

Article

3D Kinematics and Decision Trees to Predict the Impact of a Physical Exercise Program on Knee Osteoarthritis Patients

Marwa Mezghani ^{1,2} , Nicola Hagemester ³, Youssef Ouakrim ^{2,3}, Alix Cagnin ³, Alexandre Fuentes ⁴ and Neila Mezghani ^{2,3,*} 

¹ École Supérieure de la Statistique et de L'analyse de L'information de Tunis, Université de Carthage, 1080 Tunis, Tunisia; marwamezghani96@gmail.com

² LICEF Institute, TELUQ University, Montréal, QC H2S 3L5, Canada; oua.youssef@gmail.com

³ Laboratoire de Recherche en Imagerie et Orthopédie (LIO), École de Technologie Supérieure (ETS), CRCHUM, Montréal, QC H2X 0A9, Canada; nicola.hagemester@etsmtl.ca (N.H.); alix.cagnin.1@ens.etsmtl.ca (A.C.)

⁴ EMOVI Inc., Montreal, QC H2S 1B1, Canada; afuentes@emovi.ca

* Correspondence: neila.mezghani@teluq.ca

Abstract: Measuring knee biomechanics provides valuable clinical information for defining patient-specific treatment options, including patient-oriented physical exercise programs. It can be done by a knee kinesiology test measuring the three-dimensional rotation angles (3D kinematics) during walking, thus providing objective knowledge about knee function in dynamic and weight-bearing conditions. The purpose of this study was to assess whether 3D kinematics can be efficiently used to predict the impact of a physical exercise program on the condition of knee osteoarthritis (OA) patients. The prediction was based on 3D knee kinematic data, namely flexion/extension, adduction/abduction and external/internal rotation angles collected during a treadmill walking session at baseline. These measurements are quantifiable information suitable to develop automatic and objective methods for personalized computer-aided treatment systems. The dataset included 221 patients who followed a personalized therapeutic physical exercise program for 6 months and were then assigned to one of two classes, Improved condition (I) and not-Improved condition (nI). A 10% improvement in pain was needed at the 6-month follow-up compared to baseline to be in the improved group. The developed model was able to predict I and nI with 84.4% accuracy for men and 75.5% for women using a decision tree classifier trained with 3D knee kinematic data taken at baseline and a 10-fold validation procedure. The models showed that men with an impaired control of their varus thrust and a higher pain level at baseline, and women with a greater amplitude of internal tibia rotation were more likely to report improvements in their pain level after 6 months of exercises. Results support the effectiveness of decision trees and the relevance of 3D kinematic data to objectively predict knee OA patients' response to a treatment consisting of a physical exercise program.

Keywords: 3D kinematics; decision trees; knee osteoarthritis; physical exercises



Citation: Mezghani, M.; Hagemester, N.; Ouakrim, Y.; Cagnin, A.; Fuentes, A.; Mezghani, N. 3D Kinematics and Decision Trees to Predict the Impact of a Physical Exercise Program on Knee Osteoarthritis Patients. *Appl. Sci.* **2021**, *11*, 834. <https://doi.org/10.3390/app11020834>

Received: 28 September 2020

Accepted: 12 January 2021

Published: 17 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The knee is an anatomically and biomechanically complex joint that serves as 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 (WHO) estimates that 10% of the adult population in developed countries suffers from OA, 6.1% of which affects the knee [1]. In Canada, hundreds of thousands of people suffer from knee OA, which affects their functional abilities and undermines their quality of life [2].

Although there are protocols and clinical guidelines for the management of knee osteoarthritis, several studies showed that treatments are far from optimal and that significant clinical gaps exist in the therapeutic management of knee OA [3,4]. Osteoarthritis can be diagnosed by a physician (typically a general practitioner, an orthopedic surgeon or a

rheumatologist) after a musculoskeletal evaluation that can be combined with an imaging assessment (X-ray). Radiographic examinations collect information on the integrity of knee structures but because they are performed in a static condition, they do not provide information on the knee's functional status. Although such examinations allow to assess the impact of an injury on anatomical structures, they do not provide clinicians with information to support whether a conservative treatment, such as physical exercises, should be prescribed or not.

In this context, a dynamic functional evaluation of the knee provides valuable clinical information (e.g., misalignments during gait) [5]. This evaluation can be completed with a knee kinesiography exam which measures the three-dimensional rotation angles (3D kinematics) during gait, allowing to identify mechanical biomarkers directly related to the progression of the disease and the patient's symptoms. This objective assessment allows clinicians to recommend personalized exercises targeting the previously identified mechanical biomarkers. This type of evaluation can easily be performed in a clinical setting using the KneeKG™ system (EMOVI Inc., Montreal, QC, Canada; Figure 1). This system consists of passive motion sensors fixed on the knee harness, an infrared motion capture system and a computer equipped with an acquisition software. The harness is fixed quasi-rigidly on the thigh and calf to measure tibial and femoral rotation [6]. Several studies have demonstrated the accuracy, validity and reproducibility of 3D knee movements measured with this technology [6–8].



Figure 1. Knee kinematic acquisition system. Measures are captured while walking on a conventional treadmill.

Exercises are among the non-surgical treatment options with the most research evidence supporting their effectiveness [9]. Overall, they aim to strengthen muscles and improve flexibility and balance in order to alleviate symptoms and improve joint function [10]. However, the factors that could predict which patients would be more likely to respond to such a treatment have yet to be determined. Indeed, the vast majority of studies on predictive models in knee OA populations have focused on identifying risk factors to help predict disease progression [11–13]. Notably, age, body mass index (BMI), or radiographic grading appeared to play an important role, and thus they were mainly used as the only possible predictors in the few studies which actually predicted the impact of physical exercise programs on knee OA [14,15].

To our knowledge, only two studies assessed the role of knee kinematics as predictors of patient response to exercises and both suggested that they can be useful to optimize exercise recommendations [16,17]. However, these studies did not assess 3D kinematics with a high degree of accuracy. Therefore, the aim of the present study was to investigate whether 3D kinematic data can be efficiently used to objectively predict the impact of a physical exercise program.

2. Materials and Methods

The methodology used is described in the block diagram presented in Figure 2 and involves the following steps: (1) establishing a database of knee OA patients with 3D knee kinematics measurements at baseline and at the completion of a 6-month exercise program, (2) identifying patients according to the improvement in their condition, (3) extracting biomechanical factors from their kinematic data, and (4) building a prediction model based on decision trees.

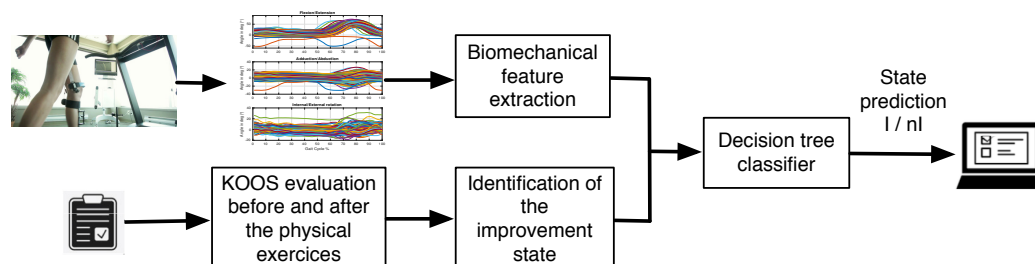


Figure 2. Block diagram of the adopted methodology.

2.1. Participants and Exercise Program

The data used in this study were collected in a cluster randomized controlled trial (RCT) approved by the institutional ethics committees of the University of Montreal Hospital Research Center (Reference numbers: CE 10.001-BSP and BD 07.001-BSP), and of the École de technologie supérieure (Reference numbers: H20100301 and H20170901). All subjects provided informed consent before participation.

Of all the participants with knee osteoarthritis who took part in this RCT [18], 221 patients completed a 6-month personalized home-based exercise program and were included in this study. The demographic characteristics (age, sex, and BMI) and one clinical feature, the radiographic OA severity grade measured by the Kellgren-Lawrence scale (KL; grade 2: mild; grade 3: moderate; grade 4: severe) [19] were collected for all participants. KL grades were evenly distributed within the cohort (KL2: 61; KL3: 82; KL4: 78).

Each patient completed a knee kinesiography exam at baseline. For each one of them, a physical therapist then created a unique program of five to ten land-based exercises. It had to combine strengthening, stretching, and gait retraining exercises addressing the mechanical biomarkers previously identified with the knee kinesiography exam (e.g., varus thrust, dynamic flexion contracture...). Each exercise was recommended with a patient-specific number of repetitions (or duration) and had to be achievable at home without supervision. Patients were simply encouraged to “do them regularly” and were asked if they followed their exercise program after 3 months and at the 6-month follow-up. Further details about the program (therapist training, examples of exercises...) are published elsewhere [18].

2.2. Identification of Patients According to the Improvement in Their Condition

The patients underwent a clinical assessment at baseline and after six months, using the Knee Injury and Osteoarthritis Outcome Score (KOOS) [20]. This questionnaire assesses five domains: the patient’s knee pain (9 items), other symptoms (7 items), function in daily life (17 items), sport and recreation (5 items) and knee-related quality of life (4 items). Each one of them can be analyzed individually with a score which ranges from 0 (indicating the worst-case scenario) and 100 (indicating the absence of knee symptomatology).

Patient improvement was determined based on the KOOS questionnaire. Based on a literature review and a preliminary study [21], the $KOOS_{pain}$ subscale was identified as the most representative score and therefore its variation between baseline and 6 months was used to determine whether a patient’s condition improved (Class I) or not (Class nI). A growing consensus in the knee OA literature suggests that identifying a single threshold to define a meaningful change in the $KOOS_{pain}$ score across varying OA severity and follow-up periods is problematic [22,23]. Furthermore, a recent meta-analysis [24] supports that

estimated thresholds increase along with patient baseline severity, and that relative changes may be a better approach than absolute differences. Based on the authors' conclusions and the variability in our participants' clinical characteristics (KL grades 2, 3, and 4), we considered that a patient's condition was improved if the $KOOS_{pain}$ score at 6 months exceeded the baseline score by 10% or more, as this value is both "intuitive and consistent with estimates for other instruments" [24].

Hence, the participants were divided into two classes based on the evolution of their $KOOS_{pain}$ between baseline and 6-month follow-up: one class of patients whose condition improved (I) and one class of patients whose condition did not improve (nI).

In other words, for each patient, the variable η was computed as follows:

$$\eta = \frac{KOOS_{pain}(t_6) - KOOS_{pain}(t_0)}{KOOS_{pain}(t_0)} \times 100$$

and, the assigned condition (class) was then:

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

Based on this assessment, the participants were categorized according to their improvement status. Tables 1–3 summarize the demographic characteristics of the two classes (i.e., I and nI) for all participants, the male and female population respectively. There was no statistical difference in terms of age and BMI distributions between classes regardless of the population considered (T -tests: all $p > 0.68$). Statistical processing was implemented via SPSS 18.0 (Statistical Package for Social Sciences). A p -value of 0.05 was set as the criterion for statistical significance.

Table 1. Demographic characteristics of the participants within each class.

Characteristics (Units)	Class I N = 112	Class nI N = 109	p -Value
Age (years)	63.87 ± 9.57	64.31 ± 7.86	0.80
BMI (kg/m ²)	29.93 ± 5.51	29.75 ± 5.63	0.70

Table 2. Demographic characteristics of the male participants within each class.

Characteristics (Units)	Class I N = 37	Class nI N = 41	p -Value
Age (years)	64.81 ± 9.56	64.80 ± 7.93	0.99
BMI (kg/m ²)	30.07 ± 5.61	29.71 ± 5.11	0.77

Table 3. Demographic characteristics of the female participants within each class.

Characteristics (Units)	Class I N = 75	Class nI N = 68	p -Value
Age (years)	63.40 ± 9.60	64.01 ± 7.85	0.68
BMI (kg/m ²)	29.87 ± 5.50	29.77 ± 5.95	0.91

2.3. Biomechanical Factors Extraction

Kinematic data describe the joint angles between the tibia and the femur in the three-dimensional space. These are in the form of 3D curves corresponding to flexion-extension in the sagittal plane, adduction-abduction in the frontal plane and external-internal rotation

in the transverse plane. These curves are mean patterns of multiple gait cycles collected per subject and normalized to a range from 1% to 100% of the gait cycle (GC). The beginning of the GC (1%) corresponds to the heel contact which is identified by the first minimum of the flexion/extension curve (Figure 3).

A set of 69 biomechanical parameters of interest was then extracted from these 3D kinematic curves for data analysis. The parameters chosen for extraction were based on variables routinely assessed in biomechanical studies of knee OA populations, such as maximums, minimums, varus and valgus thrust, angles at initial contact, mean values and ranges of motion (ROM) throughout GCs or GC sub-phases (i.e., loading, stance, swing, etc.) [25,26].

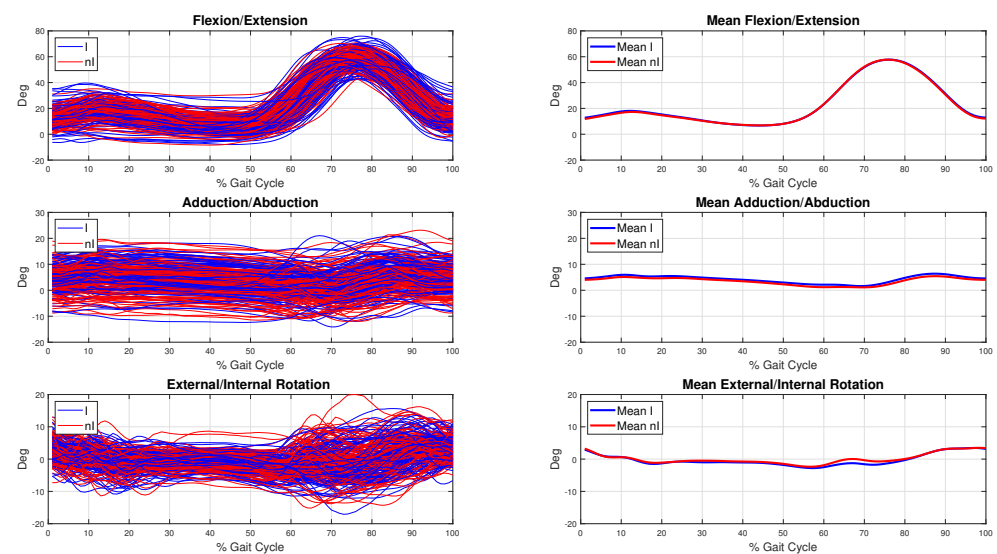


Figure 3. Kinematic curves of the 221 subjects (I: improved (blue) and nI: not-improved (red)). Angles are expressed in degrees: Flexion(+)/Extension(−), Adduction(+)/Abduction(−), External(+)/Internal(−) rotation.

2.4. Prediction Model

In order to predict patient improvement (I or nI), we developed a supervised classification system based on decision trees. The decision trees were built using the Classification And Regression Tree (CART) algorithm that can be used for classification or regression predictive modeling problems.

The algorithm to build a binary decision tree using CART operates node by node, running through the M attributes (x_1, x_2, \dots, x_M) one by one, starting with x_1 and continuing on to x_M . For each attribute, it explores all possible tests (splits) and chooses the best split, that is, the one that maximizes impurity (uncertainty) reduction. Then, it compares the best M splits to select the best one. The function that measures impurity will necessarily reach its maximum when the instances are evenly distributed among the different classes and its minimum when one class contains all the examples (the node is then considered pure). In order to build the most discriminating nodes, questions are generated by the Gini index [27].

This index measures the frequency with which a random element in the set would be misclassified if its label was randomly selected based on the label distribution in the sub-set. The index ranges from 0 to 1 and reaches its minimum value (zero) when all the elements from the set are in the same class as the target variable. The Gini diversity index

used by the CART algorithm can be calculated with the following formula: on a node t with a probability distribution for the classes in this node $P(j|t), j = 1, \dots, J$, we have [28]:

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

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 a lower error rate. If so, we trim the sub-tree using the replacement [28].

2.5. Evaluation of the Prediction Model

The evaluation of the classification system was carried out by two cross validation methods. The first is a K -fold cross validation where the database is divided 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 study, we opted for $K = 10$ and $K = 20$ divisions to evaluate the model's stability.

The second evaluation method is the Leave-one-out cross validation (LOOCV) which is a K -fold cross validation taken to its logical extreme, with K equal to N , the number of data points in the set, that is to say the number of participants.

After model training, we considered the classification rate as an evaluation criterion. This rate is the ratio between the total number of well-classified data points and the total number of data points as described in Equation (3). The classification rate is computed for each fold and the average used to evaluate the model.

$$\text{Classification rate} = \frac{\text{Well-classified observations}}{\text{Total number of observations}}. \tag{3}$$

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 $[C \times C]$ size 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. Table 4 shows an example of a confusion matrix with two classes. The precision of the classifier is calculated by Formula (4), and the sensitivity and specificity by Formulas (5) and (6), respectively [29].

Table 4. Example of a confusion matrix. I = improved; nI = not-improved.

Actual Class	Predicted Class	
	I	nI
I	True Positive	False Negative
nI	False Positive	True Negative

True Positive (TP) is the number of patients that improved and were classified in the class I. True Negative (TN) includes the patients that did not improved and were classified as nI patients. False Positive (FP) describes patients that not improved but were classified as I patients. False Negative (FN) is the number of patients that improved but were classified in the class nI.

To be considered accurate, a classifier must be both highly sensitive and highly specific. The accuracy, sensitivity and specificity were computed according to the following

mathematical formulas:

$$\text{Accuracy} = \frac{\text{TN} + \text{FP}}{\text{Total number of observation}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TN} + \text{FP}} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TN} + \text{FP}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

Significant differences in 3D knee kinematics between men and women have been reported in the literature, both in healthy and knee OA subjects [30,31]. Therefore, three different prediction systems were trained: one each for the male and the female population and one for the overall data-set (male and female together) to account for the sex-specific aspects of kinematics. The receiver operating characteristic (ROC) curve was generated to illustrate the performance of the prediction models, using the area under the curve (AUC) [32].

3. Results

Prediction models were developed using the CART algorithm. The training input vectors consisted of the 69 biomechanical parameters (F_1, F_2, \dots, F_{69}), patients' age (F_{70}), BMI (F_{71}), radiographic OA severity grade measured by the Kellgren-Lawrence scale (F_{72}), and the KOOS_{pain} (F_{73}). The training input measurements were evaluated at baseline as required by the principle of a prediction model. Analyses were undertaken with Matlab R2019b software (Mathworks, MA, USA). Specifically, we used the Toolbox Statistics and Machine Learning Toolbox.

Table 5 summarizes the classification rates of the three models developed for each validation method. Given that the prediction performances were better when the two sexes were analyzed separately, the prediction model focused on the impact of the physical exercise program for each group independently.

Table 5. Classification rates of the three models for each validation technique. Rates for K-fold techniques are presented \pm standard deviation.

Validation Techniques	Datasets		
	Male (N = 78)	Female (N = 143)	All (N = 221)
10-Fold	84.4% \pm 13.2	75.5% \pm 6.8	71% \pm 9.9
20-Fold	84.6% \pm 19.2	75.7% \pm 14.3	71% \pm 14.6
LOOCV	78.2%	73.4%	71%

Figure 4 displays the ROC curve and the AUC for both the male and female prediction models.

The confusion matrices using the LOOCV are presented in Tables 6 and 7. For instance, the prediction model within the male population reached a sensitivity of 85.4% and a specificity of 70.3%.

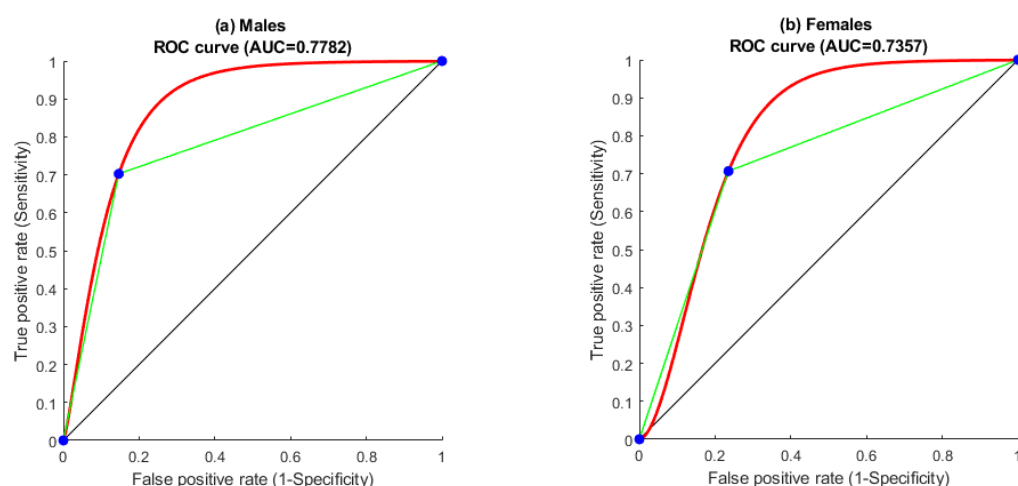


Figure 4. The ROC curves and AUC of the prediction models (a) Males and (b) Females.

Table 6. Confusion matrix of the male prediction model. Accuracy is presented for each class (I and nI).

Real Class	Predicted Class		Accuracy
	nI	I	
nI	35 (85.4%)	6 (14.6%)	85.4%
I	11 (29.7%)	26 (70.3%)	70.3%

Table 7. Confusion matrix of the female prediction model. Accuracy is presented for each class (I and nI).

Real Class	Predicted Class		Accuracy
	nI	I	
nI	52 (76.5%)	16 (23.5%)	76.5%
I	22 (29.3%)	53 (70.7%)	70.7%

For a more in-depth analysis, we identified the discriminant features retained within the 69 biomechanical factors for each prediction model (Table 8).

Table 8. Description of the retained features. All features (except for KOOS_{pain}) are expressed in degrees.

Dataset	Feature Id	Feature Name	Feature Description
Male	F ₄	Varus thrust during loading	Maximum adduction angle between 10% and 30% of the GC minus the adduction angle value at 10% of the GC
	F ₂₆	Flexion at initial swing	Flexion angle value at 69% of the GC
	F ₇₃	KOOS _{pain}	KOOS score on the pain subscale at baseline
Female	F ₁₄	External tibial rotation at initial contact	Rotation angle value at 1% of the GC
	F ₁₅	Internal tibial rotation during loading	Maximum minus the minimum rotation angle value between 1% and 20% of the GC
	F ₃₉	Maximal tibial rotation during gait	Maximum rotation angle value between 1% and 100% of the GC

Figure 5 shows the decision trees for each of the prediction models for both the male and female populations. Training and validation were performed independently for each population.

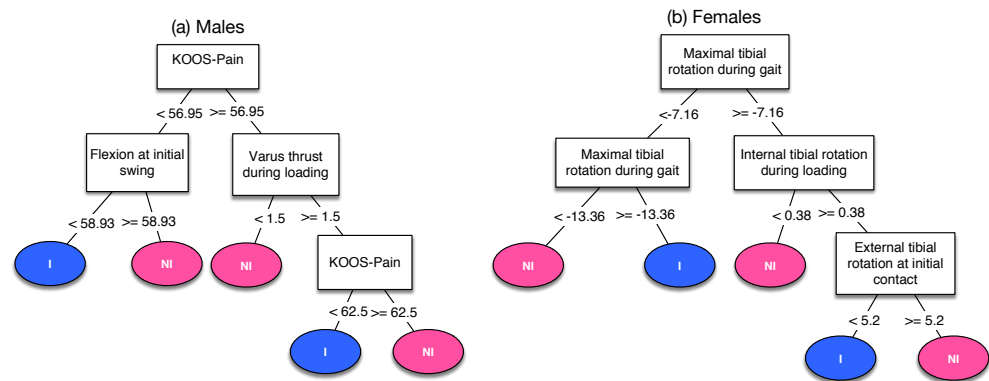


Figure 5. Decision trees for each of the prediction models (a) Males and (b) Females. All features (except for KOOS_{pain} score) are expressed in degrees. Female prediction model: external(+)/internal(−) tibia rotation.

4. Discussion

Results showed that regardless of the validation technique used, the classification rates were higher when the male and the female population were considered separately. Indeed, using a 10-fold validation, the prediction performance reached 84.4% for men and 75.5% for women but it was limited to 71 % when the two sexes were grouped together. This confirms that prediction models should be sex-specific to achieve better results and supports previous studies highlighting the differences in kinematics between men and women [30,31].

The analysis of the confusion matrices and the ROC curves showed that the specificities and the sensitivities led to accurate prediction models (i.e., 70% or higher). In both cases sensitivity (85.4% and 76.5% respectively for men and women prediction models) was better than specificity (70.3% and 70.7% respectively for men and women prediction models). Therefore, the models were better at predicting a not-Improved (nI) condition than an Improved (I) condition. Unlike many classification methods, decision trees are intuitive and provide a graphic, meaningful and easy-to-read representation compared to other well-known prediction and classification methods such as neural networks and support vector machine. The prediction model was enhanced by implementing a user-friendly graphical interface allowing clinicians to query patient characteristics and improve their understanding of the classification system's decision.

The three discriminant features retained for the male prediction model were entirely different from the three chosen for the female model, confirming that sex-specific analyses were appropriate. Interestingly, among the 69 features that were measured on 3D kinematic curves at baseline, the features retained for both decision trees were closely linked to OA progression [33]. Varus thrust is a sudden lateral shift of the knee when the weight on the limb increases during the loading phase at the beginning of the gait cycle (frontal plane). It is well known that this mechanism is associated with pain and increase risk of OA progression [34,35]. Its significant role was also highlighted in another study assessing the impact of knee kinematics on the response to exercises [17]. As shown in this study, patients with a higher varus thrust at baseline were more likely to respond to a physical exercise program. Furthermore, a lower KOOS_{pain} at baseline (<62.5) appeared to be predictive of an improvement for male patients. This was also observed by Kobsar et al. [16] who used a predictive model based on principal component analysis, as well as being reported in the hip OA literature [36]. Notably, the discriminant features retained in the male prediction model were related to frontal and sagittal plane kinematics, whereas features in the female prediction model were all related to tibial rotation.

These results reached a state-of-the-art accuracy and compared favorably with others that have considered the prediction of the impact of physical exercise. Indeed, the purpose of Reference [16] was to use pre-intervention gait kinematics and patient-reported outcome measures to predict post-intervention response to a 6-week hip strengthening exercise intervention in patients with mild-to-moderate knee OA. Using a discriminant analysis on a small dataset (39 patients), the classification accuracy was 85.4%.

Being male is associated with higher odds of presenting with a varus thrust [37], which could help explain the important role of this feature played in the male prediction model. This study suggests that exercises efficiently reduce pain in patients with varus thrust, supporting their relevance as a conservative treatment especially in the presence of this mechanical biomarker. Tibial rotation, which corresponds to external/internal rotation of the shank with respect to the thigh (transverse plane), was the most discriminant movement for the female model. To our knowledge, there has only been one other report of this feature but it was as a possible factor to discriminate surgical from non-surgical knee OA female patients [38]. Future studies exploring the biomechanical significance of movements in the transverse plane could help improve our understanding of their clinical relevance in the response to exercise programs.

Although the radiographic severity grade measured by the Kellgren-Lawrence scale (F_{72}) was incorporated as an input variable during the models' training phase, it was not identified as a discriminant variable. In the same way, for both decision trees, the age and BMI were not identified as discriminant variables, suggesting that these patient characteristics cannot be used to predict the knee OA patient response to a physical exercise program. Furthermore, the statistical analysis showed that there was no statistical difference on age and BMI between I and nI participants within the whole dataset.

This study has some limitations. First, results were based on a 10% improvement threshold on the KOOS to determine whether patients improved or not, which can be subject to debate. However, this choice was based on data from the literature and the use of this threshold showed that decision trees could help predict the patient response to an exercise program. Additional analyses with different thresholds could be performed to study their impact on prediction models. Similarly, the prediction models were only based on the KOOS_{pain} subscale, while other subscales could have been of interest to better understand the predictors of the response to a physical exercise program on other knee OA aspects. Although the KneeKG™ system used to capture the 3D knee kinematics is a validated and reliable tool, its accuracy in the transverse plane is 2.3° [6], which may have influenced the actual classification rate of the female prediction model. This limitation could affect overall decision tree methodology as the classification at each node is based on a single value and does not account for variations. Further analyses are needed to determine the impact of this variation in future algorithms. Finally, there was a notable difference between the total number of men and women included in this study (78 vs. 143), which may have impacted the performance of both prediction models.

5. Conclusions

To our knowledge, this study is the first to explore the combined use of machine learning techniques and kinematic data to predict the impact (improvement or not) of physical exercise programs in knee OA patients. To this end, a large database of subjects who had completed a personalized physical exercise program was used, and a classification system based on decision trees was developed. This classification system used 3D knee kinematic data as input to reach an objective, evidence-based decision. Using a 10-fold cross validation procedure, the decision trees achieved a classification rate of 84.4% and 75.5% respectively within the male and female populations. The decision trees suggest that men with impaired control of their varus thrust and a higher pain level at baseline, and women with a greater amplitude of internal tibia rotation were more likely to have less pain after undertaking a 6-month exercise program.

Author Contributions: Data curation, Y.O. and A.C.; Formal analysis, M.M. and Y.O.; Investigation, M.M. and Y.O.; Methodology, M.M. and Y.O.; Supervision, N.H. and N.M.; Validation, N.H., A.C., A.F. and N.M.; Writing—original draft, M.M.; Writing—review & editing, N.H., A.C. and N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Canada Research Chair on Biomedical Data Mining (950-231214).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committees of the University of Montreal Hospital Research Center (Reference numbers: CE 10.001-BSP and BD 07.001-BSP; 2015-04-30), and of the École de technologie supérieure (Reference numbers: H20100301 and H20170901; 2015-06-17).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ethical restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Woolf, A.; Pfleger, B. Burden of major musculoskeletal conditions. *Bull. WHO* **2003**, *81*, 646–656. [[PubMed](#)]
2. Tarride, J.E.; Haq, M.; Nakhaipour, H.R.; O'Reilly, D. The excess burden of rheumatoid arthritis in Ontario, Canada. *Clin. Exp. Rheumatol.* **2012**, *31*, 18–24. [[PubMed](#)]
3. Alami, S.; Boutron, I.; Desjeux, D.; Hirschhorn, M.; Meric, G.; Rannou, F.; Poiraudou, S. Patients' and practitioners' views of knee osteoarthritis and its management: A qualitative interview study. *PLoS ONE* **2011**, *6*, e19634. [[CrossRef](#)] [[PubMed](#)]
4. Allen, K.D.; Choong, P.F.; Davis, A.M.; Dowsey, M.M.; Dziedzic, K.S.; Emery, C.; Hunter, D.J.; Losina, E.; Page, A.E.; Roos, E.M.; et al. Osteoarthritis: Models for appropriate care across the disease continuum. *Best Pract. Res. Clin. Rheumatol.* **2016**, *30*, 503–535. [[CrossRef](#)]
5. Abid, M.; Mezghani, N.; Mitiche, A. Knee Joint Biomechanical Gait Data Classification for Knee Pathology Assessment: A Literature Review. *Appl. Bionics Biomech.* **2019**, *2019*, 7472039. [[CrossRef](#)] [[PubMed](#)]
6. Lustig, S.; Magnussen, R.; Cheze, L.; Neyret, P. The knee system: A review of the literature. *Knee Surgery, Sport. Traumatol. Arthrosc.* **2011**, *20*, 633–638. [[CrossRef](#)] [[PubMed](#)]
7. Labbe, D.R.; Hagemester, N.; Tremblay, M.; de Guise, J. Reliability of a method for analyzing three-dimensional knee kinematics during gait. *Gait Posture* **2008**, *28*, 170–174. [[CrossRef](#)] [[PubMed](#)]
8. Hagemester, N.; Parent, G.; Van de Putte, M.; St-Onge, N.; Duval, N.; de Guise, J. A reproducible method for studying three-dimensional knee kinematics. *J. Biomech.* **2005**, *38*, 1926–1931. [[CrossRef](#)]
9. Fransen, M.; McConnell, S.; Harmer, A.; der Esch, M.V.; Simic, M.; Bennell, K. Exercise for osteoarthritis of the knee: A cochrane systematic review. *Br. J. Sports Med.* **2015**, *49*, 1554–1557. [[CrossRef](#)] [[PubMed](#)]
10. Dong, R.; Wu, Y.; Xu, S.; Zhang, L.; Ying, J.; Jin, H.; Wang, P.; Xiao, L.; Tong, P. Is aquatic exercise more effective than land-based exercise for knee osteoarthritis? *Medicine* **2018**, *97*, e13823. [[CrossRef](#)]
11. Landsmeer, M.L.; Runhaar, J.; van Middelkoop, M.; Oei, E.H.; Schiphof, D.; Bindels, P.J.; Bierma-Zeinstra, S.M. Predicting knee pain and knee osteoarthritis among overweight women. *J. Am. Board. Fam. Med.* **2019**, *32*, 575–584. [[CrossRef](#)] [[PubMed](#)]
12. Kerkhof, H.J.M.; Bierma-Zeinstra, S.M.A.; Arden, N.K.; Metrustry, S.; Castano-Betancourt, M.; Hart, D.J.; Hofman, A.; Rivadeneira, F.; Oei, E.H.G.; Spector, T.M.; et al. Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. *Ann. Rheum. Dis.* **2014**, *73*, 2116–2121. [[CrossRef](#)] [[PubMed](#)]
13. LaValley, M.P.; Lo, G.H.; Price, L.L.; Driban, J.B.; Eaton, C.B.; McAlindon, T.E. Development of a clinical prediction algorithm for knee osteoarthritis structural progression in a cohort study: Value of adding measurement of subchondral bone density. *Arthr. Res. Ther.* **2017**, *19*, 95. [[CrossRef](#)] [[PubMed](#)]
14. Widera, P.; Welsing, P.M.; Ladel, C.; Loughlin, J.; Lafeber, F.P.; Dop, F.P.; Larkin, J.; Weinans, H.; Mobasheri, A.; Bacardit, J. Multi-classifier prediction of knee osteoarthritis progression from incomplete imbalanced longitudinal data. *Sci. Rep.* **2020**, *10*, 8427. [[CrossRef](#)] [[PubMed](#)]
15. Legha, A.; Burke, D.L.; Foster, N.E.; Windt, D.A.V.D.; Quicke, J.G.; Healey, E.L.; Runhaar, J.; Holden, M.A. Do comorbidities predict pain and function in knee osteoarthritis following an exercise intervention, and do they moderate the effect of exercise? Analyses of data from three randomized controlled trials. *Musculoskelet. Care* **2020**, *18*, 3–11. [[CrossRef](#)] [[PubMed](#)]
16. Kobsar, D.; Osis, S.T.; Hettinga, B.A.; Ferber, R. Gait Biomechanics and Patient-Reported Function as Predictors of Response to a Hip Strengthening Exercise Intervention in Patients with Knee Osteoarthritis. *PLoS ONE* **2015**, *10*, e0139923. [[CrossRef](#)] [[PubMed](#)]
17. Bennell, K.L.; Dobson, F.; Roos, E.M.; Skou, S.T.; Hodges, P.; Wrigley, T.V.; Kyriakides, M.; Metcalf, B.; Hunt, M.A.; Hinman, R.S. Influence of Biomechanical Characteristics on Pain and Function Outcomes From Exercise in Medial Knee Osteoarthritis and Varus Malalignment: Exploratory Analyses From a Randomized Controlled Trial. *Arthr. Care Res.* **2015**, *67*, 1281–1288. [[CrossRef](#)]

18. Cagnin, A.; Choinière, M.; Bureau, N.J.; Durand, M.; Mezghani, N.; Gaudreault, N.; Hagemester, N. A multi-arm cluster randomized clinical trial of the use of knee kinesiography in the management of osteoarthritis patients in a primary care setting. *Postgrad. Med.* **2019**, *132*, 91–101. [[CrossRef](#)]
19. Kellgren, J.H.; Lawrence, J.S. Radiological Assessment of Osteo-Arthrosis. *Ann. Rheum. Dis.* **1957**, *16*, 494–502. [[CrossRef](#)]
20. Roos, E.M.; Lohmander, L.S. The knee injury and osteoarthritis outcome score (koos): From joint injury to osteoarthritis. *Health Qual. Life Outcomes* **2003**, *1*, 64. [[CrossRef](#)]
21. Bensalma, F.; Mezghani, N.; Ouakrim, Y.; Fuentes, A.; Choiniere, M.; Bureau, N.; Durand, M.; Hagemester, N. A multivariate relationship between the kinematic and clinical parameters of knee osteoarthritis population. *Biomed. Eng. Online* **2019**, *18*, 58. [[CrossRef](#)] [[PubMed](#)]
22. Mills KA, G.; Naylor, J.M.; Eyles, J.P.; Roos, E.M.; Hunter, D.J. Examining the Minimal Important Difference of Patient-reported Outcome Measures for Individuals with Knee Osteoarthritis: A Model Using the Knee Injury and Osteoarthritis Outcome Score. *J. Rheumatol.* **2016**, *43*, 395–404. [[CrossRef](#)] [[PubMed](#)]
23. Harris, K.K.; Dawson, J.; Jones, L.D.; Beard, D.J.; Price, A.J. Extending the use of PROMs in the NHS—Using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: A validation study. *BMJ Open* **2013**, *3*, e003365. [[CrossRef](#)] [[PubMed](#)]
24. Devji, T.; Guyatt, G.H.; Lytvyn, L.; Brignardello-Petersen, R.; Foroutan, F.; Sadeghirad, B.; Buchbinder, R.; Poolman, R.W.; Harris, I.A.; Carrasco-Labra, A.; et al. Application of minimal important differences in degenerative knee disease outcomes: A systematic review and case study to inform BMJ Rapid Recommendations. *BMJ Open* **2017**, *7*, e015587. [[CrossRef](#)] [[PubMed](#)]
25. Astephen, J.L.; Deluzio, K.J.; Caldwell, G.E.; Dunbar, M.J. Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity. *J. Orthop. Res.* **2008**, *26*, 332–341. [[CrossRef](#)]
26. Bytyqi, D.; Shabani, B.; Lustig, S.; Cheze, L.; Gjurgjeala, N.K.; Neyret, P. Gait knee kinematic alterations in medial osteoarthritis: Three dimensional assessment. *Int. Orthop.* **2014**, *38*, 1191–1198. [[CrossRef](#)]
27. Girard, A. 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, Trois-Rivieres, QC, Canada, 2007.
28. Hawarah, L. Une Approche Probabiliste Pour le Classement d'Objets Incomplètement Connus Dans un Arbre de dÉcision. Master's Thesis, Joseph Fourier University, Grenoble, France, 2008.
29. Ouakrim, Y. Classification de Sujets Asymptomatiques et Gonarthrosiques en Fonction des données cinématiques: Comparaison de l'Approche Globale et de l'Approche Locale. Master's Thesis, École De Technologie Supérieure, Montreal, QC, Canada, 2011.
30. Clément, J.; Toliopoulos, P.; Hagemester, N.; Desmeules, F. Healthy 3D knee kinematics during gait: Differences between women and men, and correlation with x-ray alignment. *Gait Posture* **2018**, *64*, 198–204. [[CrossRef](#)]
31. Phinyomark, A.; Osis, S.T.; Hetingga, B.A.; Kobsar, D.; Ferber, R. Gender differences in gait kinematics for patients with knee osteoarthritis. *BMC Musculoskelet. Disord.* **2016**, *17*, 157. [[CrossRef](#)]
32. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [[CrossRef](#)]
33. Mezghani, N.; Ouakrim, Y.; Fuentes, A.; Mitiche Hagemester, N.; Vendittoli, P.A.; de Guise, J.A. Mechanical biomarkers of medial compartment knee osteoarthritis diagnosis and severity grading: Discovery phase. *J. Biomech.* **2019**, *52*, 106–112. [[CrossRef](#)]
34. Sharma, L.; Chang, A.H.; Jackson, R.D.; Nevitt, M.; Moio, K.C.; Hochberg, M.; Eaton, C.; Kwoh, C.K.; Almagor, O.; Cauley, J.; et al. Varus Thrust and Incident and Progressive Knee Osteoarthritis. *Arthr. Rheumatol.* **2017**, *69*, 2136–2143. [[CrossRef](#)] [[PubMed](#)]
35. Wink, A.E.; Gross, K.D.; Brown, C.A.; Guermazi, A.; Roemer, F.; Niu, J.; Torner, J.; Lewis, C.E.; Nevitt, M.C.; Tolstykh, I.; et al. Varus thrust during walking and the risk of incident and worsening medial tibiofemoral MRI lesions: The Multicenter Osteoarthritis Study. *Osteoarthr. Cartil./OARS Osteoarthr. Res. Soc.* **2017**, *25*, 839–845. [[CrossRef](#)] [[PubMed](#)]
36. Wright, A.A.; Cook, C.E.; Flynn, T.W.; Baxter, G.D.; Abbott, J.H. Predictors of response to physical therapy intervention in patients with primary hip osteoarthritis. *Phys. Ther.* **2011**, *91*, 510–524. [[CrossRef](#)] [[PubMed](#)]
37. Chang, A.; Hochberg, M.; Song, J.; Dunlop, D.; Chmiel, J.S.; Nevit, M.; Hayes, K.; Eaton, C.; Bathon, J.; Jackson, R.; et al. Frequency of Varus and Valgus Thrust and Factors Associated with Thrust Presence in Persons With or at Higher Risk for Knee Osteoarthritis. *Arthr. Rheum.* **2010**, *62*, 1403–1411. [[CrossRef](#)]
38. Ouakrim, Y.; Dunbar, M.; Fuentes, A.; Whynot, S.; Macdonald, H.; Richardson, G.; Mezghani, N. Classification tool identifying knee arthroplasty candidates based on kinematic data and gender. P064. In Proceedings of the 20th Biennial Meeting of the Canadian Society for Biomechanics CSB, Halifax, NS, Canada, 14–17 August 2018; CSB: Halifax, NS, Canada, 2018.