Evidence-Based Machine Learning Algorithm to Predict Failure following Cartilage Preservation Procedures in the Knee

Ron Gilat, Ben Gilat, Sumit Shailesh Patel, Kyle R Wagner, Eric Haunschild, Tracy Mae Tauro, Jorge A Chahla, Adam Blair Yanke[1], Brian J Cole[2]
[1]Rush University Med Ctr, [2]Rush University Medical Center

INTRODUCTION: Management of symptomatic patients who failed conservative treatment for focal chondral and osteochondral defects in the knee is complex and multifactorial. The selection process of a surgical treatment modality for a specific patient is complex. In most cases, this decision depends solely on the surgeon's knowledge, belief, and expertise. However, in recent years, machine learning tools have been increasingly utilized to improve the repeatability and accuracy of outcome predictions. The purpose of this study was to develop machine learning algorithms to predict failure of surgical procedures that address cartilage defects of the knee and detect the most valuable variables associated with failure.

METHODS: A single institution prospectively collected database of cartilage procedures was queried for procedures performed between 2000 and 2018. Failure was defined as revision cartilage surgery and/or knee arthroplasty. One hundred-one preoperative and intraoperative features were evaluated as potential predictors. The dataset was randomly divided into training (70%) and independent testing (30%) sets. Four machine learning algorithms were trained and internally validated. Algorithm performance was assessed using area under curve (AUC) and the Brier score. Local Interpretable Model-agnostic Explanations (LIME) was utilized to assess the optimized algorithm fidelity.

RESULTS: A total of 1,091 patients who underwent surgical procedures addressing cartilage defects in the knee with a minimum of 2-years of follow-up were included. The most-common procedure was chondroplasty (n=560; 51%) followed by osteochondral allograft transplantation (n=306; 28%), microfracture (n=150; 14%), autologous chondrocyte implantation (n=39; 4%), and osteochondral autograft transplantation (n=36; 3%). The mean follow up was 3.5 ± 2.8 years. The mean age was 40.5 ± 15 years. There were 205 (18.8%) patients who failed at final follow up. The Random Forest algorithm was found to be the best performing algorithm, with an AUC of 0.765 and a Brier score of 0.135. The 10 most important features for predicting failure following surgical procedures addressing cartilage defects of the knee were: symptom duration, age, body mass index (BMI), lesion grade, total lesion area (sum of all lesion areas), number of previous surgeries, number of lesions in the knee, gender, athletic level, and traumatic etiology. LIME analysis allowed for assessment of the optimized algorithm fidelity, as well as provided a patient-specific comparison for the risk of failure of an individual patient being assigned various types of cartilage procedures.

DISCUSSION AND CONCLUSION: Machine learning algorithms were accurate in predicting the risk of failure following cartilage procedures of the knee, with the most important features in descending order being symptom duration, age, BMI, lesion grade, and total lesion area. Machine learning algorithms may be used to compare the risk of failure of specific patient-procedure combinations in the treatment of cartilage defects of the knee. Integrated human and machine learning decision making may improve patient selection and bring about the new era of patient-tailored evidence-based clinical care.

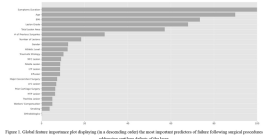

Figure 1. Global feature importance plot displaying (in a descending order) the most important predictors of failure following surgical procedures addressing cartilage defects of the knee.

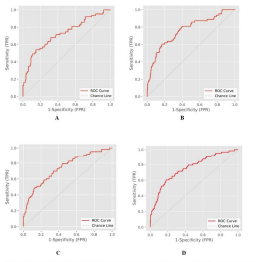

Figure 2.A-D. The corresponding receiver operating characteristic (ROC) curve analysis of the four algorithms. A. Elastic-net penalized linear regression (ENPLR, AUC=0.725); B. Random Forest (AUC=0.765); C. Neural Network (AUC=0.742); D. XGBoost (AUC=0.757).

Table 1. Demographic and Clinical Characteristics

	Overall (N=1091)
Age at time of surgery years	40.5 ± 15
Gender	
Male	550 (50.4%)
Female	541 (49.6%)
Body Mass Index	28.2 ± 6.0
Laterality	
Right	569 (52.2%)
Left	522 (47.8%)
Smoking	
Never	937 (86%)
Yes	46 (4.2%)
Former	108 (9.8%)
Diabetes Mellitus	24 (2.2%)
Hypertension	142 (13%)
Thyroid Disease	61 (5.6%)
Athlete	293 (26.9%)
Worker's Compensation	119 (10.9%)
Traumatic event	439 (40.2%)
Symptoms' duration years	2.7 ± 4.7
Osteochondritis dissecans	85 (7.8%)
Recurrent effusion	468 (42.9%)
Preoperative flexion (degrees)	127 ± 15.9
Preoperative extension (degrees)	0.9 ± 6.3
Preoperative alignment	
Neutral	977 (89.6%)
Varus	58 (5.3%)
Valgus	56 (5.1%)
Follow-up years	3.5 ± 2.8

Continuous variables presented in means ± standard deviation
Binomial variables are presented in frequencies (proportions)

Table 2. Surgical details and concomitant procedures.

Cartilage Lesion Location	
MFC	554 (50.8%)
MTP	144 (13.2%)
LFC	285 (26.1%)
LTP	145 (13.3%)
Trochlea	293 (26.9%)
Patella	329 (30.2%)
Defect Area (mm², mean±SD)	
MFC	17.8 ± 14
MTP	10.2 ± 9.4
LFC	18.7 ± 13.9
LTP	11.7 ± 10.1
Trochlea	16.4 ± 13.8
Patella	16 ± 13.5
Cartilage Lesion Grade median	4
Cartilage Procedure	
Chondroplasty	560
Microfracture	150
Osteochondral Allograft Transplantation (OCA)	306
Osteochondral Autograft Transplantation (OATS)	36
Articular Chondrocyte Implantation (ACI/MACI)	39
Concomitant Procedure	
Medial Meniscectomy	481 (44.1%)
Lateral Meniscectomy	289 (26.5%)
Medial Meniscus Repair	13 (1.2%)
Lateral Meniscus Repair	7 (0.6%)
Medial MAT	53 (4.8%)
Lateral MAT	77 (7.1%)
High Tibial Osteotomy	32 (2.9%)
Distal Femoral Osteotomy	25 (2.3%)
Tibial Tuberosity Osteotomy	51 (4.7%)
ACL reconstruction	157 (14.4%)
Platelet-rich plasma injection	14 (1.3%)
Bone Marrow Aspirate Concentrate (BMAC)	11 (1%)

ACL, anterior cruciate ligament reconstruction; LFC, lateral femoral condyle; LTP, lateral tibial plateau; MAT, meniscal allograft transplantation; MFC, medial femoral condyle; MTP, medial tibial plateau. Continuous variables presented in means ± standard deviation. Binomial variables are presented in frequencies (proportions).