Classifying neck pain status using scalar and functional biomechanical variables – development of a method using functional data boosting

Bernard X.W. Liew, David Rugamer, Almond Stocker, Alessandro Marco De Nunzio

Pll: S0966-6362(19)31773-4
DOI: https://doi.org/10.1016/j.gaitpost.2019.12.008
Reference: GAIPOS 7399
To appear in: Gait & Posture
Received Date: 12 June 2019
Revised Date: 9 October 2019
Accepted Date: 1 December 2019


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.
Title: Classifying neck pain status using scalar and functional biomechanical variables – development of a method using functional data boosting

Article type: Original Article

Bernard X.W. Liew, PhD\textsuperscript{a}, David Rugamer, PhD\textsuperscript{b}, Almond Stocker, MSc\textsuperscript{c}, Alessandro Marco De Nunzio, PhD\textsuperscript{d,e}

\textsuperscript{a}School of Sport, Rehabilitation and Exercise Sciences, University of Essex, Colchester, Essex, CO4 3SQ, United Kingdom

\textsuperscript{b}Department of Statistics, Ludwig-Maximilians-Universität München, Germany

\textsuperscript{c}Chair of Statistics, School of Business and Economics, Humboldt University of Berlin, Germany

\textsuperscript{d}LUNEX International University of Health, Exercise and Sports, 50, avenue du Parc des Sports, L-4671 Differdange, Luxembourg

\textsuperscript{e}Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Edgbaston B152TT, United Kingdom

Address correspondence and reprint requests to Dr Bernard Liew, S 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

Highlights

- Prognostic models are critical for the management of neck pain disorders
- Biomechanical variables can have a high-dimensionality
- Clinical transferability of models using biomechanical covariates may be limited
- Altered trunk kinematics and greater jerk index are predictors of neck pain status
- FDboost is a useful tool to build prognostic models with biomechanical data

Abstract

Background
Individuals with neck pain have different movement and muscular activation (collectively termed as biomechanical variables) patterns compared to healthy individuals. Incorporating biomechanical variables as covariates into prognostic models is challenging due to the high dimensionality of the data.

**Research question**

What is the classification performance of neck pain status of a statistical model which uses both scalar and functional biomechanical covariates?

**Methods**

Motion capture with electromyography assessment on the sternocleidomastoid, splenius cervicis, erector spinae, was performed on 21 healthy and 26 individuals with neck pain during walking over three gait conditions (rectilinear, curvilinear clockwise (CW) and counterclockwise (CCW)). After removing highly collinear variables, 94 covariates across the three conditions were used to classify neck pain status using functional data boosting (FDboost).

**Results**

Two functional covariates trunk lateral flexion angle during CCW gait, and trunk flexion angle during CW gait; and a scalar covariate, hip jerk index during CCW gait were selected. The model achieved an estimated AUC of 80.8%. For hip jerk index, an increase in hip jerk index by one unit increased the log odds of being in the neck pain group by 0.37. A 1° increase in trunk lateral flexion angle throughout gait alone reduced the probability of being in the neck pain group from 0.5 to 0.15. A 1° increase in trunk flexion angle throughout gait alone increased the probability of being in the neck pain group from 0.5 to 0.9.
Significance

Interpreting the physiological significance of the extracted covariates, with other biomechanical variables, suggests that individuals with neck pain performed curvilinear walking using a stiffer strategy, compared to controls; and this increased the risk of being in the neck pain group. FDboost can produce clinically interpretable models with complex high dimensional data and could be used in future prognostic modelling studies in neck pain research.

**Keywords:** Walking, Biomechanics, Neck pain, Machine learning, Functional regression

1. Introduction

   Neck pain is a common musculoskeletal disorder with up to 50% of adults experiencing neck pain in any given year [1]. Up to 85% of individuals with neck pain continue to report persistent symptoms [1], and some may go on to experience chronic pain. Neck pain also has a significant socio-economic cost with annual healthcare expenditures amounting to US$686 million [2]. Being able to predict the clinical course of neck pain is an important question because it guides clinical expectations of recovery and can help clinicians better match different clinical phenotypes to specific interventions.

   Predicting the course of neck pain requires the use of predictive models, and this type of research is termed as “prognostic model research” [3]. A predictive model contains the best combination of covariates needed to achieve the best predictive accuracy [4]. Covariates can come from various sources, such as from an individual’s socioeconomic status, and psychological health [1, 4]. Validated predictive models of neck pain recovery have reported an Area Under the Curve (AUC) ranging from 0.65 to 0.91, and the most consistent covariates were age and initial neck-pain disability [4].
Biomechanical variables have not been used in the development of predictive models in neck pain, unlike in other areas of musculoskeletal research (e.g. anterior cruciate ligament injury [5], occupational back pain [6]). Yet, it is well established that individuals with neck pain have different movement patterns and muscle behaviours (collectively termed as biomechanical variables) than asymptomatic controls [7-9], which may not be restricted to the neck. For example, individuals with neck pain walked with reduced trunk axial rotation angle range compared to controls [7], which may be attributed to greater trunk muscular co-contraction, which could have negative consequences to overall spinal health [7].

Incorporating biomechanical measures into predictive models can be challenging which may deter its more widespread inclusion in prognostic research. Firstly, technological advancement means that researchers can collect huge amounts of biomechanical data [10]. For example, up to 126 biomechanical variables can be extracted from a single accelerometer [10]. Interpreting a predictive model with many covariates is clinically challenging. Second, biomechanical variables can be scalar (e.g. peak angle) and functional (e.g. angle waveform) in nature. Functional variables may provide a richer mechanistic insight into an individual’s health, compared to scalar variables. For example, a reduced cervical extension range of motion (scalar) cannot discriminate if movement is limited at the start and/or the end of motion. Even though functional variables may provide more information than scalar variables, the former demand for special care and adequate generalizations of common statistical methods (e.g. stepwise regression).

For biomechanical measures to be considered as potential covariates in prognostic research, statistical methods that can handle functional and scalar covariates, plus being able to generate clinically interpretable models must be used. Herein, we used a state-of-the-art machine learning technique “FDboost” [11], to develop a predictive model of neck pain status using scalar and functional biomechanical covariates. The primary aim of the present
study was to investigate the predictive value of biomechanical measures collected during walking in the classification of individuals with and without neck pain.

2. Methods

2.1. Design

Data for the present study represents the result of a sub-study from a larger project investigating the effects of neck pain on cervical motor control [8]. The study obtained ethical approval from the Ethics Committee of the University of Birmingham, UK (CM06/03/17-1). All participants provided written informed consent prior to participation.

2.2. Participants

Twenty-one healthy (controls) and 26 neck pain individuals completed a single-session experimental study. Individuals with neck pain were included if they had: 1) an average neck pain intensity in the previous month of ≥ three on a Numerical Rating Scale (NRS) (0 = “no pain”, 10 = “worst pain possible” [12]), and 2) a neck pain duration for ≥ three months. Individuals with neck pain due to whiplash were included if the grade of severity was < three on the Quebec Task Force Classification. Healthy participants were included as controls if they presented with no history of neck pain during the last two years. All participants were excluded if they had: chronic respiratory, rheumatologic, or neurologic conditions, spinal surgery, or pain induced by a spinal fracture.

2.3. Descriptive characteristics

The following characteristics were collected from individuals with neck pain: 1) average and maximum pain intensity over the last four weeks using the NRS [12], 2) perceived neck disability using the Neck Disability Index (NDI), 3) fear of movement using the Tampa Scale for Kinesiophobia (TSK).
2.4. Experimental conditions

All participants performed three trials for each experimental conditions 1) rectilinear, 2) curvilinear clockwise (CW), and 3) curvilinear counterclockwise (CCW) direction walking. Both rectilinear and curvilinear walking were investigated as walking in daily life involve changes in gait path direction [13]. Participants were instructed to walk at their natural speed, along a straight path for five meters (rectilinear); or following a floor marked circle, with 1 meter radius, in a CW or CCW direction for three consecutive trials (for CW and CCW, a trial was defined as a complete loop). A one-minute resting period was provided every 5 min to avoid fatigue. Familiarisation of each condition was allowed before data acquisition. All walking conditions were performed barefooted, in a randomized order.

2.5. Biomechanical modelling

Eight infrared-based camera were used for motion capture (250Hz) (BTS Bioengineering, Milan, Italy). Twenty-six 14 mm retroreflective markers were attached on the trunk, pelvic, thigh, shank, and foot segments following the Davis protocol [14]. Head motion tracking was executed via a light rigid helmet including four reflective markers (apex, front, right and left side of the helmet). Anthropometric measurements were recorded for all subjects according to Davis's guidelines [14]. Marker trajectories were low-pass filtered at 10Hz (zero lag, 4th order Butterworth), and gait events of initial contact and toe-off were determined using a previously defined kinematic method [15].

Six bipolar electromyography (EMG) probes (16-bit resolution, 1kHz) were placed on the bilateral Sternocleidomastoid, Splenius Cervicis, Erector Spinae muscles following Barbero et al [16]. Prior to placement of the EMG sensors, the skin was prepared in accordance with the SENIAM guidelines (http://www.seniam.org/). EMG signals were rectified around the mean and low pass filtered via a fourth order Butterworth filter (9 Hz) to
create a linear envelope [17]. All participants performed 5s of antigravity contractions of each muscle during lying, and, after being processed as above, a 3s average of muscular activity envelope was extracted for use as a normalizing factor for each muscle.

2.6. Data analysis

Fifty-eight scalar and functional biomechanical variables were extracted per walking condition for all participants (Supplementary Material [SM] for description). These variables broadly represented the spectrum of biomechanical variables collected during gait (e.g. segment angles [18]; spatio-temporal variables [19]; gait variability variables [20]), and in neck pain neuromuscular research [21]. For EMG variables, activities from bilateral muscles were extracted; and for lower limb kinematics and spatio-temporal variables, only values from the right limb were extracted. For all variables, values within a right stride cycle (initial contact to initial contact) were extracted for subsequent analysis. Each biomechanical variable per condition was treated as a single covariate, making a total of 174 covariates. Treating each biomechanical variable per condition as a single covariate, will enable a clinician to prospectively collect the most important variables under specific walking conditions to use within a predictive model.

One participant was excluded as missing biomechanical data were present in the rectilinear walking trial. Eighty-four out of 174 biomechanical covariates were excluded as they exhibited a high absolute correlation of > 0.7 with all other covariates [22]. Ninety biomechanical covariates together with four demographic covariates of age, height, weight, and sex, were used as inputs for a scalar-on-function (SoFR) regression model. All biomechanical covariates were demeaned as pre-processing, so that different covariates had equal potential to be included in the model.
A SoFR model is one where the response variable takes on scalar values, and the covariates take on functional (or scalar) values. Functional regression models are extensions of standard regression models such as generalized additive models. With 94 covariates for \( N = 46 \) observations, the model cannot be estimated with conventional fitting methods without additional penalisation as the corresponding algorithm for parameter estimation suffers from a singular matrix. Hence, we used component-wise gradient boosting to estimate the model [11] to fit a functional logistic regression model. The algorithm is an iterative procedure which successively adds one covariate to the model, like a forward stepwise regression, with the ability to handle functional covariates, perform variable selection, and allow for penalized estimation. In order to estimate the optimal number of iterations, the data was divided by splitting the participants into 4 folds, each with a roughly similar ratio of individuals with neck pain, on which cross-validation was performed. The area under the Receiver Operating Characteristic curve (AUC) was used to quantify the model’s ability to discriminate the two groups. All analyses were performed using R version 3.5.3, using the “FDboost” package [11].

3. Results

Descriptive characteristics of the participants can be found in Table 1. The group averaged values for all functional covariates can be found in the SM (Figures s1, s2, s3). Two functional covariates trunk lateral flexion angle during CCW gait, and trunk flexion angle during CW gait; and a single scalar covariate, hip jerk index during CCW gait were selected as the best covariates of neck pain status. The model achieved an estimated AUC of 80.8%.

The final model in the application is:

\[
P (\text{group}_i = \text{neck pain}) = \text{Logit}^{-1}(\beta_0 + \int x_{i1}(t)\beta_1(t)dt + \int x_{i2}(t)\beta_2(t)dt + x_{i3}\beta_3)
\]
for participants \( i = 1, \ldots, 46 \) where \( \beta_0 \) is the intercept of 0.087, \( \beta_1(t) \) and \( \beta_2(t) \) are the coefficients of the two functional covariates (Figure 1a, 2a), and \( \beta_3 \) is the coefficient of hip jerk index during CCW gait with a value of 0.37.

For the scalar covariate of hip jerk index, an increase in hip jerk index by one unit (all jerk values are dimensionless due to the formulation, see SM) increased the log odds of being in the neck pain group by 0.37. To simplify the interpretation of the \( \beta \) coefficients of the functional covariates, the predicted log odds was calculated for each participant when only an instantaneous unit change occurs in a gait cycle (Figure 1b, 2b), and the cumulative increase in class probabilities was calculated when a change occurs across all time points (0% to 100%) of gait (Figure 1c, 2c). As examples, a 1° increase in trunk lateral flexion in CCW walking alone or a 1° increase in trunk flexion angle in CW walking alone only altered the log odds of being in the neck pain group by <0.02 in magnitude (Figure 1b, 2b). At the cumulative level, a 1° increase in trunk lateral flexion angle throughout gait alone reduced the \( P(group_i = \text{neck pain}) \) from 0.5 at 0% gait to 0.15 at 100% gait (Figure 1c); and a 1° increase in trunk flexion angle throughout gait alone increased \( P(group_i = \text{neck pain}) \) from 0.5 at 0% gait to 0.9 at 100% gait (Figure 2c).

**Discussion**

Prognostic research is important to guide clinical management of a complex disorder such as neck pain disorders. Much research have shown differences in movement strategies between individuals with and without neck pain [7-9]. Yet, biomechanical variables have never been incorporated into predictive models within neck pain research. The two main findings of the present study were that 1) curvilinear walking (both CW and CCW) provided the most discriminatory set of biomechanical variables, and 2) global and not local (i.e. non-
cervical) biomechanical variables were most discriminatory between individuals with and without neck pain.

Individuals with neck pain have been reported to walk with a reduced trunk axial rotation range compared to controls [7]. The differences between the present study and that of Falla et al. [7] could be attributed to at least two reasons. First, the present study investigated gait biomechanics during curvilinear walking, which requires greater trunk lateral flexion angles, than rectilinear walking which Falla et al. [7] adopted [23]. Second, the present study treated biomechanical variables as covariates in a prediction model, rather than as a response variable for hypothesis testing [7]. A biomechanical variable which is significantly different between two clinical groups may not in turn be the most discriminatory, when considered amongst a high-dimensional landscape of potential covariates.

Trunk kinematics typically work synergistically with cervical kinematics to produce head movements [24]. For example, 67% of head flexion angle is contributed by the cervical spine, with the remaining coming from the trunk [24]. The synergistic role between trunk and cervical joints suggests that a greater trunk flexion angle in individuals with neck pain compared to controls could be a compensatory strategy for a reduced cervical flexion angle in the former compared to the latter. However, this was not presently observed in that individuals with neck pain positioned their head, trunk, and even pelvic segments in a greater flexed posture, than controls. This suggests that altered trunk flexion angles during walking between individuals with and without neck may be a global “stiffening” strategy [7].

A global “stiffening” strategy is observed similarly in the frontal plane, across the head, trunk and pelvic segments. When walking in a CCW direction, the trunk normally flexes laterally towards the right to change the direction of progression of the centre of mass (COM) towards the centre of the circle in a leftward direction (termed “hip strategy” in [25]).
This implies that individuals with neck pain use less hip strategy than controls. Another strategy to alter the direction of progress is by altering foot placement during swing [25]. A smaller use of the hip strategy in individuals with neck pain was unlikely due to differences in foot placements between groups, given the similar stride width during CCW walking (both groups: mean of 0.36m). It is possible that the reduced trunk, head, and pelvic segments in individuals with neck pain could be a global response to reduce pain, and/or represent a fear avoidance behaviour [26].

The period within gait where each functional covariate had the biggest effect was between 20-25% cycle for trunk flexion angle, and between 45-50% cycle for trunk lateral flexion angle (Figures 1a, 2a). The period of 20-25% cycle represents a phase where the contralateral limb is approaching mid-swing, which requires trunk extension to raise the COM, reducing the amount of swing limb flexion needed to clear the ground. The period of 45-50% cycle represents a phase where step-to-step transition is happening, where the COM medial-lateral accelerations and postural stability demands are high [27]. Trunk kinematic differences between individuals with and without neck pain may only be partially explained by neuromuscular differences (see Figure s2 and s3 “erector spinae”). In CCW walking at 45-50% cycle, muscle activity of the left erector spinae was higher in individuals with neck pain compared to controls, which may result in a more vertically oriented trunk in the frontal plane. Inter-group differences between the erector spinae muscles during CW was similar, yet trunk flexion angle differences persist during 20-25% cycle. A more detailed neuromuscular and kinetic investigation of whole-body mechanics during curvilinear walking would be required in future research to explain the specific kinematic differences between individuals with and without neck pain.

That a stiffer walking strategy was adopted by individuals with neck pain compared to controls, was also supported by the discriminatory value of the hip jerk index. Research on
other disorders have shown that individuals with low back pain performed rectilinear walking with greater in-phase trunk-pelvis segment coordination, and running with greater leg joint stiffness [28, 29]. Greater stiffness reduces the shock attenuation capacity and increases load transmission to the proximal body segments, such as the back and neck regions [30].

Boosting as a technique is less commonly used in clinical biomechanics research, as compared to techniques such as Support Vector Machine (SVM) [31]. A disadvantage of techniques such as SVM is that the models can have a complex non-linear structure with a high number of covariates, which makes it less clinically interpretable. In contrast, the model produced in the present study allows clinicians to focus their data collection efforts to the measurement of just three movement variables in two gait tasks. Although association does not imply causation, knowing what movement variables and when movement best predicts neck pain status could help in clinical therapeutic management. For example, knowing that trunk lateral flexion angle in CCW gait is the best predictive covariate, may focus a hypothesis driven search of plausible neuromuscular impairments that could cause altered trunk kinematics for intervention; and even inspire the development of novel therapeutic strategies to correct aberrant gait kinematics [32].

A limitation of this study was the relatively small sample size compared to the number of covariates included in the model, which precluded splitting the data into a training and validation dataset. The number of participants in the present study was however, comparable to other similar research in clinical biomechanics (n = 41 in [31], n = 44 in [10]). In defence, the present study’s aim was to explore the development of predictive models using biomechanical variables, rather than aim to develop an externally validated predictive model. Another limitation was that variables included in the model were not specific to the individual’s side of pain, and specific to the direction of movement that aggravates the pain. For example, an individual may have right sided neck pain that is painful only during left
cervical rotation. Classification performance may be augmented by inclusion of subject-specific variables into FDboost, a highly relevant research area we leave for future investigations.

5. Conclusion

Three biomechanical variables (two functional and a scalar), trunk lateral flexion angle during CCW gait, and trunk flexion angle during CW gait; hip jerk index during CCW gait were selected as the best covariates of neck pain status. “FDboost” can be used in future prognostic modelling studies in neck pain, and other clinical areas, where biomechanical data are collected as part of a holistic health assessment. The clinical attractiveness of “FDboost” is that it can produce clinically interpretable models even with complex high dimensional datasets.

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

Acknowledgments: The authors of this study would like to thank Feras Alsultan for assisting in data collection. No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.
References


Figure captions

**Figure 1. CCW Trunk lateral flexion**

(a) Beta coefficient of trunk lateral flexion angle (°) during counterclockwise (CCW) walking; (b) predicted log odds of being in the neck pain group for each participant for a 1° increase in trunk lateral flexion angle per gait instance; (c) cumulative predicted probability of being in the neck pain group for each participant for a 1° increase in trunk lateral flexion angle across all gait instances; (d) visualisation of trunk lateral flexion differences between groups (not drawn to scale).
**Figure 2. CW Trunk flexion**

(a) Beta coefficient of trunk flexion angle (°) during clockwise (CW) walking; (b) predicted log odds of being in the neck pain group for each participant for a 1° increase in trunk flexion angle per gait instance; (c) cumulative predicted probability of being in the neck pain group for each participant for a 1° increase in trunk flexion angle across all gait instances; (d) visualisation of trunk flexion differences between groups (not drawn to scale).

**Figure 2.** (a) Beta coefficient of trunk flexion angle (°) during clockwise (CW) walking; (b) predicted log odds of being in the neck pain group for each participant for a 1° increase in trunk flexion angle per gait instance; (c) cumulative predicted probability of being in the neck pain group for each participant for a 1° increase in trunk flexion angle across all gait instances; (d) visualisation of trunk flexion differences between groups (not drawn to scale).
Table 1: Participants’ characteristics and results of self-report questionnaires (mean ± standard deviation, SD)

<table>
<thead>
<tr>
<th></th>
<th>Neck pain (n = 26)</th>
<th>Control (n = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>15 F, 11 M</td>
<td>10 F, 11 M</td>
</tr>
<tr>
<td>Age (years)</td>
<td>32.3 (12.6)</td>
<td>28.8 (10.8)</td>
</tr>
<tr>
<td>Maximum pain intensity (NRS)</td>
<td>6.2 (2.2)</td>
<td>-</td>
</tr>
<tr>
<td>Average pain intensity (NRS)</td>
<td>4.1 (1.7)</td>
<td>-</td>
</tr>
<tr>
<td>NDI</td>
<td>11.5 (6.7)</td>
<td>-</td>
</tr>
<tr>
<td>TSK</td>
<td>35.4 (8.3)</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: M- male; F-female; NRS-numerical rating scale (0-10); NDI – neck disability index; TSK- tampa scale of kinesiophobia