Prediction of the Impact of Physical Exercise on Knee Osteoarthritis Patients using Kinematic Signal Analysis and Decision Trees

M. Mezghani[1,2], N. Hagemeister[2], M. Kouki[1], Y. Ouakrim[2,4], A. Fuentes[3] and N. Mezghani[2,4]

[1]*École Supérieure de la Statistique et de l'Analyse de l'Information, Université de Carthage, Tunisia*

[2]*Laboratoire de Recherche en Imagerie et Orthopédie (LIO), Centre de Recherche du CHUM, Montreal, Canada*

[3] *EMOVI Inc, Quebec, Canada*

[4]*LICEF Reserach Center, TELUQ University, Montreal, Canada*

Keywords: 3D Kinematics, Decision Trees, Knee Osteoarthritis, Physical Exercise, Knee Kinesiography.

Abstract: The evaluation of knee biomechanics provides valuable clinical information. This can be done by means of a knee kinesiography exam which measures the three-dimensional rotation angles during walking, thus providing objective knowledge about knee function (3D kinematics). 3D kinematic data is quantifiable information that provides opportunities to develop automatic and objective methods for personalized computer-aided treatment systems. The purpose of this study is to explore a decision tree based method for predicting the impact of physical exercise on a knee osteoarthritis population. The prediction is based on 3D kinematic data i.e., flexion/extension, abduction/adduction and internal/external rotation of the knee. Experiments were conducted on a dataset of 309 patients who have engaged in physical exercise for 6 months and have been grouped into two classes, Improved state (I) and not-Improved state (nI) based on their state before (t_0) and after the exercise (t_6). The method developed was able to predict I and nI patien with knee osteoarthritis using 3D kinematic data with an accuracy of 82%. Results show the effectiveness of 3D kinematic signal analysis and the decision tree technique for predicting the impact of physical exercise based on patient knee osteoarthritis pain level.

1 INTRODUCTION

The knee is a joint of great anatomical and biomechanical complexity, and is the basis for the mobility and stability of the human body. This joint undergoes various static and dynamic stresses that make it subject to several degenerative diseases, including knee osteoarthritis (OA). The World Health Organization estimates that 10% of the adult population in developed countries suffers from osteoarthritis, 6.1% of which affects the knee (Woolf and Pfleger, 2003). In Canada, hundreds of thousands of people suffer from osteoarthritis of the knee, which affects their functional abilities and undermines their quality of life. The prevalence of osteoarthritis is increasing as the population ages. Indeed, professionals estimate that it will have doubled by the year 2020 (Creamer and Hochberg, 1997). Knee osteoarthritis is also one of the most important chronic diseases in terms of the use of health services (Health Council of Canada, 2007).

Although there are protocol and practice guidelines for better management of osteoarthritis, several studies show that the treatment of this pathology is far from optimal and that significant deficiencies exist both for the diagnosis and for the therapeutic management of knee osteoarthritis.The diagnosis of osteoarthritis can be made either by a family doctor (general practitioner) or by an orthopedic surgeon after a musculoskeletal evaluation that can be combined with an imaging examination (X-ray). However, because radiological examinations are performed in a static state, they collect data on the integrity of knee structures but do not describe the functional aspects of the knee. While such examinations allow us to determine the impact of an injury on knee function, they are not sensitive enough to allow clinicians to make an informed choice about the treatment to be prescribed, particularly when it comes to prescribing physical exercise. Properly formulated exercises are very important because they aim to delay the progression of the disease while improving symptoms and joint function. In addition, exercise is among the non-surgical treatments that have the most scientific evidence to support their effectiveness (Fransen et al., 2015). Despite this, no studies have investigated the effect of

115

exercises to predict their impact on the osteoarthritis and thus examine whether they could improve state of the knee.

In this context, functional evaluation of the knee provides valuable clinical information. It can be done by means of a knee kinesiography exam which measures the three-dimensional rotation angles (3D kinematics) of the knee during walking, thus making it possible to identify mechanical biomarkers directly related to the progression of the disease and the patient's symptoms. This type of evaluation can easily be performed in a clinical setting using the KneeKG^{TM} system (EMOVI Inc, Montreal, Canada).

Figure 1: Knee kinematic acquisition system.

Several studies have demonstrated the accuracy, validity and reproducibility of 3D knee movements measured by this technology (Lustig et al., 2011). For the first time, then, the use of this system allows to assess a valid and accurate functional evaluation of the knee in a clinical setting. This evaluation is non-invasive and reliable allows the biomechanical function of the knee to be evaluated and analyzed in 3D, in real time, in motion and under load. 3D kinematic data, which is quantifiable information about knee function, provides an opportunity to objectively study the impact of a physical exercise program on knee function.

2 METHODOLOGY

To achieve our objective, we adopted the methodology described in the block diagram of Figure 2 which is based on the following main steps. The first step was to establish a database of knee OA patients who had completed a 6-month exercise program. The state of the patient was assessed before starting and after six months of exercise. This assessment was based on the Knee Injury and Osteoarthritis Outcome Score (KOOS) questionnaire, which assesses the patient's knee pain (9 items), other symptoms (7 items), function and daily life (17 items), sport and recreation

(5 items) and knee related quality of life (4 items). These scores range from 0 to 100 with a score of 0 indicating the worst possible knee symptoms and 100 indicating no symptoms of knee pathology. Based on the questionnaire responses, the participants were grouped into two classes: a class of patients whose condition has been improved (I) and a class whose condition has not been improved (nI). The second step was to develop a classification system to predict the impact of physical exercise, using knee 3D kinematic data as input. The predicted state is displayed via an interactive platform in order to better serve clinicians.

2.1 Data-base

A total of 309 participants with knee osteoarthritis took part in this study. All had completed an exercise program. Participants are described by their demographic characteristics (age, gender and Body Mass Index(BMI)) and one clinical characteristic, the radiographic severity grade of their osteoarthritis measured by the Kellgren-Lawrence scale (grade 2: mild; grade 3: moderate; grade 4: severe). These characteristics were measured at the beginning of the project (t_0) and after 6 months (t_6). All participants also completed the KOOS questionnaires before (t_0) and after the exercise program (t_6), to assess whether or not there was an improvement in the patient's condition (as described in section 2.2).

Table 1: Demographic characteristics of I (improved) and nI (not-improved) classes (BMI designates the mean body mass index).

	Class I N = 141	Class nI N = 168
Age (years)	62.9 ± 9.51	63.66 ± 8.44
BMI (kg / m^2)	30.2 ± 6.54	29.89 ± 6.0
Proportion of men / group	36%	40%

2.2 Identification of Patients According to their Improvement Status

Patient improvement was determined based on the KOOS. This questionnaire provided us with a global evaluation of the knee as well as five specific evaluations of particular aspects (pain, symptoms, function and daily life, sport and leisure and quality of life). The score values measured before (at time t_0) and after the physical exercise (at time t_6) were used to deduce whether the patient's condition had improved (Class I) or not (Class nI). Indeed, according to the literature , a condition can be considered improved
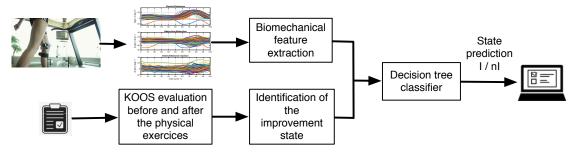
Figure 2: Block diagram of the methodology adopted.

if the end state assessment exceeds that of the initial state by a threshold of 10 % (Roos and Lohmander, 2003). This threshold can be applied to the six variables (the five KOOS subscales and the overall KOOS), which may introduce several sources of variability in the identification of patients based on their improvement state. Preliminary research (Bensalma et al., 2019) suggests that the $KOOS_{pain}$ is the most representative score. In the discussion below, states I and nI are therefore determined using $KOOS_{pain}$. In other words, for each patient, the variable η is computed as follows:

$$\eta = \frac{KOOS_{pain}(t_6) - KOOS_{pain}(t_0)}{KOOS_{pain}(t_0)}$$

and, the assigned state (class) is then:

$$\begin{cases} I, & \text{if } \eta \geq 10\% \\ nI, & \text{otherwise} \end{cases} \quad (1)$$

2.3 Kinematic Factor Extraction

Kinematic data describes the joint angles between the tibia and femur in the three-dimensional space (3D). These are in the form of 3D curves corresponding to flexion-extension in the sagittal plane, abduction-adduction in the frontal plane and internal-external rotation in the transverse plane. These curves are normalized to a range from 1% to 100% where 1 corresponds to the beginning and 100 to the end of the cycle. From these curves, parameters of interest are then extracted to characterize the pattern of each participant. The parameters in our case, are 11 kinematic factors extracted from the flexion-extension, abduction-adduction and internal-external rotation curves.

2.4 Classification System

In order to predict the patient's improvement status (I or nI), we have developed a supervised classification
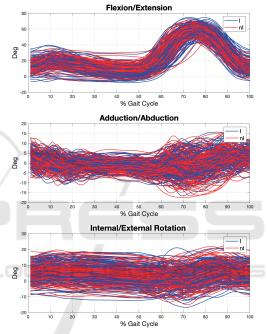

Figure 3: Kinematic curves of the 309 subjects in the database (I: improved state and nI: not improved state).

system based on decision trees. The decision trees were built using the Classification And Regression Tree (CART) algorithm and allows to reach the end of each path from the root to a leaf, a description of one of the classes.

The algorithm for building a binary decision tree using CART operates node by node, running through the M attributes $(x_1, x_2, ..., x_M)$ one by one, starting with x_1 and continuing until x_M. For each attribute, it explores all possible tests (splits) and chooses the best split, i.e., the one that maximizes impurity (uncertainty) reduction. Then, it compares the M best splits to choose the best one. The function that measures impurity will necessarily reach its maximum when the instances are equitably distributed among the different classes and its minimum when one class contains all the examples (the node is pure). In order

to build the nodes of the tree, "most discriminating" questions are chosen by the Gini index (Girard, 2007).

This index measures the frequency with which a random element of the set would be misclassified if its label were randomly selected based on the distribution of labels in the sub-set. The index ranges between 0 and 1 and reaches its minimum value (zero) when all the elements of the set are in the same class of the target variable. The Gini diversity index used by the CART algorithm can be calculated by the following formula: on a node t with a probability distribution of the classes on this node $P(j|t), j = 1, ..., J$, we have (Hawarah, 2008):

$$G(t) = i(t) = \phi(p(1/t), p(2/t), ..., p(J/t))$$
$$= 1 - \sum_j (P(j|t))^2$$

where $p(j/t)$ is the proportion of individuals belonging to class j and ϕ is the proportion function to measure the impurity $i(t)$.

It should be noted that the decision trees were pruned by the post-pruning method to avoid over-learning. This approach proceeds as follows: after completing the decision tree building process, the tree is pruned. To this end, classification errors are estimated at each node. The subset is replaced by a leaf (class) or by the most frequent branch. We then start at the bottom of the tree and examine each of the sub-trees (non-folio) to see whether replacing the sub-tree by a leaf or its most frequent branch would result in to a lower error rate. If so, we trim the sub-tree by performing the replacement(Hawarah, 2008).

2.5 Evaluation of Classification System

The evaluation of the classification system was carried out by dividing the database into two sub-databases: a training database and a test database. This division allows the model to be developed and tested on different data to verify its relevance. In our case, we opted for a division allocating 2/3 of the data for training and the remaining 1/3 for testing. This gave us 206 participants for the learning process and 103 participants for the validation of the model.

Following training, we considered the classification rate as an evaluation criterion. This rate is the ratio of the total number of well classified data points to the total number of data points.

$$\text{Classification rate} = \frac{\text{Well classified observations}}{\text{Total number of observations}}$$

The confusion matrix can also be presented for a better interpretation of the results. This is a matrix representation that determines the classification error from a set of test data. The confusion matrix is a square matrix of size $[C \times C]$ where C is the number of classes. The columns of this matrix correspond to the number of occurrences of an estimated class, while the rows correspond to the number of occurrences of an actual class. The following table shows an example of a confusion matrix with two classes. The precision of the classifier is calculated by formula (2), and the sensitivity and specificity by formulas (4) and (5), respectively.

	Predicted class	
Real class	C1	C2
C1	True Positive	False Negative
C2	False Positive	True Negative

True Positive (TP) = C1 group participant correctly classified.
True Negative (TN) = C2 group participant correctly classified.
False Positive (FP) = C2 group participant classified as C1.
False Negative (FN) = C1 group participant classified as C2.
The mathematical formulas for the evaluation parameters of a classifier are as follows:

$$Accuracy = \frac{TP + TN}{Total\ number\ of\ observations} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

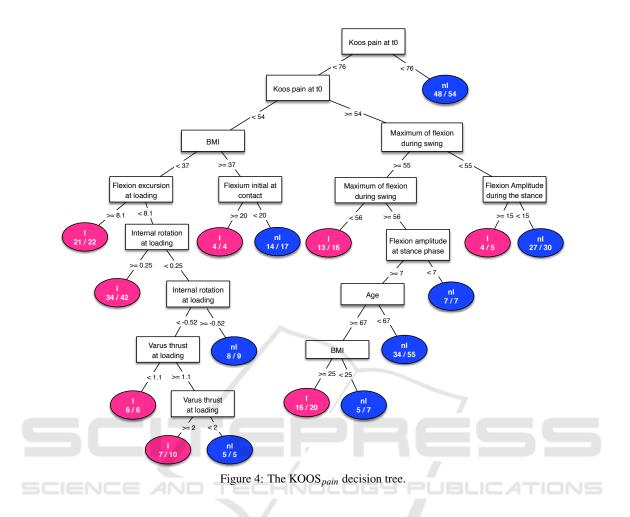$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

To be considered accurate, a classifier must be both highly sensitive and highly specific.

3 RESULTS: CLASSIFICATION BASED ON DECISION TREES

We developed the classification system using the CART algorithm and an input vector of 13 variables, i.e., 9 kinematic factors, 2 demographic variables (BMI and Age) and the $KOOS_{pain}$ score at t_0. Training was performed using both the hold out approach (2/3 of the data for training and 1/3 for testing) and leave-one out cross validation. Table 2 shows that the classification rate reaches 82% for leave-one out cross

Figure 4: The KOOS$_{pain}$ decision tree.

validation and 75% for hold out approach based on the 1/3 testing data.

Table 2: Classification rates using the CART algorithm.

Training and testing	Classification rate
2/3 of the database for training and 1/3 for testing	75%
Leave-one-out cross validation (309 patients)	82%

Figure 4 illustrates the decision tree obtained. Of the 13 input variables used, 9 were retained by the CART algorithm. The KOOS$_{pain}$ score at t_0 was identified as the most important variables ranked by impurity.

Table 3: Confusion matrix for the training based on leave-one-out cross validation.

	Predicted class	
Real class	I	nI
I	105	36
nI	20	148

4 CONCLUSION

To our knowledge, this study is the first to explore, the use of machine learning techniques to predict the impact (improvement or not) of physical exercise on the knee in a gonarthrotic population. To this end, we first developed a large database composed of subjects who had completed a personalized physical exercise program, whose condition was measured at the beginning of the project (time t_0) and after 6 months (time t_6). In a second step, we developed a classification system based on decision trees. This classification system uses 3D knee kinematic data as input to perform an objective, evidence-based decision. The decision trees achieved a classification rate of 82% based on KOOS$_{pain}$ using a leave-one-out procedure (75% based on KOOS$_{pain}$ on test data only). Unlike many classification methods, decision trees are intuitive and provide a graphic, meaningful and easy-to-read representation. This advantage has been exploited by implementing a user-friendly graphical interface that allows clinicians to query patient characteristics for a

better understanding of the classification system's decision.

ACKNOWLEDGMENTS

REFERENCES

Bensalma, F., Mezghani, N., Ouakrim, Y., Fuentes, A., Choiniere, M., Bureau, N., Durand, M., and Hagemeister, N. (2019). A multivariate relationship between the kinematic and clinical parameters of knee osteoarthritis population. *Applied Bionics and Biomechanics*, 2019:14.

Creamer, P. and Hochberg, M. (1997). Osteoarthritis. *Lancet*, 350:503–8.

Fransen, M., McConnell, S., Harmer, A., der Esch, M. V., Simic, M., and Bennell, K. (2015). Exercise for osteoarthritis of the knee: a cochrane systematic review. *British Journal of Sports Medicine*.

Girard, A. (2007). Exploration d'un algorithme génétique et d'un arbre de décision à des fins de categorisation. Master's thesis, Université du Québec à Trois-Rivières.

Hawarah, L. (2008). Une approche probabiliste pour le classement d'objets incompletement connus dans un arbre de décision. Master's thesis, Ecole Doctorale MSTII.

Lustig, S., Magnussen, R., Cheze, L., and Neyret, P. (2011). The kneekg system: A review of the literature. *Knee Surgery, Sport. Traumatol. Arthrosc.*

Roos, E. M. and Lohmander, L. S. (2003). The knee injury and osteoarthritis outcome score (koos): from joint injury to osteoarthritis. *Health Qual Life Outcomes*.

Woolf, A. and Pfleger, B. (2003). Burden of major musculoskeletal conditions. *Bulletin of the World Health Organisation*, 81(9):646–56.