35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Interpretation of Movement during Stair Ascent for Predicting Severity and Prognosis of Knee Osteoarthritis in Elderly Women Using Support Vector Machine

Tae Keun Yoo, Sung Kean Kim, Soo Beom Choi, Deog Young Kim, Deok Won Kim, *Life member, IEEE**

Abstract— **Several studies have demonstrated that pathologic movement changes in knee osteoarthritis (OA) may contribute to disease progression. The aim of this study was to investigate the association between movement changes during stair ascent and pain, radiographic severity, and prognosis of knee OA in the elderly women using machine learning (ML) over a seven-year follow-up period. Eighteen elderly female patients with knee OA and 20 healthy controls were enrolled. Kinematic data for stair ascent were obtained using a 3D-motion analysis system at baseline. Kinematic factors were analyzed based on one of the popular ML methods, support vector machines (SVM). SVM was used to search kinematic predictors associated with pain, radiographic severity of knee OA, and unfavorable outcomes, which were defined as persistent knee pain as reported at the seven-year follow-up or as having undergone total knee replacement during the follow-up period. Six patients (46.2%) had unfavorable outcomes at the seven-year follow-up. SVM showed accuracy of detection of knee OA (97.4%), prediction of pain (83.3%), radiographic severity (83.3%), and unfavorable outcomes (69.2%). The predictors with SVM included the time of stair ascent, maximal anterior pelvis tilting, knee flexion at initial foot contact, and ankle dorsiflexion at initial foot contact. The interpretation of movement during stair ascent using ML may be helpful for physicians not only in detecting knee OA, but also in evaluating pain and radiographic severity.**

I. Introduction

Knee osteoarthritis (OA) is more common in women than men and mainly occurs in the elderly [1]. Patients with knee OA complain of knee pain with physical activity [2]. Over the long term, the limitation of knee joint mobility and the loss of muscle strength around the knee result in pathologic movement changes in the lower body [3]. In particular, movement such as ascending stairs causes more significant symptoms and pathologic changes than walking on level ground [4]. Studies have shown that the early diagnosis and treatment of OA could help prevent some of the movement changes that occur early in OA that lead to higher knee joint loading and aggravation of symptoms [5].

Recently, many studies have applied the machine learning (ML) to gait analysis for prediction of diseases [6]. This technique may overcome current biomechanical methods, which are time-consuming and subjective. ML is an area of artificial intelligence research which uses statistical methods for data classification. Support vector machines (SVM) have been a widely used ML technique in medicine and bioinformatics for selecting informative variables or genes and to predict diseases more accurately [7]. The SVM is based on mapping data to a higher dimensional space through a kernel function, and choosing the maximum-margin hyper-plane that separates training data [8]. The goal of the SVM is to improve accuracy by the optimization of space separation.

Knee radiography is the primary diagnostic test for knee OA. However, the most important factors in designing a therapeutic plan are pain and limitation of daily activities [2]. Given the difficulties for physicians in developing therapeutic plans, a better method for the evaluation of pain, radiographic severity, and prognosis in knee OA is needed. We hypothesized that the movement associated with ascending stairs would be related to pain and radiographic severity. We further hypothesized that these patterns of movement would affect long-term clinical outcomes, such as total knee replacement, which is often considered the definitive treatment for severe OA. Thus, in this study, our aims were: (1) to differentiate patients with knee OA from controls using SVM; (2) to predict pain, radiographic severity, and prognosis at long-term follow-up based on the SVM model; (3) to determine kinematic factors associated with pain, radiographic severity, and prognosis in elderly women with knee OA.

II. Methods

A. Study Design

This study was a secondary analysis performed at a seven-year follow-up using the gait database of Severance Research Institute of Rehabilitation, Yonsei University College of Medicine [9]. Fig. 1 is a flow chart of the ML prediction, as well as the inclusion or exclusion criteria for study participants. This cross-sectional study included 18 patients with knee OA and 20 healthy controls enrolled between 2002 and 2004. Informed consent was obtained from

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A2A2A03045612).

T. K. Yoo is with the Department of Medicine, Yonsei University College of Medicine, Seoul, Korea (e-mail: fawoo2@yuhs.ac).

S. K. Kim is with the Graduate Program in Biomedical Engineering, Yonsei University, Seoul, Korea (e-mail: sdm04sdm@yuhs.ac).

S. B. Choi is with the Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul, Korea (e-mail: plains7@yuhs.ac).

D. Y. Kim is with the Department and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, Korea (e-mail: kimdy@yuhs.ac).

*D. W. Kim is a Professor with the Department of Medical Engineering, Yonsei University College of Medicine, Seoul, Korea (phone: 82-2-2228-1916; fax: 82-2-364-1572; e-mail: kdw@yuhs.ac).

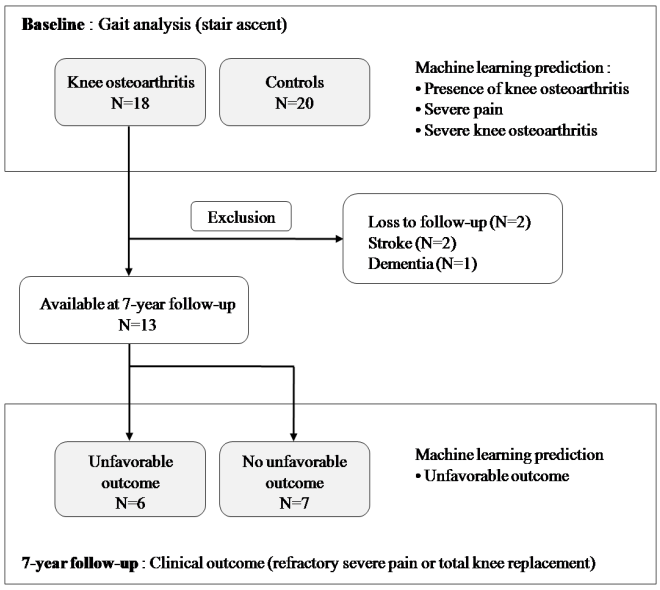

Figure 1. Flow chart of the prediction and inclusion or exclusion of study participants

all patients according to the ethical guidelines of the Human Research Ethics Committee of Yonsei University Health System.

Subjects with knee OA were eligible for enrollment if 1) they were female and at least 45 years of age; 2) they had received a diagnosis of bilateral knee OA according to the criteria of the American Rheumatism Association [10]; 3) they had the ability to climb more than five stairs without a brace; 4) they did not receive an intra-articular knee injection in the one month prior to the study; and 5) they had no hip or ankle arthritis. Controls were 1) female and at least 45 years of age; 2) they had no symptoms in either knee; and 3) they had no trauma or other medical history related to the knee. Both groups had no history of surgical intervention to either leg or neurologic disorder, including stroke and dementia.

Pain in patients with knee OA was assessed using the Western Ontario and McMaster University Osteoarthritis Index (WOMAC) [11]. The WOMAC uses five questions about pain, two questions about stiffness, and 17 questions about physical function. The degree of difficulty due to knee OA was rated on a scale from 0 (none) to 4 (extreme) for each question. Radiographic severity of knee OA was assessed based on Kellgren-Lawrence (KL) grade ranging from 1 to 4 [12]. Based on the WOMAC pain score, we divided the patient group into subgroups with less severe pain (WOMAC pain score≤6) and severe pain (WOMAC pain score>6) using the mean pain score of 6.1 as shown in Table I. Based on KL grade, we also divided the patient group into subgroups with less severe OA (KL grade≤2) and more severe OA (KL grade>2).

B. Follow-up Data Collection

To identify patients with unfavorable outcomes, we examined hospital and outpatient visit records at Severance Hospital between January 2005 and December 2011. During this follow-up period, only clinical outcome without movement analysis was obtained because many patients did not agree with the further gait analysis. At baseline, this observational follow-up consisted of 18 patients with knee OA who underwent gait analysis before 2005. Similarly to Kastelein [13], we defined unfavorable outcomes as refractory severe pain at seven-year follow-up or having undergone total knee replacement during the follow-up period. A total of 13 patients with knee OA were available for seven-year follow-up. Six of 13 patients had unfavorable outcomes at the seven-year follow-up visit. Two patients complained of severe knee pain and four patients had undergone total knee replacement during follow-up.

C. Gait Analysis

Movement during stair-ascent trials was assessed using the Vicon 370 Motion Analysis System (Oxford Metrics Inc., Oxford, U.K.). This system consists of six cameras and one force plate (Klistler Inc., Winterthur, Switzerland) placed at the first step. A total of 13 reflective markers were attached to each participant according to the Vicon protocol provided by the manufacturer. The staircase consisted of four steps. The step dimensions were designed with height of 15 cm, tread of 28 cm, and width of 100 cm according to the Korean Regulations on Standards of Housing Construction. We defined the starting point of stair ascent as the first foot contact with the first step. Similarly, we defined the endpoint as foot contact with the last stair step. Analysis of the kinematic data was performed using Vicon Polygon 2.0. Finally, we investigated the association between kinematic factors during stair ascent and pain, radiographic severity, and prognosis of knee OA in elderly women using the SVM.

D. Data Analysis

To test how well movement during stair-ascent trials indicated diagnostic detection, pain, radiographic severity, and prognosis of knee OA, we evaluated the predictive accuracy of SVM using kinematic factors using the leave-one-out cross-validation (LOOCV) procedure, which is suitable for the evaluation of small sample sizes [14]. The area under the receiver operating characteristics curve (AUC-ROC), sensitivity, and specificity were also calculated. Kinematic factors included the time of stair ascent and kinematic data, including angular features of the pelvis, hip, knee, and ankle, and are shown in Table I [15]. The LOOCV was used to validate the performance of the SVM method and the methods proposed by Miyazaki [3] and Kamruzzaman [16] in previous studies.

The goal of assessing training SVM was to maximize the sum of the predictive accuracy for four clinical problems such as diagnostic detection of knee OA, severe pain, severe radiographic OA, and unfavorable outcomes in knee OA. To obtain optimal ML method performance, we adopted a grid search in which a range of parameter values was tested using the LOOCV [6]. The optimal model of SVM was found using a Gaussian kernel function with a penalty parameter C of 10 and a scaling factor σ of 50. To overcome high dimensionality, variable selection was necessary in order to achieve effective prediction. We adopted backward elimination as a feature selection method for consistency subset evaluation [17].

Therefore, starting from all kinematic factors, the variables were removed one at a time until the accuracy of the training ML method did not improve. We determined the order of the variables with the embedded method of SVM [18].

We used MATLAB 2010a (Mathworks Inc., Natick, MA) for the analysis of ML and SPSS 18.0 (SPSS Inc., Chicago, IL) for statistical analysis. Due to the small sample size, the differences and relationship between variables were analyzed by the Mann-Whitney U-test and the Spearman's rank correlation coefficient with a significance level of 0.05, respectively. To compare the performance of the models, we plotted the performance in a ROC space due to the limitation for generating a ROC curve in cross validation as the previous study did [19].

III. RESULTS

A. Baseline Characteristics and Movement Data

Table I shows the baseline characteristics of participants in the knee OA group and the control group. The time required for stair ascent was higher in the knee OA group. In comparison to the control group, the angles of maximum anterior pelvis tilting, minimal hip flexion, maximal hip flexion, minimal knee flexion, and maximal ankle plantar flexion were significantly higher in the knee OA group. The angles of knee flexion and ankle dorsiflexion at initial foot contact were significantly lower in the knee OA group.

TABLE I. BASELINE CHARACTERISTICS AND KINEMATIC DATA ON STAIR ASCENT

	Baseline		
	Knee OA (N=18)	**Normal control (N=20)**	***P*-value**
Age (years)	65.3±9.6	61.0±5.6	0.095
Height (cm)	154.2±5.5	153.8±5.5	0.820
Weight (kg)	58.2±8.1	58.6±7.3	0.889
BMI (kg/m^2)	24.5±2.9	24.7±2.4	0.761
WOMAC total	16.6±7.3	-	-
Pain	6.1±3.1	-	-
Stiffness	0.8±0.9	-	-
Physical	9.3±5.4	-	-
Kellgren-Lawrence grade			
1	2 (11.1)	-	-
2	7 (38.9)	-	-
3	8 (44.4)	-	-
4	1 (5.6)	-	-
Time during stair ascent (sec)**	7.2±1.5	5.1±0.7	<0.001
Kinematic data (degrees)			
Maximal pelvis anterior tilting**	30.1±5.4	23.6±4.6	<0.001
Hip Fl at IC	75.1±6.5	71.4±5.0	0.059
Minimal hip Fl**	27.5±6.4	18.0±5.5	<0.001
Maximal hip Fl**	85.9±5.6	77.2±5.3	<0.001
Knee Fl at IC**	63.6±7.92	69.4±3.0	<0.001
Minimal knee Fl**	22.4±2.3	18.4±3.3	<0.001
Maximal knee Fl	99.5±7.39	99.7±5.24	0.926
Ankle DF at IC**	10.2±5.6	17.7±3.9	<0.001
Maximal ankle DF	26.1±4.6	25.6±4.6	0.707
Maximal ankle PF*	14.5±5.9	9.9±4.9	0.014

Data are presented as mean ± SD or n (%)
**P*-value<0.05, ***P*-value<0.001
OA: osteoarthritis, BMI: body mass index, Fl: flexion, IC: initial foot contact, DF: dorsiflexion, PF: plantar flexion

B. Prediction Performance

Table II summarizes the results of the LOOCV and variable selection used in the SVM and the methods in previous studies. Gray shadow represents the original implementation proposed by Miyazaki [3] and Kamruzzaman [16]. We applied this method to obtain other results. SVM performed best in the detection of knee OA with an accuracy of 97.4% (AUC: 0.972, 95% CI: 0.789-0.988) in patients and controls. SVM also performed best in the prediction of severe pain (accuracy: 83.3%, AUC: 0.833, 95% CI: 0.476-0.955) and radiographic severity (accuracy: 83.3%, AUC: 0.833, 95% CI: 0.746-0.955) among patients with knee OA. The predictors of the SVM model included time for stair ascent, maximum anterior pelvis tilting, knee flexion at initial foot contact, and ankle dorsiflexion at initial foot contact. Fig. 2 shows the performance of the prediction models in a ROC space using the LOOCV. While SVM performed better in the detection of knee OA, severe pain, and severe radiographic OA than other methods in previous studies, there was no difference between the method by Miyazaki [3] and SVM in the prediction of unfavorable outcomes (accuracy: 69.2%, AUC: 0.702, 95% CI: 0.277-0.950).

Pain and radiographic severity, which are well known features of knee OA, are indications for total knee replacement surgery. To elucidate why SVM performed better than previous methods, we studied the relationship between knee pain or radiographic severity and predictors of SVM. Knee pain was associated with time required for stair ascent (r = 0.489, P = 0.020) and ankle dorsiflexion at initial foot contact (r = -0.403, P = 0.048). Similarly, radiographic

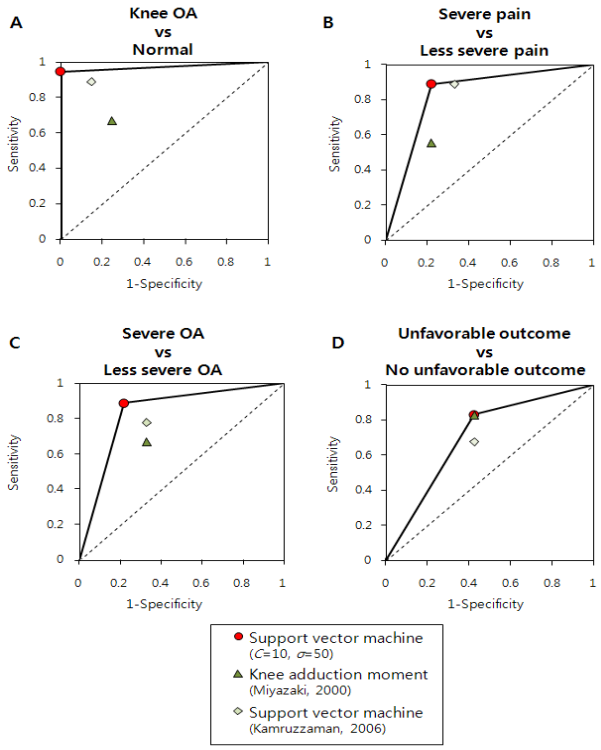

Figure 2. Classifier performance in the ROC space using leave-one-out cross-validation to detect (A) knee osteoarthritis, (B) severe pain (WOMAC pain score>6), (C) more sever osteoarthritis (Kellgren-Lawrence grade>2), and (D) unfavorable outcome (e.g. having undergone total knee replacement)

TABLE II. LEAVE-ONE-OUT CROSS-VALIDATION RESULTS AND FEATURES SELECTED BY CONSISTENCY SUBSET EVALUATION FOR SUPPORT VECTOR MACHINE (SVM)

Method	Performance	Knee OA vs Normal	Severe pain[1] vs Less severe pain	More severe OA[2] vs Less severe OA	Unfavorable outcome[3] vs No unfavorable outcome	Selected features
SVM (C=10, σ=50)	AUC	0.972	0.833	0.833	0.702	Time during stair ascent
	Acc (%)	97.4	83.3	83.3	69.2	Maximum pelvis anterior tilting
	Sen (%)	94.4	88.9	88.9	83.3	Knee Fl at IC
	Spe (%)	100.0	77.8	77.8	57.1	Ankle DF at IC
Knee adduction moment (Miyazaki, 2002)	AUC	0.708	0.667	*0.667	0.702	Maximal knee Ad moment
	Acc (%)	71.1	66.7	*66.7	69.2	
	Sen (%)	66.7	55.6	*66.7	83.3	
	Spe (%)	75.0	77.8	*66.7	57.1	
SVM (Kamruzzaman, 2006)	AUC	†0.869	0.778	0.722	0.619	Stride length
	Acc (%)	†86.8	77.8	72.2	61.5	Cadence
	Sen (%)	†88.9	88.9	77.8	66.7	Leg length
	Spe (%)	†85.0	66.7	66.7	57.1	Age

1 Severe pain was defined as WOMAC pain score>6, and less severe pain was defined as WOMAC pain score≤6.
2 Severe OA was defined as Kellgren-Lawrence grade>2, and less severe OA was defined as Kellgren-Lawrence grade≤2.
3 Unfavorable outcome was defined as reported persistent knee pain at 7-year follow-up or having undergone total knee replacement during follow-up in patients with knee OA.
*Gray shadow represents the original implementation proposed by Miyazaki. We applied this method to obtain other results.
†Gray shadow represents the original implementation proposed by Kamruzzaman. We applied this method to obtain other results.
OA: osteoarthritis, SVM: support vector machines, Acc: accuracy, Sen: sensitivity, Spe: specificity, Fl: flexion, IC: initial foot contact, DF: dorsiflexion, Ad: adduction

was associated with maximum anterior pelvis tilting (r = 0.449, P = 0.031) and knee flexion at initial foot contact (r = -0.492, P = 0.019). There was no significant relationship between pain and radiographic severity (r = 0.232, P = 0.354).

IV. DISCUSSION AND CONCLUSION

This cross-sectional study used ML analysis of movement to predict detection of knee OA, severe pain, severe radiographic OA, and poor prognosis at seven-year follow-up. We hypothesized that ML could assist in decision-making for knee OA management when a patient has knee symptoms during activities such as stair ascent. This study confirmed the relationship between movement during stair ascent and knee OA related factors in the prediction of pain and radiographic severity.

SVM differentiated patients with varying stages of knee OA from normal controls with very high accuracy of 97.4%. SVM also showed high accuracy of 83.3% in predicting of radiographic severity. These results indicate that an early stage of knee OA could be detected using this method. Since recent studies suggest that an early diagnosis of OA before irreversible degenerative changes occur is important, our methods could be used to prescreen for knee OA. To our knowledge, this is the first study to investigate the use of kinematic and long-term follow-up data on movement during activities such as stair ascent to predict pain, radiologic severity, and prognosis of knee OA using ML.

Because the SVM model considers input variables based on their own characteristics of nonlinearity and high dimension [16], the SVM model dealt with a separating space consisting of variables in high dimension and was thus able to consider all variables, improving its performance in predicting pain and radiographic severity of knee OA. One of the disadvantages of SVM was that it required many parameters to construct an optimal SVM model. Since there is no reliable method for selecting the optimal penalty parameter, C, and scaling factor, σ, of the Gaussian kernel function, we carried out an exhaustive grid search by changing the parameters.

There were several limitations to this study. First, since the sample size was too small to accurately test the hypothesis it was difficult to separate the data into training and test sets for obtaining consistent results. Therefore, we adopted the LOOCV. Second, this was a cross-sectional study. This meant that we assessed movement once at baseline, but did not assess movement changes at follow-up because many patients did not agree with the further gait analysis. Third, we did not evaluate socio-economic factors of lifestyle. Previous studies have shown that a lower economic and educational level is a significant risk factor for knee OA [13]. Finally, our inclusion criteria for unfavorable outcomes may have been affected by the treatment choice of each physician.

In conclusion, the ML method could contribute to the advancement of clinical decision-making tools and to our understanding of risk factors for knee OA progression. Assessment of time of stair ascent, maximum anterior pelvis tilting, knee flexion at initial foot contact, and ankle dorsiflexion at initial foot contact during stair ascent using SVM may be helpful in the detection of knee OA and in the evaluation of pain and radiographic severity. Further studies should be conducted with the goal of developing an extended prediction model for progressive knee OA. In addition, future studies are warranted to replicate using other machine learning methods such as artificial neural networks, random forest, and logistic regression.

REFERENCES

[1] G. Peat, R. McCarney, and P. Croft, "Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care," *Ann Rheum Dis*, vol. 60, pp. 91–97, Feb 2001.

[2] D. T. Felson, "Osteoarthritis of the knee," *N Engl J Med*, vol. 354, pp. 841-848, Feb 2006.

[3] T. Miyazaki, M. Wada, H. Kawahara, M. Sato, H. Baba, and S. Shimada, "Dynamic load at baseline can predict radiographic disease progression in medial compartment knee osteoarthritis," *Ann Rheum Dis*, vol. 61, pp. 617–622, Jul 2002.

[4] M. Guo, M. J. Axe, and K. Manal, "The influence of foot progression angle on the knee adduction moment during walking and stair climbing in pain free individuals with knee osteoarthritis," *Gait Posture*, vol. 26,

pp. 436–441, Sep 2007.
[5] C. R. Chu, A. A. Williams, C. H. Coyle, and M. E. Bowers, "Early diagnosis to enable early treatment of pre-osteoarthritis," *Arthritis Res Ther*, vol. 14, pp. 212, Jun 2012.
[6] P. Levinger, D. T. Lai, R. K. Begg, K. E. Webster, and J. A. Feller, "The application of support vector machines for detecting recovery from knee replacement surgery using spatio-temporal gait parameters," *Gait Posture*, vol. 29, pp. 91–96, Jan 2009.
[7] P. Larrañaga, B. Calvo, R. Santana, C. Bielza, J. Galdiano, I. Inza, J. A. Lozano, R. Armañanzas, G. Santafé, A. Pérez, and V. Robles, "Machine learning in bioinformatics," *Brief Bioinform*, vol. 7, pp. 86–112, Feb 2006.
[8] C. Cortes and V. Vapnik, "Support-vector networks," *Mach Learn*, vol. 20, pp. 273-297, Sep 1995.
[9] D. Y. Kim, C. I. Park, W. H. Chang, T. H. Park, S. Y. Ahn, S. K. Lee, and D. S. Lee, "Characteristics of stair ascent in patients with knee osteoarthritis," *J Korean Acad Rehabil Med*, vol. 29, pp. 654–661, Dec 2005.
[10] R. Altman, E. Asch, D. Bloch, G. Bole, D. Borenstein, K. Brandt, W. Christy, T. D. Cooke, R. Greenwald, M. Hochberg, D. Howell, D. Kaplan, W. Koopman, S. Longley 3rd, H. Mankin, D. J. Mcshane, T. Medsger, R. Meenan, W. Mikkelsen, R. Moskowitz, W. Murphy, B. Rothschild, M. Segal, L. Sokoloff, and F. Wolfe, "Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee," *Arthritis Rheum*, vol. 29, pp. 1039–1049, Aug 1986.
[11] C. Jinks, K. Jordan, and P. Croft, "Measuring the population impact of knee pain and disability with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)," *Pain*, vol. 100, pp. 55–64, Nov 2002.
[12] J. H. Kellgren and J. S. Lawrence, "Radiological assessment of rheumatoid arthritis," *Ann Rheum Dis*, vol. 16, pp. 485–493, Jun 1957.
[13] M. Kastelein, P. A. Luijsterburg, J. N. Belo, J. A. Verhaar, B. W. Koes, and S. M. Bierma-Zeinstra, "Six-year course and prognosis of nontraumatic knee symptoms in adults in general practice: A prospective cohort study," *Arthritis Care Res*, vol. 63, pp. 1287–1294, Sep 2011.
[14] R. Simon, M. D. Radmacher, K. Dobbin, and L. M. McShane, "Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification," J NatI Cancer Inst, vol. 95, pp. 14–18, Jan 2003.
[15] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," *J Biomech*, vol. 38, pp. 401–408, Mar 2005.
[16] J. Kamruzzaman and R. K. Begg, "Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait," *IEEE Trans Biomed Eng*, vol. 53, pp. 2479–2490, Dec 2006.
[17] M. Dash and H. Liu, "Consistency-based search in feature selection," *Artif Intell*, vol. 151, pp. 155-176, Dec 2003.
[18] C. H. Hsieh, R. H. Lu, N. H. Lee, W. T. Chiu, M. H. Hsu, and Y. C. Li, "Novel solutions for an old disease: diagnosis of acute appendicitis with random forest, support vector machines, and artificial neural networks," *Surgery*, vol. 149, pp. 87-93, Jan 2011.
[19] D. Agranoff, D. Fernandez-Reyes, M. C. Papadopoulos, S. A. Rojas, M. Herbster, A. Loosemore, E. Tarelli, J. Sheldon, A. Schwenk, R. Pollok, C. F. Rayner, and S. Krishna, "Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum," *Lancet*, vol. 368, pp. 1012-1021, Sep 2006.