Regression models for the quantification of Parkinsonian bradykinesia

Ji-Won Kim, Yuri Kwon, Ju-Seok Yun, Jae-Hoon Heo, Gwang-Moon Eom, Gye-Rae Tack, Tae-Hong Lim and Seong-Beom Koh

School of Biomedical Engineering, Choonju, 380-701, Korea
Research Institute of Biomedical Engineering, Konkuk University, Choonju, 380-701, Korea
Department of Biomedical Engineering, University of Iowa, USA
Department of Neurology, College of Medicine, Korea University, Seoul, Korea

Abstract. The aim of this study was to develop regression models for the quantification of parkinsonian bradykinesia. Forty patients with Parkinson’s disease participated in this study. Angular velocity was measured using gyro sensor during finger tapping, forearm-rotation, and toe tapping tasks and the severity of bradykinesia was rated by two independent neurologists. Various characteristic variables were derived from the sensor signal. Stepwise multiple linear regression analysis was performed to develop models predicting the bradykinesia score with the characteristic variables as input. To evaluate the ability of the regression models to discriminate different bradykinesia scores, ANOVA and post hoc test were performed. Major determinants of the bradykinesia score differed among clinical tasks and between raters. The regression models were better than any single characteristic variable in terms of the ability to differentiate bradykinesia scores. Specifically, the regression models could differentiate all pairs of the bradykinesia scores (p<0.05) except for one pair in the finger tapping task and one pair in the toe tapping task. In contrast, any single characteristic variable was found not sensitive enough to discriminate many of the pairs, especially in case of the toe tapping task. The results suggest that the multiple regression models reflecting these differences would be beneficial for the quantification of bradykinesia because the cardinal features included in the determination of bradykinesia score differ among tasks as well as among the raters.

Keywords: Bradykinesia, quantification, stepwise multiple regression, Parkinson’s disease, clinical rating

1. Introduction

Bradykinesia is one of the cardinal clinical features of Parkinson’s disease (PD) and is characterized by the impairment in speed and amplitude of body movement [1]. Bradykinesia is an important diagnostic and therapeutic target [1] and one of the strongest predictors of the quality of life [2, 3] in patients with PD.

To assess bradykinesia, the motor part (part III) of unified Parkinson’s disease rating scale (UPDRS-III) is commonly used. However, the rating scale is determined subjectively based on the raters’ clinical experience [4]. Furthermore, it is discontinuous and coarse-grained, limiting the detection of small changes especially in very early stage PD [5]. Accurate assessment of motor
impairment is important in monitoring the progression of disease and in evaluating the response to therapeutic interventions. Therefore, more objective and more quantitative assessment methods are desired.

Recently, there have been studies quantifying the bradykinesia using a musical instrument digital interface keyboard [6], joystick [7], accelerometer [8], magnetic sensor [9, 10] and gyro sensor [11]. Various variables to represent the characteristics of bradykinesia were derived from the measured signals. The consistency of the relationship between each characteristic variable and the clinical score (i.e. UPDRS bradykinesia score) [6, 7, 11] and the ability of each variable to discriminate different clinical scores [10, 11] were evaluated in order to assess how well each variable denote the bradykinesia-specific features. A single characteristic variable may not fully include multiple aspects of bradykinesia such as speed, amplitude, hesitations, fatiguing, and arrests, while clinicians rate a single severity score from the observation and integration of multiple aspects of bradykinesia.

A multiple regression model may be useful to describe multiple aspects of bradykinesia in only one quantitative value. So far, there has been no study where the regression model was used to quantify bradykinesia. In the previous studies, many characteristic variables were shown to be significantly correlated with the clinical score [6, 7, 11]. When multiple variables are correlated with the clinical score, we need to know the relative importance of each variable on clinical rating of the bradykinesia. In this aspect, stepwise multiple regression analysis would be desirable for the determination of the unique and relative contribution of each variable.

There are several tasks for the assessment of the upper and lower limb bradykinesia (e.g., finger tapping, hand movements and pronation-supination movement of hands and toe tapping) in UPDRS. However, previous studies have quantified bradykinesia focusing on only one clinical task such as finger tapping [8-11] or forearm rotation [12]. Important features affecting clinical rating might be different among the tasks. Furthermore, relative importance of each feature of bradykinesia might be different among the raters, indicating that the major determinants of the clinical score might be different among the clinical tasks and raters. In this study, multiple regression analyses were performed to test if a bradykinesia score can be more objectively determined by combined use of various characteristics.

2. Methods

2.1. Subjects

Forty patients with idiopathic PD (20 men and 20 women, 65.7 ± 11.1 years) participated in this study. Hoehn and Yahr stage was 2.3 ± 0.5 and mean disease duration was 4.2 ± 2.9 yrs. The levodopa equivalent dose was 512 ± 298 mg/day, and no patient showed wearing-off for the medication. The patients were recruited according to Clinical Diagnostic Criteria described by the United Kingdom Parkinson’s Disease Society Brain Bank. This study was approved by the research ethics committee of Korea University Guro Hospital (Seoul, Korea) and all patients gave written informed consent.

2.2. Experiments and outcome measures

In this study, an angular rate measurement system integrating a piezoelectric ceramic gyro sensor (CG-L53, NEC/Toxin, Japan) with a miniature size (6x10x25 mm) and light-weight (0.26 g) was used
for measuring the angular velocity of the joint. The joint movements such as finger, forearm and ankle are rotational in nature and the gyro sensor signal is free from gravitational artifact which is common cause of measurement error in accelerometers. Details of the system have been described previously [11]. The angular displacement signal was derived from the numerical integration of the angular velocity.

UPDRS-III includes many categories as motor examination. The severity of bradykinesia is commonly evaluated by summing bradykinesia scores of finger tapping, forearm rotation, hand movements and toe tapping. In this study, finger tapping, forearm rotation and toe tapping tasks were selected as the tasks of investigation from the UPDRS-III because all of them are uniaxial movements and conveniently measured by one gyro sensor. On the other hand, hand movement was excluded from this study because multiple-sensors were needed to measure the multi joint movements of hand fisting and relaxing, which was inconvenient and burdensome for patients and might alter the patients’ movement pattern. Subjects performed each task for 10 s as rapidly as possible and with as large amplitude as possible. The gyro sensor signals were recorded in a computer received through serial port with 250 Hz of sampling frequency and the limb motion was also recorded by a digital video-recorder. Both left and right limbs were tested in all tasks. After the experiment, two clinicians rated bradykinesia score from the careful observation of video files.

Some patients showed tremor with frequency range of 5–8 Hz during the task. Therefore, in case of tremor-contaminated data, we discriminated tremor from voluntary movement using low pass filter (fc=5 Hz) [11].

Fourteen characteristic variables (8 in time domain and 6 in frequency domain) were derived from the sensor signal to represent various features of bradykinesia. Three time domain variables, root mean squared (RMS) mean value, maximum value, and coefficient of variation (CV), were calculated from the angular velocity and angular displacement (six variables in total). RMS value was calculated for the entire test duration (10 s) and RMS velocity and RMS displacement were regarded to represent overall movement speed and amplitude, respectively. Maximum velocity and displacement were also extracted from the entire data and regarded to represent the best performance during the task. In this study, CV was used to represent the irregularity of the movement pattern due to the hesitations and interruptions in ongoing movement. To acquire the trend of movement from the oscillating velocity and displacement, first, RMS values were repeatedly calculated for 1-second window sliding throughout the entire data (10 s). Then, the mean and standard deviation of the RMS values were calculated and the CV was defined as the ratio of the standard deviation and the mean [13].

Two more time domain variables were added to the above, i.e., RMS acceleration and RMS jerk derived from the first order and second order numerical differentiation of angular velocity, respectively. They were expected to represent the average amount of force exerted (the effort for the task) and the change in force (the roughness of motion), respectively. Accordingly, time-domain variables were eight in total.

Three frequency domain variables, i.e., the peak power, total power and peak frequency in the power spectrum of angular velocity and displacement were calculated (six variables in total). Each of them was expected to represent the intensity of the main movement component, the total intensity of movement including all frequency components, and the main frequency of movement, respectively

2.3. Statistical analysis

Regression models were determined by stepwise multiple linear regression analysis for each clinical task and for each rater. Some characteristic variables showed exponential decay pattern with
the increase in the clinical bradykinesia score (Figure 1). When the $R^2$ of the exponential regression model for a characteristic variable was greater than that of linear regression model as much as 0.15, the variable was transformed to logarithmic index before performing stepwise multiple linear regression analysis.

An independent characteristic variable was included in the regression model in the stepwise regression process, if its F value was greater than 2.0. Mathematical explanatory weighting ($\beta$) of an independent variable in the regression equation and the significance of the weighting was examined to see the relative importance of the variable in comparison to other variables. An independent variable was excluded from the model if the adjusted $R^2$ for each step of regression was insignificant (p>0.05).

The ability of a regression model to discriminate different clinical scores was evaluated by ANOVA and post hoc test. SPSS ver. 20 for Windows (SPSS Inc., Chicago, IL) was used for all statistical analyses.

Table 1

<table>
<thead>
<tr>
<th>Limb</th>
<th>Task</th>
<th>Rater</th>
<th>Step</th>
<th>Model</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>p-Level of $R^2$</th>
<th>$\beta$</th>
<th>p-Level of $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>Finger-tapping</td>
<td>Rater 1</td>
<td>1st</td>
<td>RMS velocity</td>
<td>0.48</td>
<td>0.47</td>
<td>***</td>
<td>-0.69</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Final</td>
<td>RMS velocity</td>
<td>0.49</td>
<td>0.48</td>
<td>***</td>
<td>-0.53</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Peak power of velocity</td>
<td></td>
<td></td>
<td></td>
<td>0.20</td>
<td>p=0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rater 2</td>
<td>1st</td>
<td>RMS velocity</td>
<td>0.53</td>
<td>0.52</td>
<td>***</td>
<td>-0.73</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Final</td>
<td>RMS velocity</td>
<td>0.56</td>
<td>0.53</td>
<td>***</td>
<td>-0.47</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Peak power of velocity</td>
<td></td>
<td></td>
<td></td>
<td>0.32</td>
<td>*</td>
</tr>
<tr>
<td>Forearm-rotation</td>
<td>Rater 1</td>
<td>1st</td>
<td>RMS velocity</td>
<td>0.62</td>
<td>0.61</td>
<td>***</td>
<td>-0.79</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Final</td>
<td>RMS velocity</td>
<td>0.64</td>
<td>0.63</td>
<td>***</td>
<td>-0.52</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean acceleration</td>
<td></td>
<td></td>
<td></td>
<td>p=0.06</td>
<td></td>
</tr>
</tbody>
</table>
3. Results

The resultant models of stepwise multiple regressions for each rater’s rating score in each clinical task are shown in Table 1. Major determinants of the bradykinesia differed between raters as well as among different clinical tasks.

In case of finger tapping task, major determinants of both raters were RMS velocity and peak power of velocity. In contrast, major determinants of bradykinesia in other tasks differed between raters. In case of forearm rotation, major determinants of rater 1 were RMS velocity and mean acceleration whereas major determinants of rater 2 included many variables such as the RMS velocity, peak frequency of angle and coefficient of variation of velocity. In case of toe tapping task, major determinant of rater 1 was only logarithmic peak power of velocity. Major determinants of rater 2 included the coefficient of variation of angle in addition to the logarithmic peak power of velocity.

Figure 2 presents the relationship between regression model output and the clinically determined bradykinesia score in each clinical task. In all clinical tasks, the outputs of the regression models increased with bradykinesia score and the increasing curve pattern showed different tendency between raters.

The discrimination capability of each characteristic variable against different bradykinesia scores are shown representatively for the toe tapping task in Table 2. ANOVA showed significant difference in all single variables with regard to different bradykinesia scores (p<0.05). However, in the result of post hoc comparisons, single variables were found not be able to discriminate many non-neighboring pairs and most neighboring pairs. Also, the specific pairs that could be discriminated differed among variables. For example, coefficient of variation of angle could discriminate 1 vs. 2 and 3 vs. 4 while the peak power of velocity could not discriminate them but it could discriminate 0 vs. 1. These results suggest combination of multiple variables would be advantageous for the discrimination of overall pairs. The results for the finger tapping and forearm rotation tasks were similar to those in toe tapping task with regard to single characteristic variables.
Fig. 2. Relationship between bradykinesia score and output of regression model in each clinical task.

### Table 2

Results of ANOVA and post-hoc test for the toe-tapping task

<table>
<thead>
<tr>
<th>Variables</th>
<th>Rater</th>
<th>ANOVA</th>
<th>Post-hoc (Tukey) test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>neighboring pair</td>
<td>non-neighboring pair</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 vs. 1 1 vs. 2 2 vs. 3 3 vs. 4</td>
<td>0 vs. 2 0 vs. 3 0 vs. 4 1 vs. 3 1 vs. 4 2 vs. 4</td>
</tr>
<tr>
<td>Angular Velocity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS</td>
<td>1</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Max</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>1</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Peak power</td>
<td>1</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td>Total power</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
<tr>
<td>Peak frequency</td>
<td>1</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Angle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMS</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
<tr>
<td>Max</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
<tr>
<td>Peak power</td>
<td>1</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td>Total power</td>
<td>1</td>
<td>***</td>
<td>*</td>
</tr>
</tbody>
</table>
Table 3 shows the discrimination capability of the regression models for all tasks. ANOVA showed significant difference in the output of regression models (p<0.05). Moreover, in the post hoc comparisons, the regression models could discriminate all non-neighboring pairs and most neighboring pairs (p<0.05) except for 2 cases i.e., score 1 vs. 2 of rater 1 in the finger tapping task (p=0.15) and score 0 vs. 1 of rater 2 in the toe tapping task (p=0.36). Especially, forearm rotation task showed significant difference among all neighboring pairs as well as all non-neighboring pairs in both raters.

4. Discussion

The main findings of this study were as follows. First, task-specific and rater-specific differences existed in the regression models and accordingly, the major determinants of the bradykinesia score differed between raters as well as among clinical tasks. Second, the discrimination capability of the regression models against different bradykinesia scores was superior to those of single variables.

In the following, the regression results and validation results will be discussed for each task.

4.1. Finger tapping task

In the finger tapping task, the final model included RMS velocity and peak power of velocity in both raters. This indicates that average movement velocity and intensity of main velocity component during finger tapping (repetitive flexion and extension) are associated with clinical score ratings and that the evaluation criteria for rating were similar in both raters.

Regarding the contribution to the bradykinesia score, β of RMS velocity (rater 1: -0.53; rater 2: -0.47) was greater than that of peak power of velocity (rater 1: -0.20; rater 2: -0.32). This result
suggests that average movement velocity would be much more important component for rating of finger tapping among the major determinants.

4.2. Forearm rotation task

In contrast to finger tapping task, rater-specific dependence was shown in the regression models for the forearm rotation task. In case of rater 1, RMS velocity and mean acceleration were selected as major determinant. This indicates that rater 1 mainly considered average movement velocity ($\beta=-0.52$) and movement acceleration ($\beta=-0.30$) as important features of bradykinesia.

On the other hands, rater 2 focused on more variables such as RMS velocity, peak frequency of angle and CV of angular velocity. That is, average movement velocity ($\beta=-0.59$), speed of main angular displacement component ($\beta=-0.20$), and irregularity and fluctuation by hesitation and arrests ($\beta=0.17$) contributed to the rating of forearm rotation in rater 2.

Contribution ($\beta$) of average movement velocity for rating was greater than that of other features as was the case in the finger tapping task. This suggests that clinical rating of upper limbs including finger and forearm is mainly associated with average movement velocity among many features of bradykinesia.

4.3. Toe tapping task

As in the case of the forearm rotation task, final regression model of the toe tapping task differed between raters. Rater 2 considered not only logarithmic intensity of velocity component ($\beta=-0.59$) but also irregularity and fluctuation of movement ($\beta=-0.82$) while rater 1 considered only logarithmic intensity of main velocity component. Unlike upper limb bradykinesia, dominant contributor to clinical rating was logarithmic intensity of main velocity component. This result indicates that the raters considered main movement component more importantly than the average movement.

4.4. Validation of the regression models

It is important to validate regression models if they reflect bradykinesia scores well. In this study, ANOVA and post hoc comparisons were performed to investigate whether the regression models could discriminate different bradykinesia scores (Table 3). As expected, the discrimination capability of the regression models against different bradykinesia scores (Table 3) was superior to that of single variables (Table 2) in all clinical tasks. That is, all non-neighboring pairs could be discriminated well in all tasks ($p<0.05$). Especially, regression model of forearm rotation task could discriminate all neighboring pairs as well as all non-neighboring pairs ($p<0.05$). Even in other tasks, most neighboring pairs could be discriminated except for 1 vs. 2 of rater 1 in finger tapping task ($p=0.15$) and 0 vs.1 of rater 2 in toe tapping task ($p=0.36$). These results suggest that the regression models of this study reflect bradykinesia score well and that they are better than any single variables. This is in agreement with our hypothesis that considering multiple aspects of bradykinesia would be more beneficial for quantification in comparison to single variables.

Regarding the discrimination capability, a feasible reason why the regression models were better than single variable is that the regression model consists of complementary multiple variables for discrimination of different scores. For example, in toe tapping task, peak power of velocity could not discriminate 1 vs. 2 and 3 vs. 4 and coefficient of variation of angle could discriminate them (Table 3). That is, failure of discrimination by peak power was compensated by coefficient of variation so that
regression model consisting of the two variables had much better discrimination capability than each of single variables. Similarly, in case of forearm rotation task, failure in the discrimination of 3 vs. 4 by RMS velocity was compensated by coefficient of variation of velocity so that regression model could discriminate all pairs. Also the results of finger tapping task were similar to those of toe tapping and forearm rotation tasks.

In summary, the results of this study demonstrated the high variability in clinical rating of bradykinesia in terms of rater-specific and task-(target joint) specific selection of rating criteria. Although previous studies have tried to quantify the bradykinesia, they have neglected to consider multiple aspects of the bradykinesia [6-11]. Furthermore, they have investigated only single joint for quantification. This study has significance in that the regression models included multiple characteristic variables and the models were validated for various clinical tasks for upper and lower limb joints. According to the results of this study, clinical ratings regarded different features of bradykinesia as cardinal which are rater-specific and task-specific. We suggest that the multiple regression models reflecting these dependences would be beneficial for the quantification of bradykinesia.

The suggestion of this study needs to be additionally verified on a larger population of patients as well as many clinicians. Development of standard regression models of rating for each joint movement task would also be desirable.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2014R1A1A2057508).

Reference


