Distinguishing Patients With a Coordination Disorder From Healthy Controls Using Local Features of Movement Trajectories During the Finger-to-Nose Test

Soancatl Aguilar, Venustiano; Martinez Manzanera, Octavio; Sival, Deborah A.; Maurits, Natasha M.; Roerdink, Jos B. T. M.

Published in:
IEEE Trans. Biomedical Engineering

DOI:
10.1109/TBME.2018.2878626

Publication date:
2019

Distinguishing Patients with a Coordination Disorder from Healthy Controls using Local Features of Movement Trajectories during the Finger-to-nose Test

Venustiano Soancatl Aguilar, Octavio Martinez Manzanera, Deborah A. Sival, Natasha M. Maurits, *Senior Member, IEEE* and Jos B.T.M. Roerdink, *Senior Member, IEEE*

***Abstract*—Assessment of coordination disorders is valuable for monitoring progression of patients, distinguishing healthy and pathological conditions, and ultimately aiding in clinical decision making, thereby offering the possibility to improve medical care or rehabilitation. A common method to assess movement disorders is by using clinical rating scales. However, rating scales depend on the evaluation and interpretation of an observer, implying that subjective phenotypic assignment precedes the application of the scales. Objective and more accurate methods are under continuous development but gold standards are still scarce. Here, we show how a method we previously developed, originally aimed at assessing dynamic balance by a probabilistic generalized linear model, can be used to assess a broader range of functional movements. In this paper the method is applied to distinguish patients with coordination disorders from healthy controls. We focused on movements recorded during the finger-to-nose task (FNT), which is commonly used to assess coordination disorders. We also compared clinical FNT scores and model scores. Our method achieved 84% classification accuracy in distinguishing patients and healthy participants, using only two features. Future work could entail testing the reliability of the method by using additional features and other clinical tests such as finger chasing, quiet standing and/or usage of tracking devices such as depth cameras or force plates.**

***Index Terms*—Coordination disorders, classification, finger-to-nose test, generalized linear models, instantaneous speed and local curvature.**

I. Introduction

Human movement analysis can be valuable for diagnosis and monitoring of motor disorders; it can aid in distinguishing healthy and pathological conditions and in following the progression of patients over time and underpinning the efficacy of interventions [1]. A common method to assess human movement in a clinical setting is provided by standardized clinical rating scales, which are validated and easy to administer [2]. However, one of the main drawbacks of rating scales is that they depend on the evaluation and interpretation of an observer and thus contain a subjective component. Moreover, clinical scales are not enough to assess different motor control strategies during the execution of movements [2]–[4]. Thus, techniques for reliable and objective movement assessment could contribute to clinical practice in neurology, rehabilitation and other fields of medicine. As added benefit, they could be used in combination with clinical rating scales [5].

One of the challenges when developing quantitative and objective methods to assess human movements is the lack of methods to establish the validity of the measurements. In a previous study (see [6], [7]) we used the movement performance of younger and older participants as a proxy for better or worse movement, knowing that movement in older participants is generally worse than in younger participants. We then successfully used generalized linear models (GLMs) [8] to predict movement category (young or old) based on features derived from the movement trajectory. One characteristic that makes GLMs appropriate to classify human movement as better or worse is that their outcomes can be probability values that reflect movement performance. In case of diagnostic applications, if we assume that probability 0 represents "healthy" performance and probability 1 represents "pathological" performance, we propose that a probabilistic GLM could be used in a similar way. Intermediate probability values as estimated by the GLM would then indicate how similar the movements are to those movements that reflect pathological performance.

Here, we evaluate how the method used in our previous study [6] performs when applied to the problem of distinguishing patients, with a coordination disorder, from healthy controls. We focused on movements recorded from the finger-to-nose task (FNT) as recorded using inertial measurement units (IMUs). The FNT is a kinetic subscale measuring dysmetria and intention tremor during coordinated upper-limb movements between the tip of the nose of the participant and the tip of an examiners index finger [9]. This subscale score is included as part of the summed scores of the International Cooperative Ataxia Rating Scale (ICARS) [10] and the Scale for the Assessment and Rating of Ataxia (SARA) [11]. This test can thus be used to quantify coordination impairment in patients with Early-Onset-Ataxia (EOA) or Developmental Coordination Disorder (DCD) [12]. In the present study, we aimed to compare quantitative FNT-IMU data of pediatric and adult patients with ataxia or DCD to those of healthy

Manuscript received February 14, 2018; revised April 27, 2018 and October 25, 2018.

Venustiano Soancatl Aguilar was supported by the Mexican National Council of Science and Technology (CONACYT) under scholarship number 313791 (correspondence e-mail: v.soancatl.aguilar@rug.nl)

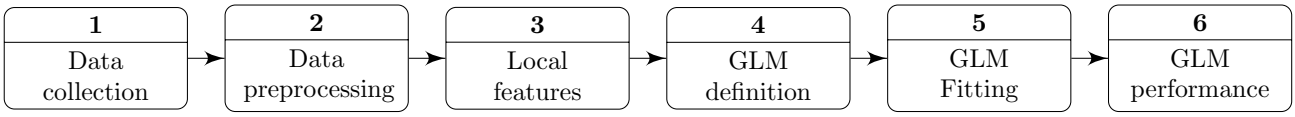

Fig. 1. Steps of the method to assess movement trajectories.

controls [12]. If both groups could be reliably discerned by FNT-IMU data, we would conclude that the test could provide an objective instrument to support the distinction of patients with coordination problems from controls.

Figure 1 illustrates the general steps of FNT-IMU measurements in a heterogeneous group of pediatric and adult patients with coordination impairment (ataxia or DCD) and healthy controls. In step 1, the three-dimensional (3D) trajectories are collected (during a functional task) using tracking technology such as inertial measurement units or depth cameras. Step 2 involves smoothing the signals, segmenting the movement trajectories (if necessary), or any other preprocessing of the data. Step 3 involves the extraction of local features from the 3D trajectories. This is one of the most important steps in the methodology, as the local features must characterize the movements under study. Three important local features are curvature, torsion, and velocity, as these features fully characterize a curve in 3D space [13]. In addition, depending on the task other features can also be included for the assessment of the movements such as the number of velocity peaks, the target error, and spatial overshoot [2]. In step 4, one or more (probabilistic) GLMs are defined as a function of the local features estimated in step 3. The mathematical definition of GLMs can be given following the steps in [8]. In step 5, the GLMs are fitted using R, Matlab or any other specialized statistical software. In step 6, the performance of the models on new data is assessed using the Akaike or Wanatabe-Akaike information criteria [14], [15]. Cross-validation [16] can also be used for model comparison, although this is computationally more expensive.

In Section II we detail the steps in Figure 1 to investigate the predictive accuracy of a GLM to distinguish between healthy participants and patients with coordination disorders, regardless of age and class of coordination disorder. In general we expect that FNT movement trajectories are smoother and faster in healthy participants than in patients. Results are presented in Section III. Finally, in Section IV a general discussion and ideas for future work are presented.

II. Methods

This study was performed using data acquired in the context of the project *Quantification of symptoms of movement disorders employing motion sensors* [17]. Part of the data in [17] were previously used to investigate whether a random forest classifier employing 14 features derived from 3D movement trajectories during the FNT could classify children with coordination problems and age-matched healthy control children. In the present study we only use two features, which are different from those used in [17] and include additional participants.

A. Participants

In the present study we included two groups of participants. The first group concerns the data used in the study mentioned above, consisting of 34 children: 16 patients with a coordination disorder, of which nine with EOA (mean age 13.3 years, SD 4.0 years) and seven with DCD (mean age 9.4 years, SD 2.2 years), and 18 healthy age-matched controls (mean age 11.8 years, SD 3.4 years). The second set involved 36 participants: 34 patients with a coordination disorder, of which 12 children with EOA (mean age 13.5 years, SD 2.8 years) and 22 adults with Adult Onset Ataxia (AOA; mean age 54.9 years, SD 14.7 years), and two healthy participants 20 and 21 years old. DCD participants were included as patients with mild coordination disorders increasing the difficulty to distinguish between patients and healthy participants.

By including patients and controls over a wider age range, we increased the complexity of the data compared to the previous study (see [17]). All parents of the children and all adult participants provided written informed consent. Children who were 12 years or older provided informed assent. Inclusion criteria for ataxia patients were a clinical diagnosis of pediatric ataxia or recognition of ataxia as a primary movement disorder as assessed by three experts in movement disorders.

Presence of ataxia was clinically established in accordance with the definition of ataxia (see [18]). As part of their diagnostic assessment, included EOA and AOA patients had received radiologic, metabolic and generic assessments at the department of (pediatric) neurology at University Medical Center Groningen (UMCG). According to the official DSM-V criteria, DCD is a developmental disorder, characterized by non-progressive motor incoordination, interfering with daily activities or academic achievement, not attributable to a neurological, intellectual or visual condition. This implicates that a neurologic investigation always precedes the referral to the pediatric rehabilitation doctors. All included DCD patients were thus assessed at the outpatient clinic of the department of pediatric neurology at UMCG, as well. When necessary to exclude other neurologic disorders, these assessments also included MRI, electromyography, muscle ultrasound and/or laboratory tests. Exclusion criteria for healthy participants were a neurological and/or orthopedic disorder and/or any medication with a negative effect on coordination. Furthermore, healthy children were declared to be healthy by their parents.

We included these three groups (EOA, AOA and DCD) as they are associated with a different factor that could affect coordination performances and/or SARA scores: ataxia, clumsiness in general and immaturity of the central nervous system. In adults, these three factors would be expected to play a smaller role, because the effect of DCD is clinically known to diminish in adulthood. Furthermore, effects of immaturity

on coordination performances generally disappear after puberty [19]. Finally, pediatric onset ataxia (EOA) concerns a highly heterogeneous group, which is more often associated with other (comorbid) movement disorder features (mixed movement disorder) than ataxia onset after the 25th year of life (AOA). In children, it is thus much more difficult to distinguish between the different factors that can affect coordination.

B. Data collection and preprocessing (steps 1 and 2)

Participants were asked to perform at least 10 FNT cycles with both hands if they were able to do so, on average the trials lasted 21.8 sec (SD 7.8 sec), left and right. The trials were video recorded. Three pediatric neurologists additionally assessed the FNT executed by EOA and DCD participants, according to the official SARA guidelines [11], for the first group of participants only. SARA assessment was not performed for the second group of patients because the goal of that study was not to compare results to the SARA score, but to distinguish between groups. During task execution, participants wore three inertial measurement units (IMUs Shimmer3, Shimmer, Dublin, Ireland-based Realtime Technologies) on the upper arm, fore arm, and index finger. The data collected by the IMUs at 51.2 Hz were used to estimate 3D trajectories of the participants' index finger using an upper limb model [17] implemented in Labview (Austin, Texas, USA). We subsequently applied a moving-average filter of the 3D trajectory data using a window of 15 samples to smooth the signals.

C. Estimating local features (step 3)

Local features are those that can be estimated for short segments taken from the 3D trajectories, such as curvature, torsion, instantaneous speed, and their time-derivatives [20], [21]. Compared to global features, local features have the added value that they offer the possibility to assess performance in "real-time" and provide immediate feedback. As local features we selected local curvature and instantaneous speed because they allowed high classification accuracy in our previous study involving movement of younger and older participants [6] and because they are expected to provide relevant information about the ability of the current participants to perform the FNTs. The curvature of a trajectory is defined as a measure of deviation of a curve from a straight line. In this sense, a straight trajectory has a constant zero curvature. One way to estimate the curvature of a trajectory at a specific point in time x_t is by fitting a circle to x_{t-1}, x_t and x_{t+1} and taking the inverse of its radius. Thus, large circles will reflect small curvature values and small circles will reflect high curvature values. Curvature values of non-intermittent (smooth) movement trajectories can be expected to be small, whereas for the intermittent (non-smooth) cases are expected to be high. In this way, curvature can indeed provide information related to movement smoothness. Thus, local curvature measures how smooth a 3D trajectory is for each three consecutive points, while instantaneous speed is determined between each two consecutive points. By visualizing curvature and speed signals and identifying the repetitive FNT movement, samples that preceded or followed FNT execution were excluded from further analysis. After sample exclusion, trials lasted 17.5 sec on average (SD 6.4 sec). Then, local curvature (κ) and instantaneous speed (s) were estimated from the FNT trajectories according to the method of Soancatl-Aguilar et al. [20] and subsequently log-transformed. Finally, mean speed ($\bar{s}$) and mean curvature ($\bar{\kappa}$) were estimated for each participant k and used as predictors in the GLM.

D. GLM definition and GLM fitting (steps 4 and 5)

Following the steps described in [8] we specified a GLM as follows. First, we defined an outcome variable (d) as binary (0 - healthy class, 1 - coordination disorder class) and assumed that it follows a Bernoulli distribution. Second, a linear model was specified as a function of mean speed $\bar{s}$ and mean curvature $\bar{\kappa}$. Third, the logit function [22] was used to transform the probability distribution constrained between 0 and 1 into a function that can take any real value. Mathematically:

$$\begin{aligned} & d_k \sim \mathit{Bernoulli}(\mathbb{P}_k), \qquad k = 1 \ldots n \\ & \mathrm{logit}(\mathbb{P}_k) = \alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k, \\ & \alpha \sim \mathcal{N}(0, 10), \quad \beta_1 \sim \mathcal{N}(0, 50), \quad \beta_2 \sim \mathcal{N}(0, 50) \end{aligned} \tag{1}$$

where n is the number of participants, α is the intercept, β_1 and β_2 are the slopes, and k is a participant index. The logit function is defined as the logarithm of the odds *(log-odds)* [22], where the odds of $\mathbb{P}_k$ is $\mathbb{P}_k/(1-\mathbb{P}_k)$. Thus,

$$\mathrm{logit}(\mathbb{P}_k) = \log\left(\frac{\mathbb{P}_k}{1-\mathbb{P}_k}\right) = \alpha + \beta_1 \cdot \bar{\kappa}_k + \beta_2 \cdot \bar{s}_k, \tag{2}$$

and solving for $\mathbb{P}_k$

$$\mathbb{P}_k = \frac{e^{\alpha+\beta_1\cdot\bar{\kappa}_k+\beta_2\cdot\bar{s}_k}}{1+e^{\alpha+\beta_1\cdot\bar{\kappa}_k+\beta_2\cdot\bar{s}_k}} \tag{3}$$

$\mathcal{N}$ represents a normal distribution with 0 mean and standard deviation 10 for the intercept (α) and standard deviation 50 for the slopes (β_1 and β_2). These features and values are just initial assumptions. For simplicity, we used the same probability distributions for the model as in our previous study [6] where we used high posterior density interval (HPDI) bands, which are conceptually similar to confidence interval bands. The GLM fitting method will iteratively estimate the final values of these parameters [23]. To fit the GLM (Eq. (1)) we built a model in *Stan*, which is a probabilistic programming language [24], using the *rethinking* R package [23].

E. GLM performance (step 6)

We performed leave-one-out cross validation (LOOCV) [25] to test the performance of the model on new data. Suppose that the set U contains the pairs $(\bar{\kappa}_k, \bar{s}_k)$ collected from the two groups of participants ($k = 1 \ldots 70$). Then, for each participant k in U we fitted a model on the set $\{U - (\bar{\kappa}_k, \bar{s}_k)\}$ and used the fitted model to predict the probability that participant k belongs to the coordination disorder class. The predicted probabilities were used to estimate an optimal *threshold* to classify FNT trials as belonging to a healthy or coordination disorder participant. This threshold was estimated as the point with the best sum of *sensitivity* and *specificity* known as the

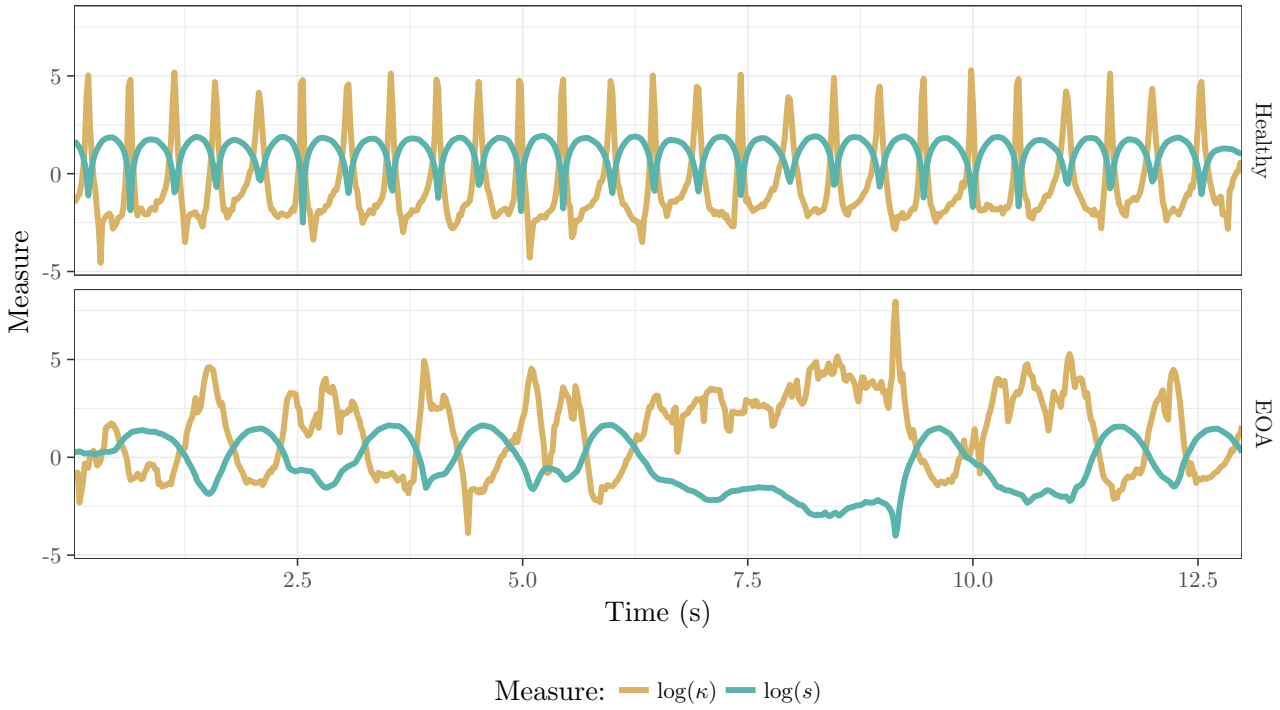

Fig. 2. Local curvature (κ) and instantaneous speed (s) during FNT execution, using the right hand, for a healthy (top) and an EOA (bottom) participant.

Youden index [26], closest to the point (0,1) of the receiver operating characteristics (ROC) curve [27]. Sensitivity is the proportion of correctly classified patients. Specificity is the proportion of correctly classified healthy participants. The threshold was estimated using the pROC R-package [28].

1) SARA FNT scores compared to GLM scores: To gain further understanding of any misclassifications, SARA FNT scores and model scores were compared for the first group of patients only. The mean SARA FNT score across observers for each patient in the first data set was determined. Then, to investigate to what extent SARA FNT scores coincide with model scores a scatter plot was used. For specific cases, we visualized 3D trajectories and the distribution of curvature and speed values as violin plots to gain further understanding.

III. Results

Figure 2 provides an example of local curvature and instantaneous speed (in log-scale) as a function of time for a healthy participant and a participant with EOA. Both measures, speed and curvature, are clearly regular and repetitive for the healthy participant. For the patient, however, both measures behave more irregularly. Taking into account the range of the measures, the healthy participant displayed faster and smoother movements than the EOA participant, as indicated by higher speed values and lower curvature values. It can also be observed that high speed values coincide with low curvature values and, vice versa, low speed values coincide with high curvature values. This is known as the power law relation between curvature and speed in log-scale [29].

A. GLM classification

After performing LOOCV and using the ROC curve, the probability threshold that best separates healthy participants from patients was found to be 0.587 (Figure 3). Using this threshold, 84% of the healthy participants were correctly classified and 84% of the patients were correctly classified. Figure 4 shows the LOOCV predictions of model (1). Most of the healthy participants are grouped in the top left corner of the graph; this group represents participants who scored probability values lower than the threshold, and were classified as healthy participants. Most of the patients are in the group dispersed between the center and the bottom right corner of the graph; this group represents participants who scored probability values higher than the threshold and were classified as patients. These findings again illustrate that in general healthy participants displayed faster and smoother FNT movements than patients. Some overlap between the two groups of participants, however, prevents a better separation. For example, some healthy participants (5, 23, and 25) score similar probabilities as patients, while one DCD patient and one EOA patient (32 and 61, respectively) score similar probabilities as healthy participants. Thus, according to model (1) participants 32 and 61 behave very much as healthy participants.

B. SARA FNT scores compared to GLM predictions

In Figure 5 GLM scores are plotted against SARA FNT scores to gain further understanding of misclassified patients from group 1. From this figure we can observe that most of the misclassified patients had relatively low SARA FNT scores, meaning that the observers noticed only small or no tremor at

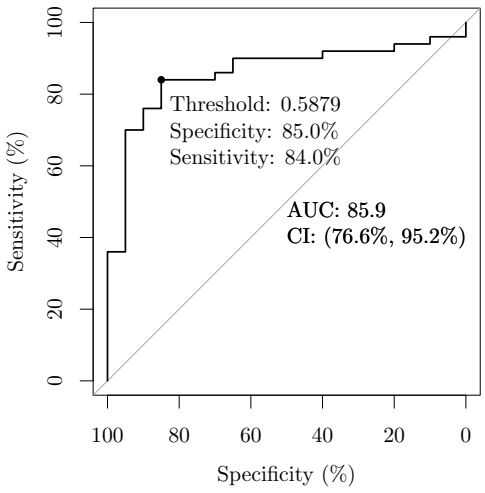

Fig. 3. ROC curve (non-smoothed) of the LOOCV showing the threshold that best separates patients from healthy participants. The plot was created using the pROC-Rpackage [28] and the tikzDevice package [30]. CI: confidence interval.

all (smooth FNT trajectories). However, some patients received a low SARA score (suggesting that the FNT trajectories looked smooth to the observers) whereas the model score was high (participants 4, 24 and 30, in the top left corner of Figure 4). From a classification point of view, this suggests that the model classifies patients not only on the presence of irregularities (tremor) which would result in high curvature.

To understand why the model classifies some patients with no visible or minor tremor as healthy and others (correctly) as patients, we present violin plots of the curvature and speed distributions in Figure 6 and 3D trajectories in Figure 7, for participants 4, 8, 14, 24, 30, and 31 (both hands). We included participant 8 as an example where no visible tremor was observed and the model classified the patient as healthy, and participants 4 and 24 as examples where no visible tremor was observed but the model correctly classified them as patients. For comparison, we also included patients 30 and 14 where minor to moderate tremor was observed and the model correctly identified them as patients. The level of tremor is based considering the average SARA FNT subscores of the three pediatric neurologists. Thus, there was assumed to be no tremor if the FNT score equals 0, minor tremor ($< 2cm$) if the score is less than 1, and moderate tremor ($< 5cm$) if the score is < 2. There were no scores of 3 ($> 5cm$) or 4 (unable to perform 5 pointing movements). Finally, participant 31 was included as an example of a healthy participant. To start with this participant, the trajectory is regular and smooth (Figure 7) with relatively high speed values and low curvature values (Figure 6). In strong contrast, participant 14, who exhibited "moderate" tremor (based on the average score 1.3 provided by the neurologists, amplitude between 2-5cm), had relatively low speed and high curvature values during very irregular trajectories resulting in a high model score. A similar, although more subtle, difference compared to the healthy participant (31) can be observed for participant 30, who exhibited minor tremor, but also had relatively low speed and high curvature values during trajectories that were also irregular, although less than for participant 14. This explains the high model score for this patient as well. Participants 4 and 24, who had no visible tremor, did have relatively low speed and high curvature values, while their trajectories looked very similar to those of the healthy participant (31), explaining the high model score as well as why no tremor was observed. Finally, participant 8, who had no visible tremor either, but was scored as healthy by the model, indeed had curvature and speed values that were very similar to those of the healthy participant (31).

In summary, the model seems to classify some patients with no visible or minor tremor as patients, because it picks up features from the movement trajectories that have been recorded by the IMUs and that are not visible to the naked eye. On the other hand, if the trajectory of a patient is similar to that of a healthy participant in terms of speed and curvature values, as may be the case for some of the (mildly affected) DCD patients, it seems the patient will be classified as healthy.

IV. Discussion

The goal of the present study was to apply a recently developed method for distinguishing between patients with coordination disorders and controls who performed the FNT. We expected that FNT movement trajectories would be smoother and faster for healthy participants than for patients and that these movement characteristics should be reflected in lower local curvature and higher instantaneous speed values in healthy participants, which was indeed confirmed. Using local curvature and instantaneous speed as features the method achieved 84% accuracy distinguishing patients and controls.

First a (probabilistic) GLM was defined as a function of curvature and speed to estimate the probability that the FNT trajectories were collected from a patient. Then, to test the GLM on new data we performed LOOCV resulting in 84% accuracy. In addition, we expected that misclassified patients would exhibit FNT trajectories similar to those of healthy participants, exhibiting smooth trajectories as reflected in low model scores. For further understanding of misclassifications, we plotted SARA scores against model scores, as well as violin plots of the local curvature and instantaneous speed distributions of selected participants. This suggested that the model classifies some patients with no visible or minor tremor as patients, because it detects features from the movement trajectories recorded by the IMUs that are not visible to the naked eye. On the other hand, if the trajectory of a patient is similar to that of a healthy participant in terms of speed and curvature, it seems the model classifies the patient as healthy.

Our accuracy results are consistent with other studies [17], [31], [32] that tried to distinguish between healthy and pathological FNT trials. Table I displays characteristics of earlier studies and ours using different techniques to distinguish patients from healthy participants, allowing a direct comparison. First, because we combined participants of two different studies, our study includes the largest number of healthy participants and patients, so far. Second, our method involved only two variables whereas the other methods [17],

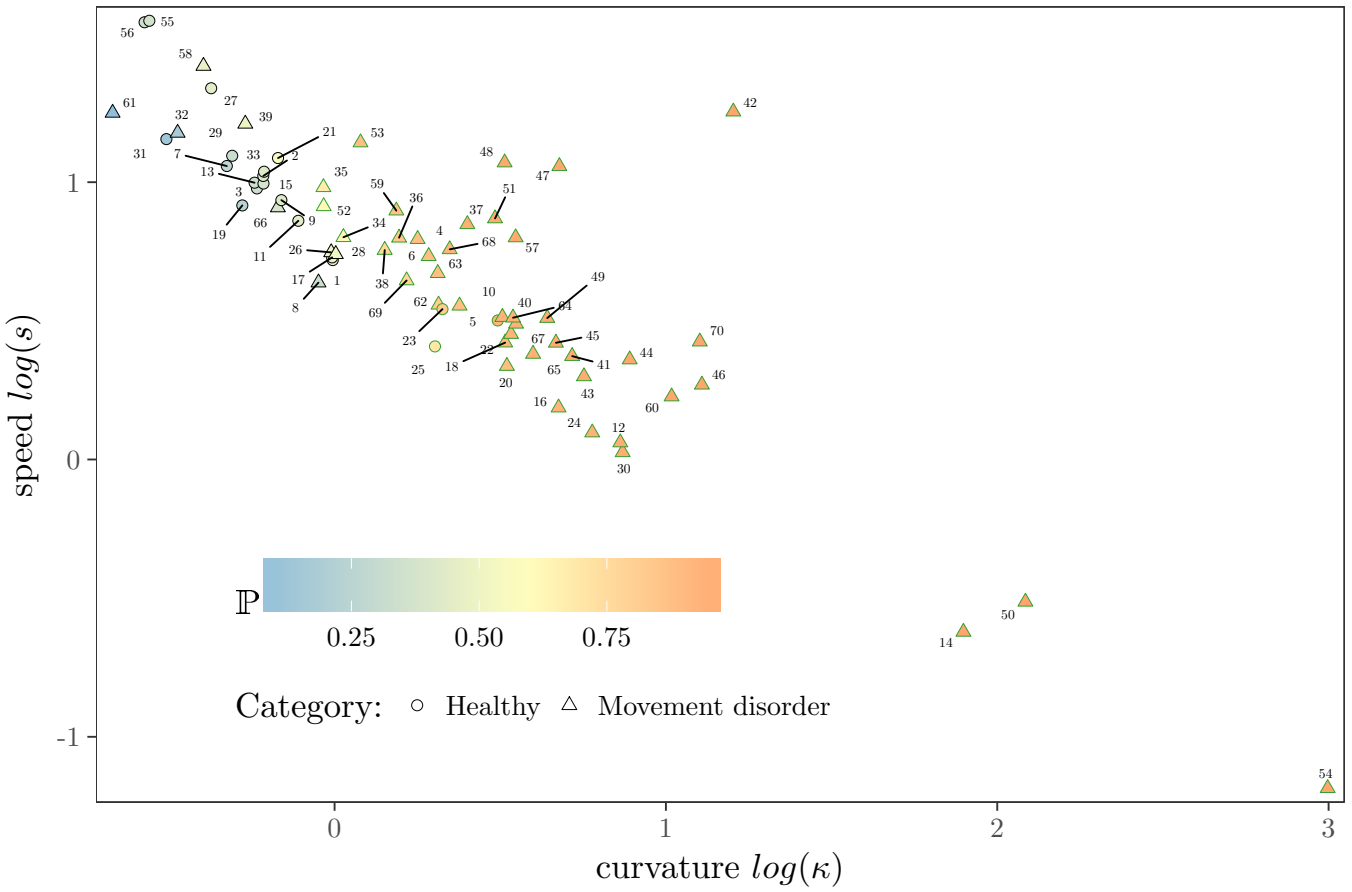

Fig. 4. Visualization of the predictions of model (1). Black edges of the shapes represent participants classified as healthy, while green edges represent participants classified as patients.

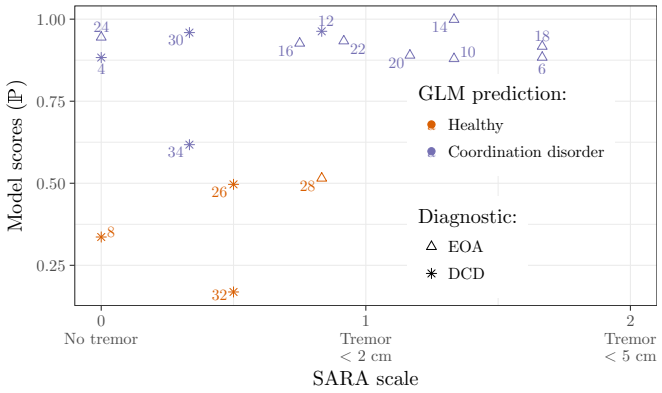

Fig. 5. Model scores against SARA scores. SARA scores are averaged over left and right hands and observers. The vertical axis indicates the probability of having a coordination disorder. The numbers represent participants. The misclassified patients from group 1 are in the lower left corner of the plot (numbers 8, 26, 28 and 32).

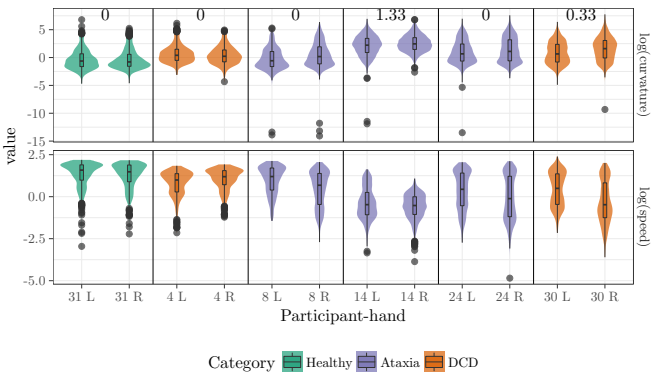

Fig. 6. Violin plots of the distribution of curvature and speed values for 6 participants. L: left hand, R: right hand. The numbers on top represent SARA scores. Violin plots provide a way to visualize the differences between a healthy participant and patients in terms of distributions of curvature and speed values. For example, clear differences can be noticed (visually) between the healthy participant and patients 14 and 30. Statistical descriptors are included in terms of the boxplots showing medians and interquartile ranges. Medians show that the healthy participant had higher speed values than patients. In contrast, the healthy participant scored lower curvature values than patients.

[31], [32] involved 11 variables or more. Third, because of the relatively large number of variables the other methods used dimensionality reduction techniques such as PCA and Sammon's map. As our method involves only two variables,

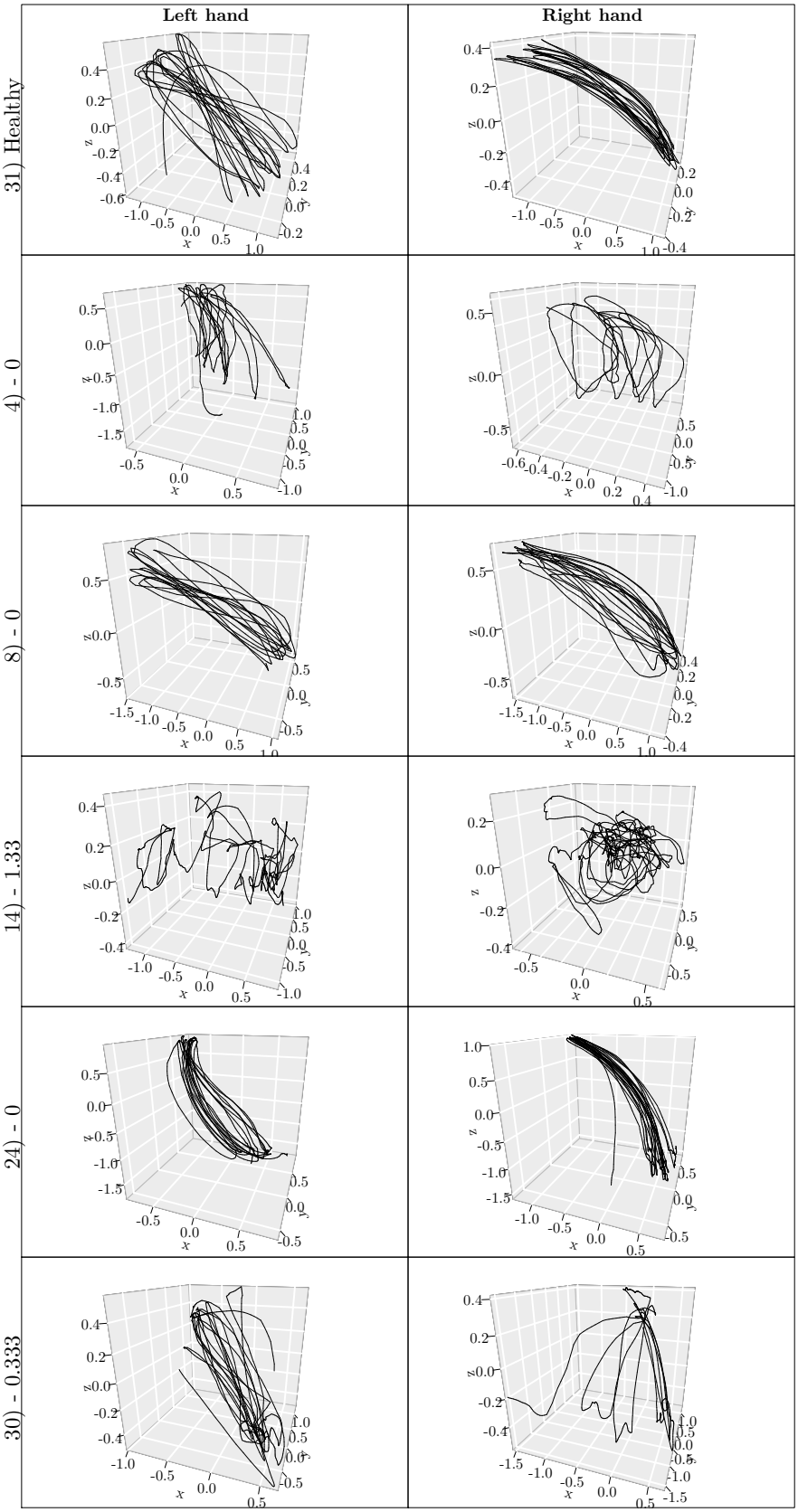

Fig. 7. 3D trajectories collected from a healthy participant and 5 patients with a diagnosed coordination disorder. The labels on the left represent participants and SARA scores. The values of the axis are unitless, as the upper limb model provided normalized three-dimensional coordinates based on the orientation of the IMUs, see [17] for further mathematical details.

TABLE I
MAIN FEATURES OF RECENT FNT STUDIES TRYING TO DISTINGUISH HEALTHY PARTICIPANTS AND PATIENTS WITH A COORDINATION DISORDER. KFCV: K-FOLD CROSS VALIDATION.

Study	Controls	Patients	Variables	Techniques	Accuracy
[31]	10	10	16	PCA, k-means MS-clustering	91.4% controls 98.5% patients
[32]	10	28	16	Sammon's map, KFCV	72%
[17]	16	11	14	PCA, LOOCV, random forest	83% controls 78.5% patients
present	20	68	2	GLM, ROC LOOCV	84% controls 84% patients

there was no need to apply such techniques allowing for straightforward visualization and interpretation of the results. The other studies apply different classification algorithms such as random forest and k-means. One advantage of GLMs is that the probabilistic outcomes are useful not only for classification (using an optimal threshold) but also for gaining insight into the performance of the FNT. For example, regardless of a clinical diagnosis, values close to 0 would suggest "good" performance during the test while values close to 1 would indicate "bad" performance during the test. Fourth, the classification accuracy is on average 84.4% with SD 8.5% among the four studies. The highest accuracy (more than 91%) was reported by [31]. We believe that this highest accuracy can be explained in part by the smallest number of participants, 10 controls and 10 patients, among the four studies and because the patients might have had more severe coordination disorders. As a result, there was less overlap in study [31] between patients and controls than in the other studies, including ours.

It is remarkable that by using only two features we still achieve high accuracy levels as in studies in Table I. We suggest that because our method strongly focuses on local features of the FNT trajectories, the method is able to detect smaller changes than when using global features (as mostly used in [17], [31], [32]). In addition, we also believe that local curvature, which to our knowledge has not been used before to study FNT trajectories, also plays an important role to assess the smoothness of the FNT trajectories. These aspects of our method (not fully present in the others) and although only two features were used, thus lead to results comparable to those of the state of the art.

IMUs are practical devices having several benefits such as being affordable, non-invasive and the whole setup can be placed in the same location where clinical evaluation takes place. An advantage of our method is that after data collection, only simple preprocessing steps are performed such as smoothing of the signal (no need to apply PCA, for example) and segmentation of the signal by removing extremes of the signal which are not part of test. There is no need to detect features such as peaks or inflection points. GLMs offer the possibility to include many variables in the models and the probability values as outcomes have meaningful and straightforward interpretation. In addition, the threshold estimated using the ROC curve (Figure 3) can already be considered as a cut-off score. These features of the method are promising but for clinical application this approach needs to be developed further because calibration is limitative now and it is not fully automatic. Additionally, the main question is not only to distinguish healthy persons from patients with a coordination disorder, but also to distinguish between different movement disorders. This was a first step towards that goal.

One limitation of the present study is that we pre-assigned the subjects to the EOA or DCD group on clinical grounds. In absence of a gold standard, we can never exclude the possibility that clinical assignment is not 100% correct, either. Another limitation is that the severity of ataxia in our group of participants, as assessed by neurologists, does not cover the whole range of the SARA FNT scale. The highest score provided by the neurologists is not higher than 2 (tremor smaller than 5 cm), whereas the maximum score is 4 (unable to perform the pointing movements). Including participants with more severe symptoms of a coordination disorder may change the results. However, the classification accuracy should be similar, as more severe symptoms should result in even higher curvature values and slower speed values. In other words, we could expect such participants to be in the lower right corner of Figure 4 (high curvature and slow speed values), where we expect the model to classify them as patients.

The need for objective and quantitative assessment of human movement to reinforce and support the use of clinical rating scales is evident [5], [33], [34]. A benefit of the presented method is that it can be applied to a broad range of human movements commonly used in clinical tests such as gait, static postural control, dynamic postural control, finger chasing, path drawing spirals, circles, squares, or figure-8 shapes, and fast alternating hand movements [11]. In addition, the methodology can be applied independently of the tracking device such as force plates or depth cameras.

V. CONCLUSION

In conclusion, we have shown that the (probabilistic) GLM we developed to assess dynamic balance can also be used to assess patients with coordination disorders. The GLM as a function of only two features (instantaneous speed and local curvature) is useful to distinguish patients and healthy participants based on an instrumented version of the FNT. The quantification of smooth movements plays an important role in the assessment of coordination disorders [35]. In previous studies [7], [20] we proposed curvature as a measure of smoothness of body movements. Here, we provide additional evidence of the usefulness of this measure to assess body movements and to differentiate pathological from healthy movements. Future work could entail adding more features (local and global) to the GLM, testing the reliability of the method for distinguishing groups of patients and/or controls using other clinical tests such as finger chasing or quiet standing, and using other tracking devices such as force plates or depth cameras that can be used to track body movements without the need for markers or wearable measurement devices.

REFERENCES

[1] N. Yang, M. Zhang *et al.*, "Motion quality evaluation of upper limb target-reaching movements," *Medical Engineering and Physics*, vol. 24, no. 2, pp. 115–120, 2002.

[2] A. D. L. Reyes-guzmán, I. Dimbwadyo-terrer *et al.*, "Quantitative assessment based on kinematic measures of functional impairments during upper extremity movements: A review," *Clinical Biomechanics*, vol. 29, no. 7, pp. 719–727, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.clinbiomech.2014.06.013

[3] F. Venetsanou, A. Kambas *et al.*, "Can the Movement Assessment Battery for Children-Test be the "gold standard" for the motor assessment of children with Developmental Coordination Disorder?" *Research in Developmental Disabilities*, vol. 32, no. 1, pp. 1–10, 2011.

[4] U. Röijezon, R. Faleij *et al.*, "A new clinical test for sensorimotor function of the hand – development and preliminary validation," *BMC Musculoskeletal Disorders*, vol. 18, no. 1, p. 407, 2017. [Online]. Available: http://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/s12891-017-1764-1

[5] C. E. Dean, J. M. Russell *et al.*, "Clinical Rating Scales and Instruments," *Journal of Clinical Psychopharmacology*, vol. 24, no. 3, pp. 298–304, 2004. [Online]. Available: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00004714-200406000-00010

[6] V. Soancatl Aguilar, J. J. van de Gronde *et al.*, "Assessing Dynamic Balance Performance during Exergaming based on Speed and Curvature of Body Movements," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 171–180, 2018.

[7] V. Soancatl Aguilar, C. J. C. Lamoth *et al.*, "Assessing Dynamic Postural Control During Exergaming in Older Adults: a Probabilistic Approach," *Gait & Posture*, vol. 60, no. 1, pp. 235–240, 2018.

[8] A. F. Zuur, E. N. Ieno *et al.*, *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer Science + Business Media, 2009.

[9] J. Kreutzer, D. John, and C. Bruce, *Encyclopedia of Clinical Neuropsychology*, 1st ed. New York: Springer-Verlag, 2011.

[10] P. Trouillas, T. Takayanagi *et al.*, "International cooperative ataxia rating scale for pharmacological assessment of the cerebellar syndrome," *Journal of the Neurological Sciences*, vol. 145, no. 2, pp. 205 – 211, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022510X96002316

[11] T. Schmitz-Hübsch, S. T. Du Montcel *et al.*, "Scale for the assessment and rating of ataxia," *Neurology*, vol. 66, no. 1, pp. 1717–1720, 2006. [Online]. Available: http://www.neurology.org/content/66/11/1717.abstract

[12] A. Mannini, O. Martinez-Manzanera *et al.*, "Automatic classification of gait in children with early-onset ataxia or developmental coordination disorder and controls using inertial sensors," *Gait & Posture*, vol. 52, pp. 287–292, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.gaitpost.2016.12.002

[13] P. Morasso and V. Tagliasco, Eds., *Human Movement Understanding. From Computational Geometry to Artificial Intelligence*. Amsterdam: Elsevier Science Publishers, 1986.

[14] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[15] S. Watanabe, "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory," *Journal of Machine Learning Research*, vol. 11, pp. 3571–3594, 2010. [Online]. Available: http://www.jmlr.org/papers/volume11/watanabe10a/watanabe10a.pdf

[16] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *International Joint Conference on Artificial Intelligence*, 1995.

[17] O. Martinez Manzanera, "Quantification of symptoms of movement disorders employing motion sensors," Ph.D. dissertation, University of Groningen, 2017.

[18] A. E. Harding, "Clinical features and classification of inherited ataxias," *Advances in Neurology*, vol. 61, pp. 1–14, 1993.

[19] T. F. Lawerman, R. Brandsma *et al.*, "Age-related reference values for the pediatric scale for assessment and rating of ataxia: a multicentre study," *Developmental Medicine & Child Neurology*, vol. 59, no. 10, pp. 1077–1082, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/dmcn.13507

[20] V. Soancatl Aguilar, J. J. van de Gronde *et al.*, "Visual Data Exploration for Balance Quantification in Real-Time During Exergaming," *PLOS ONE*, vol. 12, no. 1, p. e0170906, jan 2017. [Online]. Available: http://dx.plos.org/10.1371/journal.pone.0170906

[21] S. Wu and Y. F. Li, "On Signature Invariants for Effective Motion Trajectory Recognition," *The International Journal of Robotics Research*, vol. 27, no. 8, pp. 895–917, 2008. [Online]. Available: http://ijr.sagepub.com/content/27/8/895.full.pdf

[22] D. R. Cox, "The Regression Analysis of Binary Sequences," *Journal of the Royal Statistical Society*, vol. 20, no. 2, pp. 215–242, 1958. [Online]. Available: http://www.jstor.org/stable/2983890

[23] R. McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC, 2015. [Online]. Available: https://www.crcpress.com/Statistical-Rethinking-A-Bayesian-Course-with-Examples-in-R-and-Stan/McElreath/p/book/9781482253443

[24] A. Gelman, D. Lee, and J. Guo, "Stan: A Probabilistic Programming Language for Bayesian Inference and Optimization," *Journal of Educational and Behavioral Statistics*, vol. 40, no. 5, pp. 530–543, 2015.

[25] G. James, D. Witten *et al.*, *An introduction to Statistical Learning with Applications in R*. New York: Springer Science + Business Media, 2013.

[26] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, no. 1, pp. 32–35, 1950.

[27] T. Fawcett, "An introduction to ROC analysis Tom," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.

[28] X. Robin, N. Turck *et al.*, "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformatics*, vol. 12, no. 1, pp. 1–8, 2011. [Online]. Available: http://dx.doi.org/10.1186/1471-2105-12-77

[29] P. L. Gribble and D. J. Ostry, "Origins of the power law relation between movement velocity and curvature: modeling the effects of muscle mechanics and limb dynamics." *Journal of neurophysiology*, vol. 76, no. 5, pp. 2853–2860, 1996.

[30] C. Sharpsteen and C. Bracken, *tikzDevice: R Graphics Output in LaTeX Format*, 2017, r package version 0.10-3. [Online]. Available: https://github.com/yihui/tikzDevice

[31] M. Velázquez-Mariño, M. Atencia *et al.*, "Cluster Analysis of Finger-to-nose Test for Spinocerebellar Ataxia Assessment," in *IWANN 2015. Lecture Notes in Computer Science*, C. A. Rojas I., Joya G., Ed. Springer, Cham, 2015, pp. 524–535.

[32] A. R. P. Machado, H. C. Zaidan *et al.*, "Feature visualization and classification for the discrimination between individuals with Parkinson's disease under levodopa and DBS treatments," *BioMedical Engineering OnLine*, vol. 15, no. 1, p. 169, 2016. [Online]. Available: http://biomedical-engineering-online.biomedcentral.com/articles/10.1186/s12938-016-0290-y

[33] B. R. Swaine and S. J. Sullivan, "Reliability of the scores for the finger-to-nose test in adults with traumatic brain injury." *Physical therapy*, vol. 73, no. 2, pp. 71–8, 1993. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/8421720

[34] M. van Dijk, N. Smorenburg *et al.*, "How clinicians analyze movement quality in patients with non-specific low back pain: a cross-sectional survey study with Dutch allied health care professionals," *BMC Musculoskeletal Disorders*, vol. 18, no. 1, p. 288, 2017. [Online]. Available: http://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/s12891-017-1649-3

[35] S. Balasubramanian, A. Melendez-Calderon *et al.*, "On the analysis of movement smoothness," *Journal of NeuroEngineering and Rehabilitation*, vol. 12, no. 1, p. 112, 2015. [Online]. Available: http://www.jneuroengrehab.com/content/12/1/112