Title: Classifying individuals with and without Patellofemoral Pain Syndrome using ground force profiles – development of a method using functional data boosting.

Running title: Knee pain classification using force profiles

Article type: Original Article

Bernard X.W. Liew, PhD 1
David Rugamer, PhD 2,3
Deepa Abichandani, MSc 4
Alessandro Marco De Nunzio, PhD 5

1 School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, Essex, CO4 3SQ, United Kingdom
2 Department of Statistics, Ludwig-Maximilians-Universität München, Germany
3 Chair of Statistics, School of Business and Economics, Humboldt University of Berlin, Germany
4 Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Edgbaston B152TT, United Kingdom
5 LUNEX International University of Health, Exercise and Sports, 50, avenue du Parc des Sports, L-4671 Differdange, Luxembourg

Address correspondence and reprint requests to Dr Bernard Liew, School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, Essex, CO4 3SQ, United Kingdom; E-mail: bl19622@essex.ac.uk, liew_xwb@hotmail.com

Source of funding: None

Conflict of interest: All authors declare that they have no conflicts of interest.

Highlights

- Prognostic models are critical for the management of patellofemoral pain syndrome
- Many physiological predictors are functional in nature
- Functional predictors cannot be used in traditional modelling methods
- Simple ground reach force signatures provide excellent classification performance
- FDboost may be useful for modelling when functional predictors are involved
Abstract

Background

Predictors of recovery in patellofemoral pain syndrome (PFPS) currently used in prognostic models are scalar in nature, despite many physiological measures originally lying on the functional scale. Traditional modelling techniques cannot harness the potential predictive value of functional physiological variables.

Research question

What is the classification performance of PFPS status of a statistical model when using functional ground reaction force (GRF) time-series?

Methods

Thirty-one individuals (control = 17, PFPS = 14) performed maximal countermovement jumps, on two force plates. The three-dimensional components of the GRF profiles were time-normalized between the start of the eccentric phase and take-off, and used as functional predictors. A statistical model was developed using functional data boosting (FDboost), for binary classification of PFPS statuses (control vs PFPS). The area under the Receiver Operating Characteristic curve (AUC) was used to quantify the model’s ability to discriminate the two groups.

Results

The three predictors of GRF waveform achieved an average out-of-bag AUC of 93.7%. A 1% increase in applied medial force reduced the log odds of being in the PFPS group by 0.68 at 87% of jump cycle. In the AP direction, a 1% reduction in applied posterior force increased the log odds of being classified as PFPS by 1.10 at 70% jump cycle. For the vertical GRF, a
1% increase in applied force reduced the log odds of being classified in the PFPS group by 0.12 at 44% of the jump cycle.

**Significance**

Using simple functional GRF variables collected during functionally relevant task, in conjunction with FDboost, produced clinically interpretable models that retain excellent classification performance in individuals with PFPS. FDboost may be an invaluable tool to be used in longitudinal cohort prognostic studies, especially when scalar and functional predictors are collected.

**Key words:** Patellofemoral pain syndrome; Jumping; Biomechanics; Machine learning; Functional regression.
1. Introduction

Patellofemoral pain syndrome (PFPS) is a common musculoskeletal disorder, with a prevalence as high as 22.7% [1]. Individuals with PFPS commonly present with retro or peri-patellar pain during activities which incur high patellofemoral joint (PFJ) loads [2, 3]. Despite receiving evidence-based interventions, up to 40% of patients have persistent symptoms 12 months later [4]. Being able to predict who are at risk of poor outcomes will guide clinical expectations of recovery and assist clinicians in matching different clinical phenotypes to specific interventions.

Predicting the course of PFPS typically requires predictive models, and the development and validation of such models is termed as “prognostic model research” [5]. A predictive model contains the best combination of predictors needed to achieve the best predictive accuracy. Predictors of poor clinical outcomes in PFPS can come from various sources, such as an individual’s demographic characteristic, neuromuscular and biomechanical function [6]. Many clinical predictors of poor outcomes, such as baseline pain, can be considered as a scalar variable – meaning they only reflect magnitude. However, since most neuromuscular and biomechanical variables (herein termed broadly as physiological variables) are collected during movement, physiological predictors are often functional in nature – meaning they reflect magnitude over time and/or space.

The statistical approach typically used in predictive modelling is logistic regression [6], where only scalar predictors can be used. If a predictor was originally collected on a functional scale, such as ankle eversion angle over a gait cycle, then it must first be discretized into a singular value (e.g. taking the peak value) [7]. Increasingly, investigations have reported physiological differences between individuals with and without PFPS over different periods of movement, and not restricted only to differences in peak values [8]. Also,
given that some static assessments of posture have been criticized as being weakly correlated with dynamic movements [9], the discretization of functional into scalar variables may be suboptimal when incorporating physiological predictors into prediction models. One common strategy for accommodating functional predictors in predictive modelling is by dimension reduction (e.g. principal components analysis (PCA)), followed by using the principal component (PC) coefficients as the “new” predictors [10, 11]. A disadvantage of using the PC coefficients as predictors, is that interpreting the final statistical model’s solutions is more challenging, than if the model was built using the original functional predictors.

Recent advancements in machine learning techniques have enabled the simultaneous incorporation of functional and scalar predictors into statistical models for predictive modelling. One such technique is functional data boosting (FDboost) [12], a composite technique which combines functional regression with component-wise boosting. The idea of boosting is to train a strong ensemble model by combining weaker and simpler models, with the added model trying to correct the prediction errors made by the preceding model. An advantage of FDboost over contemporary machine learning techniques in biomechanics [10, 11], is that since the functional predictors remain on the original scale, clinical interpretation of the ensuing model is more straightforward than when transformed predictors are used.

Prediction models of clinical outcomes ultimately require building a model on a prospective cohort study and validating the model’s performance on an independent prospective cohort. Prior to the conduct of a more costly longitudinal study, it is wise to evaluate a novel predictive modelling technique on a cross-sectional cohort of individuals with and without PFPS. Hence, the aim of the present study was to develop and evaluate the performance of FDboost using simple ground reaction force (GRF) waveforms collected during countermovement jumps (CMJ), as predictors for classifying the presence of PFPS. The gluteal and quadriceps are important muscles for generating medial-lateral (ML) and
vertical GRF [13], and impairments to these muscles have been reported in individuals with PFPS [14, 15]. Hence, we hypothesized that the vertical and ML GRF variables would emerge as important discriminators of individuals with and without PFPS.

2. Methods

2.1. Design and participants

Male and female participants were eligible for inclusion in the PFPS group if they were: 1) between 18 - 45 years old; 2) ≥ 6 points on the SNAPPS questionnaire (Survey instrument for Natural history, Aetiology and Prevalence of Patellofemoral pain Studies) [16]; 3) have a minimum knee pain intensity of 3/10 on the visual analogue scale (VAS) during at least two of the following activities – jumping, running, squatting, prolonged sitting, or stair climbing. Participants were eligible to be included in the control group if they had no anterior knee pain within the past 12 months. Participants were excluded from the study if they had 1) knee pain from an acute injury, patellar tendinopathy, iliotibial band syndrome, ligamentous, or degenerative pathology; 2) history of a traumatic patellar dislocation; 3) previous knee surgeries within the past 12 months; and 4) females currently pregnant. Ethical clearance was obtained from the Ethics Committee of University of Birmingham, United Kingdom (MCR041218-1). All participants provided written informed consent prior to study enrolment.

The following measures were collected to characterize the nature of pain for individuals with PFPS: current pain intensity on a visual analogue scale (0 no pain-10 maximum pain), current knee related function using the Knee Injury & Osteoarthritis scale (KOOS) [17], and an added KOOS patellofemoral subscale (KOOS-PF).

2.2. Motor tasks
Countermovement jumps (CMJ) were performed on two 60 x 40 cm in-ground force plates sampling at 500 Hz (BTS P6000, BTS Bioengineering, Italy), in their own comfortable exercise attire and sporting shoes. Participants stood with one foot on each force plate, with their arms fixed at 90° abduction to minimize the influence of arm swing on jumping mechanics. Participants were asked to perform three trials of maximal CMJ with a one minute of rest provided between each trial. The depth reached during the countermovement phase was self-determined and practised by each participant.

2.3. Processing

GRF data were low-pass filtered at 75Hz (4th order, zero-lag, Butterworth), time-normalised to 101 data points (cycle) between the start of the eccentric phase (drop in vertical GRF > 2.5% of body weight (BW) ) and toe-off (vertical GRF < 20 N), and scaled to each individual’s static standing weight (N). The mean GRF variables over three CMJ trials were derived for each participant, resulting in six GRF predictors for each participant (three anatomical directions and two sides). For the ML GRF, a positive value reflects a medially directed force; for the antero-posterior (AP) GRF positive values reflect an anterior force, and in the vertical direction positive values reflect a proximal upward force. To reduce high collinearity in the predictors, only GRF variables from one side (right or left) were selected. For healthy controls and individuals with bilateral PFPS, GRF variables from the right side were selected. For individuals with unilateral PFPS, GRF variables from the side of pain were selected. This resulted in three functional predictors serving as input for each participant.

2.4. Statistical learning

A scalar-on-function (SoFR) logistic regression model was used for binary classification. A SoFR model is one where the response variable takes on scalar values, and the predictors take on functional values. All three functional variables were demeaned as a
pre-processing step, so that different predictors had equal potential to be included in the model. We used component-wise gradient boosting for model fitting [12]. The algorithm is an iterative procedure which successively adds one predictor to the model with the ability to handle functional predictors, perform variable selection, and allow for penalized estimation. The order of predictor entry into the model is dependent on which is the best predictor at each iteration.

To estimate the optimal number of iterations, cross-validation was performed on 25 bootstrap samples of the data, each with a roughly similar ratio of individuals in each group. In each bootstrap sample, some participants will be represented multiple times while others will not be selected at all. The samples not selected are referred to as the “out-of-bag” (OOB) samples. For each iteration of bootstrap resampling, a model is built on the selected samples, and the area under the Receiver Operating Characteristic curve (AUC) was used to quantify the model’s ability to discriminate the two groups on the OOB sample.

The results of FDboost is best illustrated with several types of graphical plot. The principle plot is that of the $\beta$ coefficient time-series for each predictor. Like a standard logistic regression, the $\beta$ coefficient reflects the increase in log odds of being in one group given a unit change of the predictor. Given the functional nature of our predictors, the $\beta$ coefficient plots reflect the increase in log odds of being in the PFPS group given a unit change in the predictor at each 100 time-normalized points. Another useful plot is the partial dependence plot, which shows the marginal effect one predictor have on the probability of being in the PFPS or control group. In this instance, we simulated two “new” individuals, one with and another without PFPS, both jumping with GRF values which reflected the average GRF values of the group with and without PFPS, respectively. The point-by-point product between the $\beta$ coefficient and the predictor value is taken, summed over all 100 time-
normalized points, and transformed to probabilities of being in the PFPS and control groups (see equation).

The duration of the CMJ, defined from the start of the eccentric phase to toe-off was quantified, and compared between groups with a two-sample t-test, with significance defined as $P < 0.05$. All analyses were performed using R version 3.5.3, using the “FDboost” package [12], and the codes with accompanying data are included in the supplementary material.

3. Results

Descriptive statistics of the demographic data can be found in Table 1. The group average GRF plots are found in Figure 1. The jump duration of individuals with PFPS was significantly longer than healthy controls (PFPS: mean [SD] 1.01s [0.29s] vs control: 0.80s [0.19s], $t = -2.47$, $P = 0.02$). The three predictors of GRF waveform achieved an out-of-bag AUC of 93.7% (2.7%). The final model in the application is:

$$ P (\text{group}_i = \text{PFPS}) = \text{Logit}^{-1}(\beta_0 + \int x_{11}(t)\beta_1(t)dt + \int x_{12}(t)\beta_2(t)dt + \int x_{13}(t)\beta_3(t)dt) $$

for participants $i = 1, ..., 31$ where $\beta_0$ is the intercept of -0.097, $\beta_1(t)$, $\beta_2(t)$ and $\beta_3(t)$ are the coefficients of the three functional predictors (Figure 2). An example of an interpretation of Figure 2 is as follows. A 1% increase in applied medial force reduced the log odds of being in the PFPS group by 0.68 at 87% of jump cycle (Figure 2). In the AP direction, a 1% reduction in applied posterior force increased the log odds of being classified as PFPS by 1.10 at 70% jump cycle (Figure 2). For the vertical GRF, a 1% increase in applied force reduced the log odds of being classified in the PFPS group by 0.12 at 44% of the jump cycle (Figure 2).
Table 1. Mean (standard deviation) of patient and pain characteristics.

<table>
<thead>
<tr>
<th>Variables</th>
<th>PFPS (n = 14)</th>
<th>Control (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>20.86 (1.83)</td>
<td>23.47 (2.67)</td>
</tr>
<tr>
<td>Sex</td>
<td>6M, 8F</td>
<td>9M, 8F</td>
</tr>
<tr>
<td>Painful side</td>
<td>6R, 4L, 4Bilateral</td>
<td>-</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.71 (0.10)</td>
<td>1.70 (0.08)</td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>64.96 (10.51)</td>
<td>67.02 (10.87)</td>
</tr>
<tr>
<td>Pain VAS (0 no pain-10 max pain)</td>
<td>3.71 (2.02)</td>
<td>-</td>
</tr>
<tr>
<td>KOOS-adl (0 indicating extreme symptoms-100 no symptoms)</td>
<td>85.29 (17.83)</td>
<td>100 (0)</td>
</tr>
<tr>
<td>KOOS-pain (0 indicating extreme symptoms-100 no symptoms)</td>
<td>74.60 (16.12)</td>
<td>98.69 (2.62)</td>
</tr>
<tr>
<td>KOOS-qol (0 indicating extreme symptoms-100 no symptoms)</td>
<td>58.04 (17.24)</td>
<td>97.43 (6.65)</td>
</tr>
<tr>
<td>KOOS- sports (0 indicating extreme symptoms-100 no symptoms)</td>
<td>68.93 (26.90)</td>
<td>98.53 (3.43)</td>
</tr>
<tr>
<td>KOOS-symptoms (0 indicating extreme symptoms-100 no symptoms)</td>
<td>71.17 (16.42)</td>
<td>96.85 (5.19)</td>
</tr>
<tr>
<td>KOOS-pf (0 indicating extreme symptoms-100 no symptoms)</td>
<td>68.99 (18.90)</td>
<td>99.47 (1.71)</td>
</tr>
</tbody>
</table>

Abbreviations: VAS – visual analogue scale; KOOS - Knee Injury & Osteoarthritis Outcome; adl – activities of daily living; qol – quality of life; pf – patellofemoral sub-scale
Figure 1
Figure 2
To visualize the application of the trained model, Figure 3 illustrates the cumulative change in probability of being in either PFPS or control group, if two simulated “new” participants were assessed, each with GRF values which reflected the average GRF values of the group with and without PFPS. The greatest increase in probability occurred during the period of 25% to 75% of the jump cycle (Figure 3), with the vertical GRF driving the change in certainty of group classification.

4. Discussion

A potentially useful source of predictors that could be used to develop prognostic models is functional physiological variables, although its predictive utility remains unexplored given the limitations of currently adopted statistical approaches. In partial support of our hypotheses, the vertical GRF was the most important driver, but the ML GRF was the
least important of change in probabilities in the classification of individuals with and without PFPS.

The period where vertical GRF had the greatest influence on outcome probability – i.e. between 25% to 62.5% of the jump cycle, coincides with the eccentric phase of CMJ [18, 19]. During this phase, individuals with PFPS applied lesser vertical GRF than controls to slow the descent of the centre of mass (COM) towards its lowest depth in the CMJ. Previous studies have shown that greater countermovement descent is associated with smaller vertical GRF [18, 19]. This implies that individuals with PFPS perform a self-selected CMJ using a greater countermovement depth than controls. A greater countermovement depth used by individuals with PFPS was supported by a previous study which reported greater peak hip and knee flexion angle during a single leg jump in individuals with PFPS compared to controls [2].

The eccentric phase of the CMJ is used to increase the time over which lower limb extensor muscles can generate force [20]. Based on a muscle’s force-velocity relationship [21], a longer duration over which tension can be built means less shortening velocity, and more force being generated. Evidenced by the jump duration, individuals with PFPS increase the time over which force is generated by increasing their countermovement depth, potentially to minimize disturbance to vertical impulse (impulse = force * time) generation. This change in movement strategy could be due to a reduced ability to generate maximal force – attributable partly to muscle atrophy [15], and/or a reduced ability to rapidly generate force [14, 22].

GRF reflects the net force exerted by all muscles, indicating that a reduced lower limb strength would result in reduced GRF [23]. Hence, the influence of vertical GRF as a predictor presently has indirect support from the literature. A longitudinal study on runners
with PFPS reported that reduced baseline knee extensor strength predicted poorer self-reported recovery in individuals with PFPS, than those with greater baseline strength [6]. Interestingly, another longitudinal study reported that changes to knee extensor strength had no relationship with self-reported change to function and pain in individuals with PFPS [24]. The inconstant predictive potential of lower limb strength indices on recovery in PFPS could be due to most assessments of strength being undertaken at a single joint angle (i.e. isometric test). Quantifying strength at a single joint angle may not be a sensitive tool of quantifying neuromuscular impairment, where deficits at other angles may be masked [25]. However, such static neuromuscular analyses have the advantage of providing scalar results that can be used in traditional regression techniques, such as logistic regression. On the contrary, the methods presented has the advantage of rapidly providing a dynamic measure of global lower limb force capacity, and that the entire functional measure can be used directly during statistical modelling.

Although the vertical direction is the largest of the three GRF components, the fact that FDboost selected the horizontal (ML and AP) components in the model, suggests the importance of assessing lower limb neuromuscular function in three-dimensions. Alterations to joint torques are more sensitive to small changes in horizontal GRF components, compared to the vertical component [26]. This is because the lever arm from the horizontal GRF components to the joint centres are greater than that of the vertical component [26], necessitating greater change to joint torque for a unit horizontal GRF change than vertical GRF change. The importance of considering the predictive influence of horizontal force components is also supported by a wealth of studies (e.g. reviewed in [27]) which reported non-sagittal plane biomechanical and neuromuscular differences between individuals with and without PFPS. Interestingly, Boling et al. [28] reported the important influence of hip
external rotation torque in influencing the risk of PFPS onset, although the study did not include horizontal GRF components during model development.

Despite the excellent classification performance, the cross-sectional nature of this study means that extrapolating the results to predicting longitudinal recovery outcomes should be done with caution. In defence, the aim of the present study was to provide preliminary pilot evidence in the use of a novel machine learning technique, FDboost, at integrating functional biomechanical variables into statistical prediction models, rather than on developing and validating a formal clinical prediction tool. The results of the present study have two significant scientific implications. First, similar to the use of knee extensor strength [6], functional GRF variables should be considered as candidate predictors of PFPS recovery for use when developing prognostic models in longitudinal studies. It remains to be investigated if GRF predictors would be selected once inclusion of a thorough set of biopsychosocial predictors are considered. Second, FDboost is useful as a statistical technique in prognostic studies, as it has the capacity to integrate both traditional scalar, and novel functional, predictors into the model. Also, given that the predictors in FDboost lie on their original scale with meaningful units, unlike using PC coefficients as predictors as an example, the mapping between a change in predictor against a change in outcome can be easily determined.

A limitation of the present study was the relatively small sample size, which precluded splitting the data into a training and an independent validation dataset. The number of participants in the present study was however, comparable to other research within the biomechanics-machine learning literature (e.g. n = 41 in [29]). The present results will enable future researchers to fit the current model’s learning curve to inverse power law models [30], to estimate the sample size needed to achieve a desired prediction performance (see supplementary material for example sample size calculation). Another limitation is that the
underlying neuromechanical impairments that drive the observed differences in GRF variables are unknown. Although using GRF cannot identify the specific muscle or joint impairments, it can inform clinicians of the dominant anatomical plane of impairment. Since sagittal plane GRF variables (vertical and AP) were the two most influential predictors in the present study, this would mean clinicians should focus their investigative efforts to muscle and joint impairments in this plane (e.g. quadriceps dysfunction).

5. Conclusion

Future prognostic studies in PFPS may benefit from a quick method of assessing dynamic lower-limb force generating capacity, by measuring GRF during jumping. Our approach of using simple functional GRF variables collected during functionally relevant task, in conjunction with FDboost, produced clinically interpretable models that retain excellent classification capability. The prognostic utility of using functional physiological variables as candidate predictors of recovery in PFPS needs to be validated in prospective cohort studies.

**Figure captions**

**Figure 1:** Group mean ground reaction force (GRF) during countermovement jumps (a) control group, (b) patellofemoral pain group. Abbreviations: ML = medial lateral, AP = anterior posterior, Vert = vertical.

**Figure 2:** Beta coefficients (log odds) at each 1% cycle of selected predictors in the model. Abbreviations: ML = medial lateral, AP = anterior posterior, Vert = vertical.

**Figure 3:** Predicted cumulative probability of being in the PFPS group given an input of each group’s (con and PFPS) average waveform for each selected predictor. Abbreviations: ML = medial lateral, AP = anterior posterior, Vert = vertical, con = control, pfps = patellofemoral pain syndrome.
REFERENCES


