

RESEARCH

Open Access



# A multivariate relationship between the kinematic and clinical parameters of knee osteoarthritis population

Fatima Bensalma<sup>1,2\*</sup> , Neila Mezghani<sup>1,2</sup>, Youssef Ouakrim<sup>1,2</sup>, Alexandre Fuentes<sup>2</sup>, Manon Choinière<sup>3,5</sup>, Nathalie J. Bureau<sup>3,6</sup>, Madelaine Durand<sup>3,7</sup> and Nicola Hagemester<sup>2,4</sup>

\*Correspondence:

fatima.bensalma@gmail.com  
<sup>1</sup> Centre de Recherche LICEF,  
TELUQ, 5800, rue Saint-Denis,  
bureau 1105, Montreal,  
Canada

Full list of author information  
is available at the end of the  
article

## Abstract

**Background:** Biomechanical and clinical parameters contribute very closely to functional evaluations of the knee joint. To better understand knee osteoarthritis joint function, the association between a set of knee biomechanical data and a set of clinical parameters of an osteoarthritis population (OA) is investigated in this study.

**Methods:** The biomechanical data used here are a set of characteristics derived from 3D knee kinematic patterns: flexion/extension, abduction/adduction, and tibial internal/external rotation measurements, all determined during gait recording. The clinical parameters include a KOOS questionnaire and the patient's demographic characteristics. Canonical correlation analysis (CCA) is used (1) to evaluate the multivariate relationship between biomechanical data and clinical parameter sets, and (2) to cluster the most correlated parameters. Multivariate models were created within the identified clusters to determine the effect of each parameter's subset on the other. The analyses were performed on a large database containing 166 OA patients.

**Results:** The CCA results showed meaningful correlations that gave rise to three different clusters. Multivariate linear models were found explaining the subjective clinical parameters by evaluating the biomechanical data contained within each cluster.

**Conclusion:** The results showed that a multivariate analysis of the clinical symptoms and the biomechanical characteristics of knee joint function allowed a better understanding of their relationships.

**Keywords:** Biomechanics, Canonical correlation, Multivariate analysis, Multiple regression, Kinematic gait analysis, Knee osteoarthritis (OA)

## Background

Biomechanical knee assessment is increasingly used in gait analysis as a tool for characterizing the knee function [1], understanding pathological knee alterations [2], and assessing the progression of knee pathologies and their impact on gait [3]. It has already been suggested that the type and severity of biomechanical changes should be assessed since they can impact treatment outcomes [4]. Mechanical factors linked to the progression of osteoarthritis (OA) [5] and its treatment [6] have also been identified. Still, the relationship between the kinematic and clinical parameters of knee OA populations has not been sufficiently explained and remains incompletely understood. Only a few studies



have investigated the relationship between 3D knee kinematic parameters and clinical data [7, 8]. These studies have been limited to a univariate analysis implying the correlation between one kinematic parameter and one specific clinical parameter. Such analysis is not adapted to the complexity of biomechanical data [9] and can even mask several strong relationships if the parameters are considered independently.

The objective of this study is (1) to evaluate the multivariate relationship (compared to the univariate approach) between a set of biomechanical data and a set of clinical parameters of an osteoarthritis population, and (2) to cluster the most correlated parameter. The biomechanical data are a set of characteristics extracted from 3D knee kinematic patterns during gait recording: flexion/extension, abduction/adduction, and tibial internal/external rotation measurements. The clinical parameters were acquired via the Knee Osteoarthritis Outcome Score (KOOS) questionnaire. Through this questionnaire, the patient provides a valid and reliable assessment of his/her health status relative to the pathology [10]. Our hypothesis is that these subjective clinical measures may complement objective biomechanical measures for a better understanding of knee joint function.

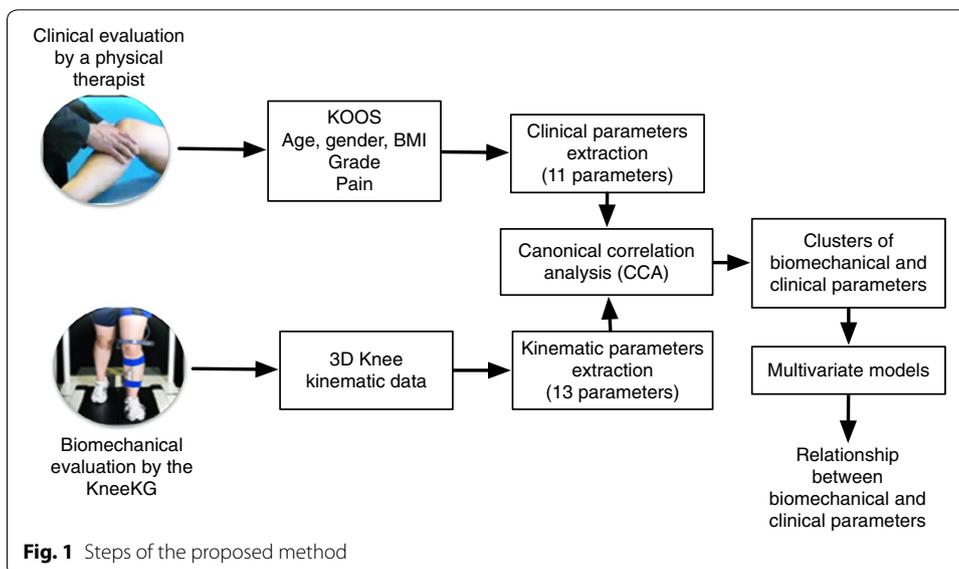
This study utilizes a canonical correlation analysis (CCA) to evaluate the relationship between a set of biomechanical data and a set of clinical parameters of an osteoarthritis population. CCA is a method for exploring the relationship between two multivariate sets of variables all measured on the same individual. Although the CCA has already been successfully applied to several applications in image processing [11] and in the domain of ecology [12], its use remains almost limited in the biomedical field. This situation could be due to the difficulty of interpreting results. To our knowledge, this study is the first to consider such a multivariate analysis combined with multivariate modeling in the biomechanical domain.

## Methods

The flowchart of the proposed method is shown in Fig. 1. The first step consists of biomechanical and clinical data acquisition and parameter extraction. Next is a multivariate analysis using a CCA, which aims at clustering the most correlated parameters. Multivariate models are then developed within the identified clusters to determine the correlation and relationships between biomechanical and clinical data.

### Biomechanical and clinical data collection

One hundred and sixty-six patients with clinically and radiographically confirmed knee osteoarthritis participated in the study [mean age of 62 years old (SD = 9.2), body mass index (BMI) of 32 kg/m<sup>2</sup> (SD = 7.3), 99 women (59.6%)]. The experimental data were collected by the KneeKG (Emovi, Canada), a knee marker attachment system (Fig. 2) designed to reduce skin-motion artifacts during motion [13]. The KneeKG was installed on participants knees to record 3D kinematics (flexion/extension, abduction/adduction, and internal/external rotation) during gait trials. The kinematic data were represented over several gait cycles (GCs) and averaged to obtain mean GCs per participant. This was followed by re-sampling of from 1 to 100% of the GCs with 100 measurement points for each participant in each plane.



Participants were also asked to answer the KOOS questionnaires. The KOOS is a valid and reliable instrument which assesses the impact of knee OA on five domains: symptoms, pain, activities of daily living (ADL), sports and recreation (Sports/Rec), and quality of life (QoL). Scores on the subscales range from 0 (extreme symptoms) to 100 (no symptoms) [10].

**Biomechanical and clinical parameters’ extraction**

For each participant, a set of 69 parameters that correspond to biomechanical values were measured on the kinematics curves from gait analysis. These parameters were chosen based on variables routinely assessed in clinical biomechanical studies of knee OA populations, such as maximums, minimums, varus and valgus thrust, angles at initial contact, mean values, and range of motion (ROM) throughout GCs or GC sub-phases

(i.e., loading, stance, swing) [14]. Thirteen parameters among this set have been identified by Mezghani et al. [3] as having the potential to serve as diagnostic and burden of disease biomarkers of knee OA. The kinematic parameters considered as biomarkers were identified by incremental selection on a regression tree determining the best set of biomechanical parameters for each biomarker type: knee OA disease diagnosis and severity grading. This has been done in accordance with the standard BIPED (burden of disease, investigative, prognostic, efficacy of intervention, and diagnostic) OA biomarker classification scheme [15]. Table 1 describes the clinical meaning of the 13 biomechanical data considered in this study.

The participants in this study were selected if the OA was the main cause of their knee pain. The exclusion criteria were considered for the subjects being on a waiting list for total knee replacement. Patients being pregnant, suffering from rheumatoid arthritis, and active cancer were also excluded. A standardized radiographic examination of both knees was performed after the patient had given written informed consent. Only patients who had a Kellgren–Lawrence (KL) grade  $\geq 2$  on radiographs were considered and only data from the most painful knee were collected.

The set of clinical parameters contains 11 measurements: the patients' demographic characteristics (sex, age, BMI), the degree of osteoarthritis severity variable (grade), the variable (pain) which is measured on the Pain Numerical Scale (NS) for Knees (on which no pain is marked 0 and the worst pain imaginable is marked 10), and 6 scores generated using the KOOS questionnaire [10, 16]. These scores assess the five dimensions mentioned above. An overall KOOS score is then generated and normalized to give a maximum of 100 points in the absence of pain or other knee dysfunction. Through this questionnaire, the patient provides a valid and reliable assessment of his health status relative to the pathology [10, 16]. A summary of the clinical parameters and their descriptions is provided in Table 2.

### Canonical correlation analysis (CCA)

The CCA is a multivariate statistical technique that explores the correlations between two sets of variables observed on the same individual [17]. The

**Table 1 Description of the 13 biomechanical parameters**

Y	Name	Description	Mean (SD)
Y <sub>1</sub>	Rot_RomSw	Range of motion of internal/external rotation during swing	1.9 (1.2)
Y <sub>2</sub>	Abd_MeanSt	Mean abduction/adduction angle during the stance phase	4.4 (5.7)
Y <sub>3</sub>	Rot_Init	Internal/external rotation angle at initial contact	3.0 (3.6)
Y <sub>4</sub>	Flex_Max	Flexion angle maximum	59.5 (5.7)
Y <sub>5</sub>	Flex_EndSt	Flexion angle at the end of the stance phase	12.1 (7.2)
Y <sub>6</sub>	Abd_MaxSw	Maximum of the abduction/adduction angle during swing phase	7.9 (5.6)
Y <sub>7</sub>	Abd_MinLo	Minimum abduction/adduction angle during the loading phase	4.3 (5.5)
Y <sub>8</sub>	Abd_Init	Abduction/adduction angle at initial contact	4.8 (5.3)
Y <sub>9</sub>	Abd_Lo	Abduction/adduction angle at the end of the loading phase	5.6 (5.9)
Y <sub>10</sub>	Abd_Rom	Range of motion of the abduction/adduction angle	9.3 (3.1)
Y <sub>11</sub>	Rot_InitAbs	Internal/external rotation absolute angle value at initial contact	3.8 (2.8)
Y <sub>12</sub>	Rot_Rom	Range of motion of the internal/external rotation	11.3 (3.2)
Y <sub>13</sub>	Abd_RomLo	Range of motion of the abduction/adduction angle during loading phase	-1.3 (2.4)

**Table 2 Description of the clinical parameters**

<b>X</b>	<b>Name</b>	<b>Description</b>	<b>Mean (SD)</b>
X <sub>1</sub>	Grade (1 to 4)	Degree of osteoarthritis severity	3.0 (0.8)
X <sub>2</sub>	Pain (4 to 10)	Outcome score for pain	6.6 (1.8)
X <sub>3</sub>	Sex (1: Men, 0: Women)	Gender	Men: n = 68; Women: n = 98
X <sub>4</sub>	Age	Age (years)	61.9 (9.2)
X <sub>5</sub>	BMI	Body mass index (kg/m <sup>2</sup> )	31.8 (7.3)
X <sub>6</sub>	KOOS_Symptoms	KOOS score for symptoms	62.7 (17.4)
X <sub>7</sub>	KOOS_Pain	KOOS score for pain	60.5 (17.3)
X <sub>8</sub>	KOOS_Adl	KOOS score for daily living	67.4 (18.3)
X <sub>9</sub>	KOOS_Sport	KOOS score for sport and recreation function	38.7 (25.7)
X <sub>10</sub>	KOOS_Qol	KOOS score for quality of life	52.3 (22.8)
X <sub>11</sub>	KOOS	Normalized overall KOOS score	56.3 (17.1)

theoretical development of CCA can be found in [18–20]. Let  $\mathbf{Y} = [Y_1, Y_2, \dots, Y_q]$  and  $\mathbf{X} = [X_1, X_2, \dots, X_p]$  denote the two data vectors to be analyzed, i.e., the biomechanical parameter vector and the clinical parameter vector. In our case  $q = 13$  and  $p = 10$ . The  $\mathbf{Y}$ -variables can be thought of as response (or dependent) variables, but in fact, the  $\mathbf{X}$  and the  $\mathbf{Y}$  sets can be interchanged without affecting the results. The aim of the CCA is to project  $\mathbf{X}$  and  $\mathbf{Y}$  datasets onto basis vectors  $\mathbf{A}$  and  $\mathbf{B}$ , respectively, in a way that the correlations between the projections of the variables onto these basis vectors are mutually maximized [21]:

$$\rho = \max_{\mathbf{A}, \mathbf{B}} \text{corr}(\mathbf{XA}, \mathbf{YB}), \tag{1}$$

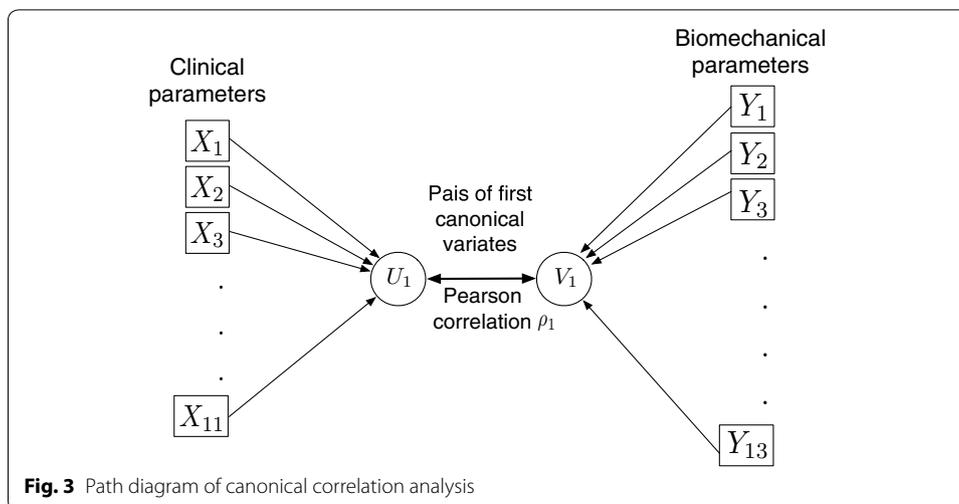
where  $\rho$  is the Pearson correlation coefficient vector and  $\mathbf{U}$  and  $\mathbf{V}$  are linear combinations of the original variables  $\mathbf{X}$  and  $\mathbf{Y}$  (Eqs. 2, 3), respectively.

$$\mathbf{U} = a_1X_1 + a_2X_2 + \dots + a_pX_p = \mathbf{XA} \tag{2}$$

$$\mathbf{V} = b_1Y_1 + b_2Y_2 + \dots + b_qY_q = \mathbf{YB}. \tag{3}$$

$\mathbf{U}$  and  $\mathbf{V}$  are the canonical variate vectors. The coefficient vectors  $\mathbf{A}$  and  $\mathbf{B}$  are known as canonical weights, canonical vectors, or canonical coefficients. The procedure is to find the first two canonical variates  $U_1$  and  $V_1$  that have the largest correlation as illustrated in Fig. 3. The maximized correlation between these two canonical variates is the first canonical correlation  $\rho_1$ . The canonical coefficients are normalized such that each canonical variate has a variance of 1. The procedure continues by finding a second pair of canonical variates  $U_2$  and  $V_2$ , uncorrelated with the first pair, that produces the second highest correlation coefficient  $\rho_2$ . The process continues until the number of pairs of canonical variables reaches pre-set  $\min(p, q)$ .

To evaluate the statistical significance of the canonical correlation model, we use the Wilks’ Lambda statistic ( $\lambda$ ). This is a multivariate statistic that uses approximations based on the Fisher distribution for the null hypothesis, i.e., all canonical correlations are zero in the population. The small  $p$  values for this test ( $< 0.05$ ) suggest a rejection of the null hypothesis and that the first canonical correlation is significant.



In our study, the analysis was conducted using the R software environment for Statistical Computing (R version 3.4.3) [22].

#### Comparison between the multivariate analysis (CCA) and a univariate analysis

The results of the CCA analysis were compared to those of a univariate analysis based on the pairwise correlation matrix calculated using the Pearson correlation coefficient. The objective of this comparison is to show that the univariate analysis cannot adapt to the complexity of biomechanical data [9] and can even mask several strong relationships if parameters are considered individually.

#### Clustering via correlation biplot

The results of a CCA are visualized by a correlation biplot graph, which represents the between-set correlation matrix  $R_{XY}$  by a joint plot. This format allows for the visualization of the intra-set correlation for the original variables and the corresponding canonical variates and of the correlation between the original variables and the opposite canonical variates. The main features of a correlation biplot are the angles between the variables from sets  $X$  and  $Y$  in the biplot, which reflect their correlations [12]. The combined angle and direction of the  $X$  and  $Y$  variables indicate the importance of the positive and negative correlations of the two sets. Strongly correlated variables are very close to each other. More specifically, in our case, the correlation biplot graph is used to cluster biomechanical data and clinical parameters. The identified clusters are then used to explain the relationships between the sets of parameters within the clusters.

#### Canonical prediction model and regression within clusters

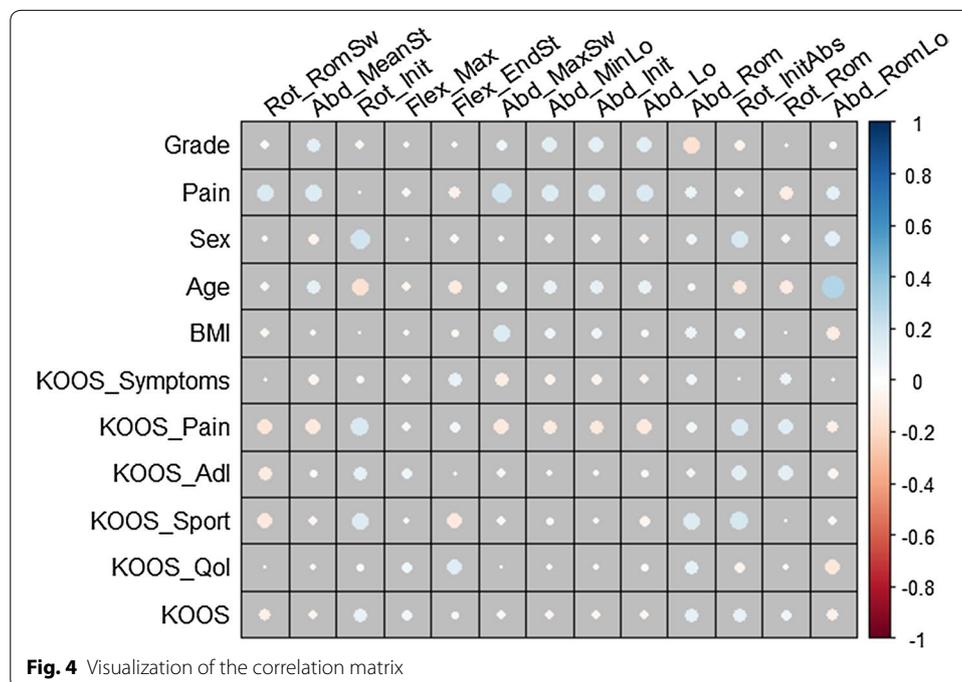
Once the clusters are identified, we can explain the relationship between the parameters within the clusters using a regression analysis. This analysis aims at estimating the coefficients of the linear equation, involving one or more independent variables (clinical parameters) that best predict the value of the dependent variables (biomechanical data). The purpose of regression is to predict  $X$  on the basis of  $Y$  within the clusters.

In order to determine which variable should be considered as dependent and which as independent, we performed a redundancy analysis. This analysis measures the proportion of variance of one original variable explained by the canonical variate of the other set. The original variables of one set are well represented by the canonical variate of the other set when the redundancy index is higher. A relational model is then proposed to determine which of the variables best explains the other. A redundancy coefficient close to 1 is considered to be the highest, and shows that the amount of the dependent (original) variable’s variance shared with the independent (canonical) variable is significant, and vice versa; a coefficient close to zero means that there is no significance in the shared variance.

## Results

### Univariate correlation analysis

The univariate correlation matrix is visualized in a graphical display in Fig. 4. The 10 clinical parameters are in rows and the 13 biomechanical parameters are in columns. Positive correlations are displayed in blue and negative correlations in red color. Color intensity and the size of the circle are proportional to the correlation coefficients. The correlations between the biomechanical data and clinical parameters are moderate. The largest correlation value is between age ( $X_4$ ) and the range of motion of the abduction/adduction angle during loading phase ( $Y_{13}$ : Abd\_RomLo) ( $r = 0.3$ ). Indeed, the univariate analysis considers the pairwise correlation of only two parameters. This result supports the need for a multivariate investigation.



### Canonical correlations and multivariate statistic

The Wilks' Lambda statistics of the canonical correlation model was  $\lambda = 0.32$ ,  $p = 0.04$ . This confirms that canonical correlations are worthy of consideration and the between-set correlations are significant. The two first higher canonical correlations are  $\rho_1 = 0.52$  and  $\rho_2 = 0.44$ .

### Correlation clustering via biplot

The correlation biplot graph of Fig. 5 represents the between-set correlation matrix  $R_{XY}$ , i.e., the correlation between 13 biomechanical parameters (in black) and 10 clinical parameters (in red) via their canonical variates. It identifies three clusters, each grouping biomechanical and clinical parameter. Recall that, the 13 biomechanical parameters among this set have been identified by Mezghani et al. [3] as having the potential to serve as diagnostic and burden of disease biomarkers of knee OA.

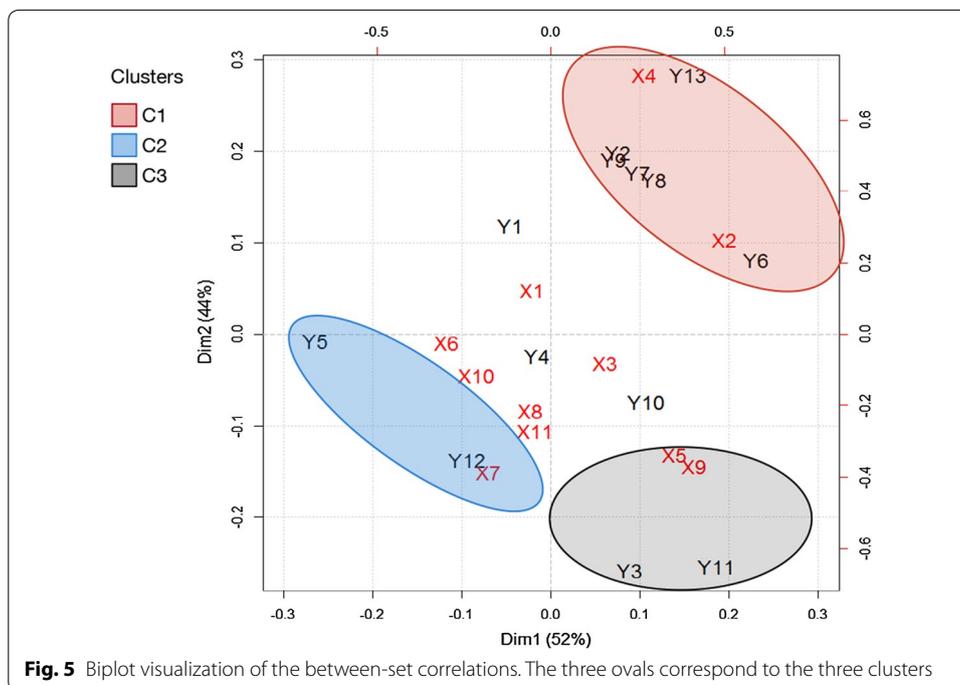
Note that the correlations located around the center are negligible. For example, the parameters  $X_1$ ,  $X_3$ ,  $Y_4$ , and  $Y_{10}$  are very close to the origin in the biplot, which shows that they are not important in the CCA. Meanwhile, the parameters  $Y_2$ ,  $Y_7$ , and  $Y_9$  are highly correlated with  $Y_8$ . Thus, these parameters have been removed from the subsequent analysis, and the identified clusters between  $X$  and  $Y$  are summarized in Table 3.

### Canonical correlation model

The canonical model of the first canonical variates is summarized in Fig. 6. This model describes the most strongly correlated variables with their appropriate weights or canonical coefficients. It provides the following two relations with a significant canonical correlation relation ( $\rho_1 = 0.52, p = 0.04$ ):

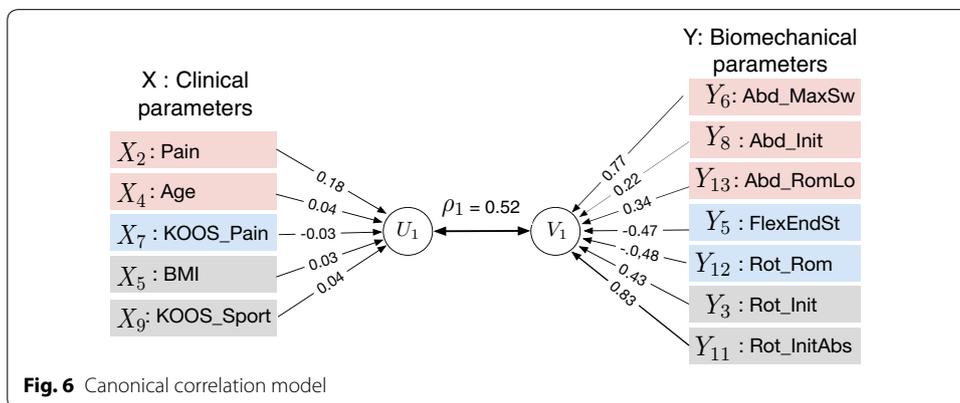
$$V_1 = 0.43 Y_3 - 0.47 Y_5 + 0.77 Y_6 + 0.22 Y_8 + 0.83 Y_{11} - 0.48 Y_{12} + 0.34 Y_{13}$$

$$U_1 = 0.18 X_2 + 0.04 X_4 + 0.03 X_5 + 0.04 X_7 - 0.03 X_9$$



**Table 3 Description of the retained correlation clusters between X and Y**

Cluster	X	↔	Y
C <sub>1</sub>	X <sub>4</sub> , X <sub>2</sub>	↔	Y <sub>6</sub> , Y <sub>8</sub> , Y <sub>13</sub>
C <sub>2</sub>	X <sub>7</sub>	↔	Y <sub>5</sub> , Y <sub>12</sub>
C <sub>3</sub>	X <sub>5</sub> , X <sub>9</sub>	↔	Y <sub>3</sub> , Y <sub>11</sub>



**Redundancy coefficients**

The total redundancy corresponds to 8.98% of the variance of X explained by the opposite canonical variate V, and to 8.71% of the variance of Y explained by the opposite canonical variate U. We can therefore affirm the equality of the indices of shared variances; more specifically, both clinical and biomechanical parameters may be considered as dependent or independent.

**Regression within the clusters**

Following the redundancy analysis, we developed a multivariate regression model within clusters to investigate the relationships between clinical and biomechanical parameters. Table 4 summarizes the regression model developed for each cluster. All the estimated regression models were significant with an Adjusted R<sup>2</sup> ≥ 0.68. The residual standard errors RSEs of the models were 0.63 ≤ RSE ≤ 1.05 indicating a perfect fit to the data by the estimated models.

**Discussion**

**Cluster 1 analysis**

The cluster C<sub>1</sub> regroups biomechanical data corresponding to kinematic parameters in the frontal plane (abduction/adduction) (Y<sub>6</sub>: Abd\_MaxSw, Y<sub>8</sub>: Abd\_Init and Y<sub>13</sub>: Abd\_ROMLo) and the level of pain (X<sub>2</sub>) as described in Tables 1, 2, and 3. The results of the multivariate regression of pain as a function of three parameters of the abduction/adduction movement (the X<sub>2</sub>: Pain regression model in Table 4) indicate that the pain felt is negatively correlated with Y<sub>8</sub> and Y<sub>13</sub>, while positively correlated with Y<sub>6</sub>.

**Table 4 Multiple linear regression models**

Cluster	Variables	Coefficient (Std. error)	P value	Residual Std. rrror	Adjusted $R^2$	P value of F statistic
<b>C<sub>1</sub></b>	$X_2$ : Pain regression model			0.68	0.70	< 0.01
	$Y_6$ : Abd_MaxSw	0.19 (0.02)	< 0.01			
	$Y_8$ : Abd_Init	− 0.33 (0.08)	< 0.01			
	$Y_{13}$ : Abd_RomLo	− 0.14 (0.06)	0.035			
<b>C<sub>2</sub></b>	$X_7$ : KOOS_Pain regression model			0.63	0.89	< 0.01
	$Y_{12}$ : Rot_Rom	0.26 (0.02)	< 0.01			
	$Y_5$ : Flex_EndSt	0.06 (0.01)	< 0.01			
<b>C<sub>3</sub></b>	$X_9$ : KOOS_Sport regression model			1.05	0.68	< 0.01
	$Y_3$ : Rot_Init	0.12 (0.05)	0.02			
	$X_5$ : BMI	1.39 (0.11)	< 0.01			

**Cluster 2 analysis**

From the second cluster **C<sub>2</sub>**, the Flexion angle at the end of the stance phase ( $Y_5$ :Flex\_EndSt), the Range of motion of the internal/external rotation ( $Y_{12}$ :Rot\_Rom), and the pain measured by the score KOOS ( $X_7$ :KOOS\_Pain) were very directly related. The association between the improvement in KOOS\_Pain score and changes in the range of motion (ROM) in the transverse plane was identified by Makovey et al. [23]. The subjective value of KOOS\_pain is positively correlated with parameters in the sagittal (flexion/extension) and transverse (internal/external rotation) plane as shown by the  $X_7$ :KOOS\_Pain regression model in Table 4.

**Cluster 3 analysis**

From the third cluster **C<sub>3</sub>**, only kinematic parameters in the transverse plane (internal/external rotation)  $Y_3$  presented correlations with  $X_5$  (BMI) and  $X_9$  (KOOS\_Sport), more precisely the internal/external rotation angle at initial contact. The improvement in KOOS\_Sport score was identified by Makovey et al. [23] as being related to the changes in the range of motion (ROM) in the transverse plane. Therefore, the model explaining the value of KOOS\_Sport and recreation score as a function of the kinematic parameters in the transverse plane and the BMI showed (Table 4) positive correlations.

When comparing the multiple regression models in **C<sub>1</sub>** and **C<sub>2</sub>** (Table 4), we note that they are both related to pain scores ( $X_2$ : Pain Numerical Scale and  $X_7$ : KOOS\_Pain) but they are not associated with the same kinematic parameters. Indeed, these two scores, i.e.,  $X_2$  and  $X_7$ , are quite different because they are evaluated based on different symptoms: the Pain numerical Scale variable ( $X_2$ ) was evaluated on a 0–10 pain intensity scale and concerns a general pain felt for knees, whereas KOOS\_Pain variable ( $X_7$ ) was evaluated based on (9) questions, especially relative to the knee injury [10].

## Conclusion

The CCA results showed a moderate correlation that gave rise to three clusters of the most closely related parameters. Multivariate linear models were found complementing the subjective clinical parameters by the biomechanical data using the correlation clusters.

Only the age, BMI, pain which is measured based on Pain Numerical Scale (NS), KOOS\_Pain, and KOOS\_Sport scores were correlated with the kinematic parameters (mechanical biomarkers). Biomechanical data corresponding to kinematic parameters in the frontal plane (abduction/adduction) during swing phase, the kinematic parameters in the sagittal plane (flexion/extension) at the end of the stance phase, and the kinematic parameters in the transverse plane (internal/external rotation) were positively correlated with pain. In other words, pain increased when kinematic parameters in those planes also increased. On the other hand, Biomechanical data corresponding to kinematic parameters in the frontal plane (abduction/adduction) at initial contact and during the loading phase were correlated negatively with pain. This means a decrease in the frontal plane at those phases is related with an increase in pain level. Kinematic parameters in the transverse plane (internal/external rotation) were correlated positively with the KOOS sport and recreation function. This means that KOOS\_Sport increased when movement in the transverse plane was more increased in kinematic parameters.

Finally, the results show that a multivariate analysis of the clinical symptoms and the biomechanical characteristics of knee joint function allows a better understanding of their relationships and would help to better understand how biomechanical characteristics can be used in guiding clinical decision making in OA management.

## Acknowledgements

The authors would like to thank Alix Cagnin for data collection and management.

## Authors' contributions

Conceptualization, formal analysis, and methodology: FB. Resources, supervision and project administration: NM. Data curation: YO. Software, validation, and visualization: FB. Writing—original draft preparation: FB. Writing—review and editing: FB and NM. Revising and correcting: NM, AF, MC, NB, MD, and NH. Clinical interpretation: NM, AF, and NH. Discussion of findings and their relevance: NM, AF, MC, NB, and NH. All authors read and approved the final manuscript.

## Funding

This research was supported by the Canada Research Chair on Biomedical Data Mining (950-231214) and FPQIS\_Fond de partenariat pour un Québec innovant et en santé.

## Availability of data and materials

The participants data are of the confidential category and cannot be put in an open repository for unrestricted public access. This restriction is imposed by the ethics committee of the University of Montreal Hospital Research Center. However, they can be made available upon request provided a statement of confidentiality is signed. Please contact the Quality system manager Clarisse Bascans (clarisse.bascans@etsmtl.ca).

## Ethics approval and consent to participate

The data collection was approved by 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).

## Consent for publication

All authors give their consent for publication.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Centre de Recherche LICEF, TELUQ, 5800, rue Saint-Denis, bureau 1105, Montreal, Canada. <sup>2</sup> Laboratoire de recherche en Imagerie et Orthopédie (LIO), Research Center of the Centre hospitalier de l'Université de Montréal (CRCHUM), 900, rue Saint-Denis, Montreal, Canada. <sup>3</sup> Research Center of the Centre hospitalier de l'Université de Montréal (CRCHUM), 900, rue Saint-Denis, Montreal, Canada. <sup>4</sup> École de Technologie Supérieure, 1100, rue Notre-Dame Ouest, Montreal, Canada. <sup>5</sup> Department of Anesthesiology University and Pain Medicine, Faculty of Medicine, Université de Montréal,

2900, boul. Édouard-Montpetit, Montreal, Canada. <sup>6</sup> Department of Radiology, Faculty of Medicine, Université de Montréal, 2900, boul. Édouard-Montpetit, Montreal, Canada. <sup>7</sup> Department of Medicine, Université de Montréal, 2900, boul. Édouard-Montpetit, Montreal, Canada.

Received: 25 October 2018 Accepted: 30 April 2019

Published online: 15 May 2019

## References

1. Clement J, Toliopoulos P, Hagemeister N, Vendittoli PA. Healthy 3d knee kinematics during gait: differences between women and men, and correlation with x-ray alignment. *Gait Posture*. 2018;64:198–204.
2. Bytyqi D, Shabani B, Lustig S, Cheze L, Gjurgjeala NK, Neyret P. Gait knee kinematic alterations in medial osteoarthritis: three dimensional assessment. *Int Orthop*. 2014;38(6):1191–8.
3. Mezghani N, Ouakrim Y, Fuentes A, Mitiche A, Hagemeister N, Vendittoli P, de Guise JA. Mechanical biomarkers of medial compartment knee osteoarthritis diagnosis and severity grading: discovery phase. *J Biomech*. 2017;52:106–12.
4. Webster ke, Wittwer JE, O'Brien J, Feller JA. Gait patterns after anterior cruciate ligament reconstruction are related to graft type. *Am J Sports Med*. 2005;33(2):247–54.
5. Kurz MJ, Stergiou N, Buzzi UH, Georgoulis AD. The effect of anterior cruciate ligament reconstruction on lower extremity relative phase dynamics during walking and running. *Knee Surg Sports Traumatol Arthrosc*. 2005;13(2):107–15.
6. Gaudreault N, Mezghani N, Turcot K, Hagemeister N, Boivin K, de Guis JA. Effects of physiotherapy treatment on knee osteoarthritis gait data using principal component analysis. *Clin Biomech*. 2011;26(3):284–91.
7. Bensalma F, Dunbar M, Whynot S, Fuentes A, Macdonald H, Ouakrim Y, Richardson G, Mezghani N. Correlations between kinematics and clinical measures in end-staged knee osteoarthritis patients. In: The 20th biennial meeting of the Canadian society for biomechanics. Halifax, Nova Scotia. 2018.
8. Wilson JA, Deluzio K, Dunbar M, Caldwell G, Hubley-Kozey C. The association between knee joint biomechanics and neuromuscular control and moderate knee osteoarthritis radiographic and pain severity. *Osteoarthr Cartil*. 2011;19(2):186–93.
9. Mezghani N, Mechmeche I, Ouakrim Y, Mitiche A. An analysis of 3d knee kinematic data complexity in knee osteoarthritis and asymptomatic controls. *PLoS ONE*. 2018;13(10):e0202348.
10. Roos EM, Lohmander LS. The knee injury and osteoarthritis outcome score (koos): from joint injury to osteoarthritis. *Health Qual Life Outcomes*. 2003;1(1):64.
11. H, S. Discriminant canonical correlation analysis for interactive satellite image change detection. In: 2015 IEEE international geoscience and remote sensing symposium (IGARSS). 2015. p. 2789–92.
12. Legendre P, Legendre LL. Numerical ecology. Second English edition. Amsterdam: Elsevier; 1998.
13. Lustig S, Magnussen R, Cheze L, Neyret P. The kneeKG system: a review of the literature. *Knee Surg Sports Traumatol Arthrosc*. 2012;20(4):633–8.
14. Astephen JL, Deluzio KJ, Caldwell GE, Dunbar MJ. Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity. *J Orthop Res*. 2008;26(3):332–41. <https://doi.org/10.1002/jor.20496>.
15. Bauer D, Hunter D, Abramson S, Attur M, Corr M, Felson D, Heinegård D, Jordan J, Kepler T, Lane N, Saxne T, Tyree B, Kraus V, For the Osteoarthritis Biomarkers Network, V. Classification of osteoarthritis biomarkers: a proposed approach. *Osteoarthr Cartil*. 2006;14(8):723–7. <https://doi.org/10.1016/j.joca.2006.04.001>.
16. Kanhonou M. Méthode d'évaluation fonctionnelle de la gonarthrose par étude de la pseudo-cinématique d'accroupissement. Ph.D thesis, École de technologie supérieure. 2017.
17. Hotelling H. Relations between two sets of variates. *Biometrika*. 1936;28(3/4):321–77.
18. Gittins R. Canonical analysis: a review with applications in ecology, vol. 12. Berlin: Springer; 2012.
19. Mardia K, Kent J, Bibby J. Multivariate analysis. London: Academic Press Inc; 1979.
20. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate data analysis, vol. 5. Upper Saddle River: Prentice hall; 1998.
21. Hardoon DR, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: an overview with application to learning methods. *Neural Comput*. 2004;16(12):2639–64.
22. González I, Déjean S, Martin PG, Baccini A. CCA: an R package to extend canonical correlation analysis. *J Stat Softw*. 2008;23(12):1–14.
23. Makovey J, Fuentes A, Williams M, Hunter D. Improved knee kinematics is associated with improved clinical outcomes in persons with symptomatic knee osteoarthritis. *Osteoarthr Cartil*. 2014;22:90–1.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.