cases, especially those intended for patient engagement or treatment planning, models incorporating a higher degree of subject-specific information should be favored.

**SIGNIFICANCE/CLINICAL RELEVANCE:** This study introduces a KOA prediction model ideal to be used to predict the disease progression at a population level or provide directive KOA risk assessment for individuals.

**REFERENCES:** [1] Mononen, Sci Rep, 2023; [2] Kokkotis, OA Cartilage Open, 2020; [3] Mononen, Ann Biomed Eng, 2019; [4] Kellgren & Lawrence, Ann Rheum Dis, 1957; [5] Cai, Arthritis Res Ther, 2019; [6] Karvonen, J Rheumatol, 1994; [7] Bertoncelli, Ther Adv Musculoskelet Dis, 2022

Figure 1. The workflow for using the $GPR_{DIM}$ and $GPR_{KOA}$ models.

| GPR_DIM models (validation data) | | |
|---|---|---|
| Target variable | R² | RMSE |
| $JS_{MED}$ [mm] | 0.67*** | 0.67 |
| $JS_{LAT}$ [mm] | 0.72*** | 0.47 |
| $W_{Xray}$ [mm] | 0.90*** | 2.26 |

Table 1. Accuracy of the trained $GPR_{DIM}$ models with $R^2$ (Pearson correlation, *** $p < 0.001$) and RMSE between model and measured dimensions.
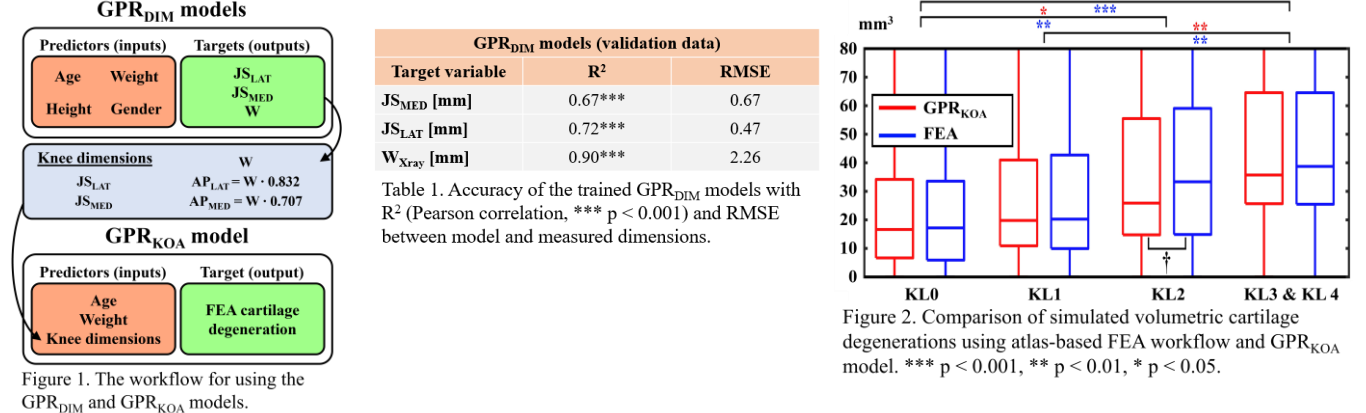


Figure 2. Comparison of simulated volumetric cartilage degenerations using atlas-based FEA workflow and $GPR_{KOA}$ model. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.