



Getting on the Right Foot: Using Observational and Quantitative Methods to Evaluate Movement Disorders

James Spann
js Spann2@cs.rochester.edu
University of Rochester
Rochester, New York, USA

Tetsuo Ashizawa
tashizawa@houstonmethodist.org
Houston Methodist
Houston, Texas, USA

Sarah Chen
schen137@u.rochester.edu
University of Rochester
Rochester, New York, USA

Ehsan Hoque
meHoque@cs.rochester.edu
University of Rochester
Rochester, New York, USA

ABSTRACT

Currently doctors rely on tools such as the Unified Parkinson's Disease Rating Scale (MDS-UDPRS) and the Scale for the Assessment and Rating of Ataxia (SARA) to make diagnoses for movement disorders based on clinical observations of a patient's motor movement. Observation-based assessments however are inherently subjective and can differ by person. Moreover, different movement disorders show overlapping symptoms, challenging neurologists to make a correct diagnosis based on eyesight alone. In this work, we create an intelligent interface to highlight movements and gestures that are indicative of a movement disorder to observing doctors. First, we analyzed the walking patterns of 43 participants with Parkinson's Disease (PD), 60 participants with ataxia, and 52 participants with no movement disorder to find ten metrics that can be used to distinguish PD from ataxia. Next, we designed an interface that provides context to the gestures that are relevant to a movement disorder diagnosis. Finally, we surveyed two neurologists (one who specializes in PD and the other who specializes in ataxia) on how useful this interface is for making a diagnosis. Our results not only showcase additional metrics that can be used to evaluate movement disorders quantitatively but also outline steps to be taken when designing an interface for these kinds of diagnostic tasks.

CCS CONCEPTS

• Human-centered computing → Visualization systems and tools.

KEYWORDS

Multi-modal Interfaces, Intelligent visualization, Quantitative methods

ACM Reference Format:

James Spann, Sarah Chen, Tetsuo Ashizawa, and Ehsan Hoque. 2024. Getting on the Right Foot: Using Observational and Quantitative Methods to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IUI '24, March 18–21, 2024, Greenville, SC, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0508-3/24/03
<https://doi.org/10.1145/3640543.3645160>

Evaluate Movement Disorders. In *29th International Conference on Intelligent User Interfaces (IUI '24)*, March 18–21, 2024, Greenville, SC, USA. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3640543.3645160>

1 INTRODUCTION

For most movement disorders there is no definitive test to verify a prognosis for the disorder, and for Parkinson's disease (PD) there is no standardized test at all. Doctors currently rely on tools such as the Unified Parkinson's Disease Rating Scale (MDS-UDPRS) [10, 33] to make diagnoses based on clinical observations of a patient's motor movement, rated on a scale from "Normal" to "Severe" on one out of four checkboxes. Similarly, the Scale for the Assessment and Rating of Ataxia (SARA) rubric [31] measures a patient's observed performance on a series of tasks, generating a total score between 0 and 40.

Basing patient diagnosis on human eyesight alone can be problematic and often leads to misdiagnosis or a waste of resources. One challenge is the use of Likert rating scales by physicians, whose interpretations of the scale are inherently subjective and may vary on a doctor-by-doctor basis. What one doctor sees as "Moderate" another might see as "Severe" and can change a patient's diagnostic outcome. This problem is made worse when considering that the tremors appearing with Parkinson's disease can appear to be visually indistinguishable from those of another movement disorder like Ataxia. Ataxias have been known to be misdiagnosed for other disorders that have similar forms of seizures, such as epilepsy [34]. Similarly, 15% of patients with tremors have been misdiagnosed as having PD when they did not fully meet the clinical criteria for the disease [32], while in other studies as many as one in three patients met the criteria for PD but were diagnosed with something else [13].

Measuring the subtlety of motions associated with movement disorders based on eyesight alone is not sufficient for what can be a life-altering diagnosis. Previous studies have tried to address this issue with on-body sensors or computer vision to unilaterally distinguish between people who have a movement disorder and people who do not, but these systems are not designed to categorize between multiple movement disorders. To the best of our knowledge, there is no prior work that investigates comparing the movements of patients against multiple movement disorders and a baseline control group (no movement disorder).

We propose creating an interface that can provide objective gait metrics during the assessment process and high-level interpretations of those metrics. We seek to amplify the ability of experienced specialists by showing them physiological information about the patient, in the context of how similar their gestures are to those of a patient with a particular movement disorder. This allows patients to be more informed about their diagnosis with more transparency in how a doctor could have arrived at a diagnosis and helps to support an objective standard for movement disorder diagnoses.

In this paper, we explore how an interface can assist diagnostic practitioners when making a diagnosis for a particular movement disorder. We describe our contributions as follows:

- (1) We report metrics, grounded in clinical literature, for distinguishing patients with Parkinson's disease, ataxia, or neither (Control), that can be found with computer vision, but may be difficult to distinguish observationally.
- (2) We create an interface that helps the physician visualize the distinguishing metrics and insights about the patient's diagnostic task.
- (3) We assess how useful this interface is by interviewing doctors who are specialized in diagnosing Parkinson's disease and ataxia about their experience with it.

2 RELATED WORK

2.1 Movement Disorders and their Tools for Diagnosis

Parkinson's Disease is a neurodegenerative disorder marked by involuntary tremors and limb stiffness [5, 8, 17]. Bradykinesia, the slow movements such as dragging of feet or freezing of facial expressions, is known to be a parkinsonian trait, and this stiffness is expected to be present when diagnosing a patient for the disease [5]. To diagnose a patient with Parkinson's Disease, specialized doctors utilize the criteria created by the International Parkinson and Movement Disorder Society (MDS), known as the MDS-UPDRS [10, 33]. The MDS-UPDRS provides questions to evaluate the motion of a patient and rate it on a five-point scale ranging from "Normal" to "Severe".

There are several kinds of Ataxias, each with its own grouping of symptoms, however, they all are degenerative diseases that affect the central nervous system [4]. They generally affect gait, balance, and can progressively cause nerve damage [4, 23]. To evaluate a patient's likelihood of Ataxia, diagnosticians use the SARA scale [31] which measures a patient on 8 categories (Gait, Stance, Sitting, Speech Disturbance, Finger Chase, Nose-finger test, Fast alternating hand movement, Heel-shin slide) on a Likert scale (either four, six, or eight choices depending on the category).

When evaluating a patient for either Parkinson's disease or Ataxia, the MDS-UPDRS and SARA criteria respectively assist in deciding how severe a patient's symptoms are, relative to a patient who does not have a movement disorder. These criteria however do not rule out the possibility of another movement disorder (i.e., a patient who is being tested for Parkinson's disease but has Ataxia). While there are distinctive traits for these movement disorders, such as Ataxia causing the gait cycle to be more variable [11] and Parkinson's disease-causing more rigidity in movement [5], it can be difficult for doctors to notice the subtle differences in movement

tremors. The difficulty in noticing these differences can lead to misdiagnosis, and in one study 15% of patients were misdiagnosed for Parkinson's disease when they did not fully meet the criteria for the disease [32].

2.2 Automated Tools for Diagnosis of Movement Disorders

To compensate for the disadvantages that come with making a diagnosis with eyesight alone, researchers have explored using computer vision [12, 20, 27] and sensors [7, 22] to make assessments on the severity of movement disorders or to predict the movement disorder altogether [2, 3, 29]. On-body sensors provide a means to get information about a diagnostic task in real-time. To evaluate its use for classifying Friedreich's Ataxia, LeMoyné et al. [16] measured gait with inertial sensors. Another study done by Phan et al. [26] placed IMU sensors on the chest and ankles of 29 ataxia patients, to generate an Ataxia severity score.

The ubiquity of smartphones has prompted studies to also look at how they can be used to detect Movement Disorders. The GEORGE smartphone application [35] used sensor data to assess individuals for Huntington's disease by using chorea, finger tap rate, and step count. Similarly, the Roche PD Mobile app [18] passively used the phone's gyroscopes and accelerometers to distinguish between participants with Parkinson's disease and participants who did not.

As Work-from-Home becomes a growing option for employees, at-home testing looks to be the next trend of diagnostic procedures. Liu et al. [19] explored using radio waves, similar to that of a Wi-Fi router, to passively monitor a patient's movement trajectories and gait speeds as they move around their home. This tool was able to successfully keep track of patients and found gait speed to be correlated to Parkinson's Disease severity.

Langevin et al. [15] addressed individuals' inability to see a neurologist by creating a tool to let participants receive a remote diagnosis at home using a webcam. The framework assesses participants' facial features and motor movements, such as finger tapping and hand movements, to study which traits were closely associated with Parkinson's disease and provide analysis on the usability of a web-based application for diagnosis.

Our approach builds on these concepts by distinguishing between the two movement disorders and presenting relevant movement information to make it easier to infer which is more likely, rather than classifying a patient as one or another. Unlike previous work, our dataset uses videos from a frontal plane, similar to how a specialist would observe a patient in a clinical setting. Similarly, rather than provide an immediate classification of a patient's movements, our work quantifies a patient's gestures and objectively conveys that evidence in an assistive format. Lastly, to the best of our knowledge, there is no other work that creates diagnostic interfaces for movement disorder specialists and evaluates them.

In the following sections, we describe our study in full. In section 3 we outline our dataset and the demographics that constitute it. In section 4 we describe how we created features to distinguish patients and the interface we designed to show those features. In section 5 we report our findings and provide more context to them in section 6.

3 DATA

3.1 Data Overview

Our initial dataset is made up of two groups of patients doing a diagnostic task for either Parkinson’s disease or Ataxia. In total, we have 45 minutes (13 minutes [23,439 frames] PD, 32 minutes [57,733 frames] Ataxia/Control) of videos of patients doing the same task of walking back and forth down a hallway, recorded from a frontal perspective. On average, these videos are 5.6 seconds long. The patients are recorded at different intervals of time (e.g. 6 months, 12 months, 24 months) since their initial visit.

Working with doctors from two major medical centers in the USA, we obtained a dataset of patients with Parkinson’s disease. The initial dataset of videos of patients with Parkinson’s disease is made up of 129 videos of 51 patients who were completing a diagnostic gait task which is evaluated by the criteria in sections “3.10 - Gait” and “3.11 - Freezing of Gait” of the MDS-UPDRS [10, 33]. After excluding videos that did not fully contain the patient completing the walking task and splitting the videos based on the direction the patient was walking, as described in section 3.3, we had 65 videos of patients walking towards the camera and 64 videos of patients walking away from the camera, for Parkinson’s disease.

For our Ataxia dataset, we worked with 11 clinical sites across eight states in the United States to collect videos of patients doing a diagnostic gait task. These walks were evaluated against the “Gait” section SARA rating scale [33] which, like the MDS-UPDRS, is evaluated via eyesight alone. The videos were processed using the same methods as the Parkinson’s disease dataset’s videos, as described in section 3.3, and from the initial 177 videos recorded, we obtained 98 videos walking away from the camera and 95 walking towards the camera.

Lastly, we used the same methods above to create a dataset of Control diagnostic walks (patients who were not observed to have any movement disorder), which was gathered in the same settings as the Ataxia dataset. This resulted in 71 videos of patients walking away from the camera and 71 videos of a patient walking towards the camera.

3.2 Data Acquisition

The patients for the Parkinson’s dataset group were recruited via clinician referrals, registry postings, and flyers. Participants who left their contact information were then screened over the phone for eligibility, and if eligible, were invited to an in-person screening visit where they completed an informed consent form. This study and its data security procedures were reviewed and approved by our institution’s Human Subjects Review Board (IRB).

When recording patients completing the walking task in a clinic, we specified a study-specific video camera to be used, and that as much of the patient’s body should be recorded. For virtual visits, the same recording procedures are used and are recorded over Zoom. The videos are captured and handled by the clinical coordinator and stored securely on the secure data storage service Box ¹.

For our Ataxia dataset group, we used data from a prior study where patients were recorded during their annual visits with a neurologist during which they completed the SARA Gait task[28]. The

participants involved in the study were diagnosed with Spinocerebellar Ataxia types 1 or 3 (SCA1 and SCA3). Criteria concerning genotypes consistent with inherited ataxia, diseases that may affect Ataxia diagnosis, and changes in physical and occupational ataxia therapy, were used to assess eligibility for the study. Patients were recorded in a clinical setting. After recording, the participants’ faces are blurred in the video using Google’s Mediapipe mesh tool to prevent identification.

3.3 Data Cleaning

To get the videos of patients into a format for analyzing their movements, we broke the data-cleaning task into several steps. First, we manually clip the full-length videos into sections based on whether the patient is walking towards the camera or away from it. Second, we use a DeepSort-YOLOv3 Model ² trained on humans to identify people in the video and use that information to remove everything in the video that is not the detected person. Next, to filter out which detected person was the patient we measure how the area of the bounding box around the person changed during the video, with the person who had the largest decrease being categorized as the “patient”. This step was necessary as the doctor or several nurses could sometimes be seen standing in the video, watching the patient complete the diagnostic walking task. To ensure the filtering method worked correctly, the resulting videos were manually checked against the originals to confirm only the patients were selected. Finally, we then run CMU Openpose [6] on each frame of the video to record the movements of the patient in each frame of the video. Our study was run on the resulting key point information that was found for the patient.

4 METHODS

In this section, we describe the features that were found to distinguish patients of PD/Ataxia/Control groups, and how the interface was designed to show this information to diagnosticians.

4.1 Computed Features

To find features that are relevant to the diagnostic process to help distinguish between movement disorders with visually indistinguishable tremors, we consulted previous studies, clinical literature, and current rubrics to find attributes relevant to detecting Parkinson’s disease [5, 8, 14, 17, 23] and Ataxia [1, 4, 25, 30]. We found most visual features were related to the rigidity or range of motion associated with a patient’s limbs. We ran a trial and error process of finding statistically significant features that would distinguish patients of certain groups from other groups. These features were curated by evaluating statistical significance across all six patient groups (PD/Ataxia/Control and Away/Towards) and evaluated via a series of ANOVA tests to assert their effectiveness when distinguishing patients from different groups. To verify these results were not biased by individual patients, we repeated this statistical analysis five times with a random 20% of each participant group of videos removed from the compared videos, which led to the same conclusions.

We found ten features that are statistically significant to compute to diagnostic specialists:

¹<https://box.com>

²<https://github.com/anushkadhiman/ObjectTracking-DeepSORT-YOLOv3-TF2>

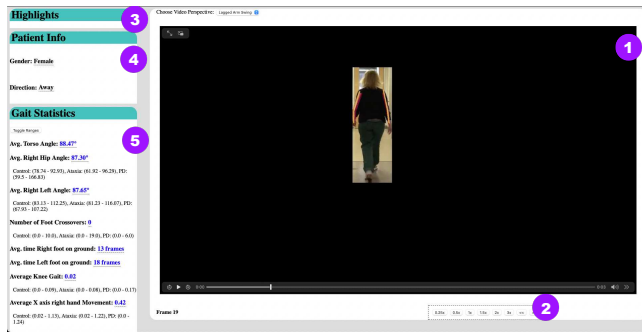


Figure 1: Interface to see information about the participant's diagnostic task. (1) A view to watch the participant's walk with 3 different kinds of overlays, (2) Controls to speed up, slow down, or step through the video, (3) A panel containing natural language descriptions of significant attributes of the patient's diagnostic walk, (4) A panel containing demographic information about the participant and the direction the participant is walking in, (5) A panel containing computed statistics about the participant's walk.

Left/Right Wrist Movements - The difference of the Euclidean distance of the movement of the patient's wrist, relative to their shoulder width

Foot Crossover Count - The number of times the patient's foot crosses in front of the other in the X-axis.

Right/Hip Angle - The average angle between the knee and mid-waist key points (i.e. the hip), for every ten frames.

Knee Gait - The distance between the patient's knees, relative to the width of the person's waist.

Pause time while Walking - The number of frames a patient stopped walking. This was computed for videos with at least 30 frames.

Shoulder Waist Alignment - The average distance between the mid-shoulder and mid-hip of the patient's body.

Left/Right Foot relative Travel - The height of the person's foot relative to their body height. This was computed for videos with at least 30 frames.

The first three could be found with just eyesight alone to assist when making a diagnosis while the remaining features require a computer to find the values.

4.2 Interface Design

We designed a web page interface to show a diagnostician the values of features described in Section 4.1 which can be used to distinguish patients who have PD, Ataxia, or neither. The interface displays information for one patient at a time and only displays videos and statistics for that patient. Our interface's design was based on the principles of presenting all of the information about a patient's walk in one central location. We also included the video of the patient walking, as we were inspired by the aid that visual representations of the data can provide during a diagnosis [9].

The center of the interface shows a video of a patient in one of four perspectives: the unaugmented video of the patient walking ("Original"), walking with an Openpose skeleton overlay ("Full



Figure 2: The interface tool shows the video of the patient's movements in three different perspectives: "Lagged Arm Swing", "Ankle View" and "Full Pose", shown here from left to right respectively. In the interface, this annotated video would appear in part (1) of Figure 1.

Pose"), with a line drawn between the patient's ankles ("Ankle View"), and with a colored overlay of the range of the last five locations of the patient's arms ("Lagged Arm Swing"). The Mediapipe framework [21] was used to overlay the highlights used in the Full Pose and Lagged Arm Swing views. These views were added to give a diagnostician a visual interpretation of how the patient's walking changes during the diagnostic walking task. If the diagnostician wants to look at these videos more closely, they also have the controls to speed up (1.5x, 2x, or 3x), slow down (0.25x or 0.5x), or step frame-by-frame through the video.

The left column of the interface contains three collapsible sections with information related to the patient and their walking task: "Highlights", "Patient Info" and "Gait Statistics".

The *Highlights* section provides full-sentence interpretations of distinguishing features for the observed patient. For example, if the patient had a knee gait width that is more like that of a patient with Ataxia, the phrase "Knee Gait - Higher value is more indicative of Ataxia" would be automatically added to the section.

The *Patient Info* section provides demographic information about the patient, including age (if available) and gender, and video information such as the direction that the patient is walking in the video.

The *Gait Statistics* section outlines the values for the features found in Section 4.1 along with the ranges of those values for PD, Ataxia, and Control Patients. For each instance where a foot crossover occurs, the interface also displays the frame numbers where that occurs, which the user can jump to in the patient's video.

To add additional transparency, if the user hovers over a feature's value for the given patient, the interface will display information about how that feature was calculated.

5 RESULTS

We evaluated our computed features against all six categories of video (PD/Ataxia/Control, walking away or towards the camera). The t-test comparisons of these features are shown in Table 1. Further, we measured the effectiveness of these features in distinguishing these patient groups by computing an ANOVA test for

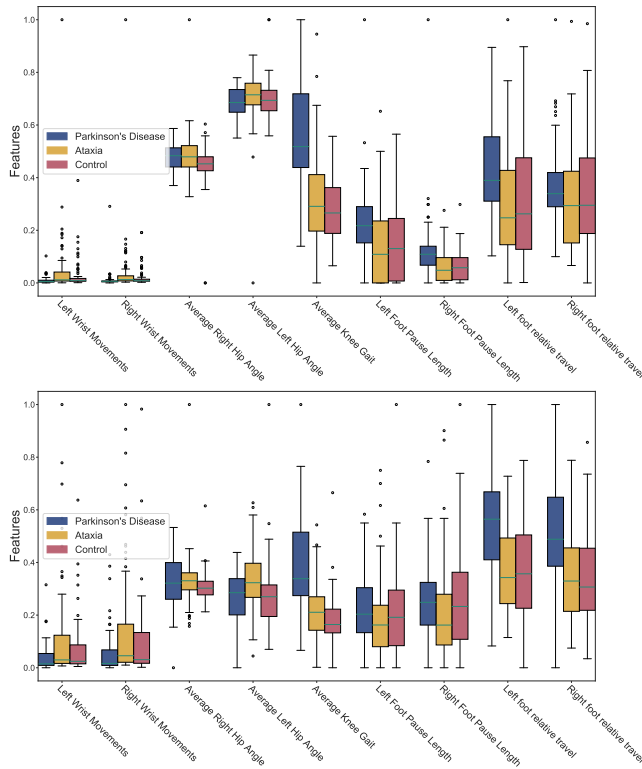


Figure 3: Normalized features for all of the patients as they completed their gait tasks walking towards (Top) and Away (Bottom) from the camera

each feature, the results of which can be seen below in Table 2, and a Kruskal test which can be seen in Table 3. For all of our statistical tests, we used a significance threshold of $p < 0.05$.

We additionally evaluated the video based on the number of times one of a patient's feet crossed in front of another. A chi-square test was used under the hypothesis that the proportions of patients surveyed who had one foot cross in front of the other while walking at least once was equal. We looked at this for the group of patients walking towards the camera, away from the camera, and the entire set of patients as a whole. We found the proportions to be unequal for the patients who were walking away from the camera ($p = 0.000138$) and for the entire population of the patients ($p = 7.190822e-05$), but not for the participants who were walking towards the camera ($p = 0.067138$). The plots for the observed and expected values of these tests can be found in Figure 4.

6 DISCUSSION

6.1 Finding Features to Differentiate Movement Disorders

Our study found different features that can be used for detecting Parkinson's disease from Ataxia, and both movement disorders from a control group. We sought out features that would be consistently significant ($p < 0.05$) when a participant was walking towards a camera or away from it. This occurred for Control/PD

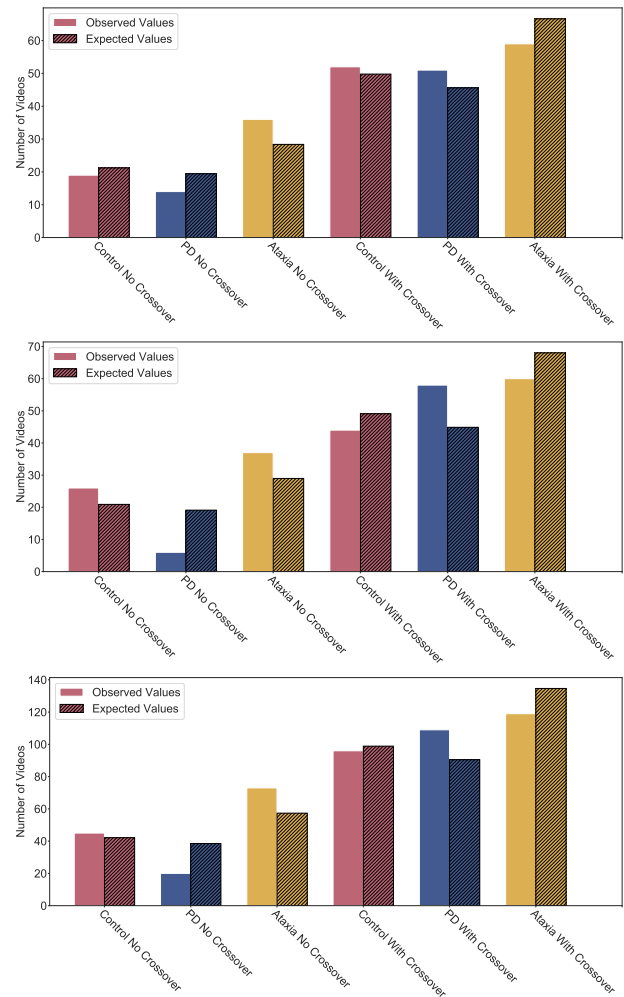


Figure 4: Chi-Squared test of participants having one foot cross in front of the other while walking at least once when walking towards the camera (Top), away from the camera (Middle), or both sets of videos together (Bottom).

with the Average Knee Gait and relative travel with the patient's left foot. We also saw the patient's hip movements as a relevant statistic when comparing Control/Ataxia in either direction. The most attributes that could be used to distinguish between the patient groups in either direction were found between the PD/Ataxia groups, where four features (Left wrist movements, Average Knee Gait, and left/right relative foot travel) could be used to distinguish patients in either direction.

When looking at the ability to distinguish the patients when moving away from the camera we found Average Knee gait to be a successful metric. We attribute this to participants in the control group having a significantly lower mean knee gait relative to their body size, represented by shoulder width. This could be explained by the smaller stance that is needed to attain better balance, as would be the opposite case with a movement disorder that is prone to rigidity and stiffness. Similarly, the consistency of relative foot travel as a

	Away			Towards		
	Control PD	Control Ataxia	PD Ataxia	Control PD	Control Ataxia	PD Ataxia
Left Wrist Movements	0.055	0.311	0.000766	0.0191	0.75	0.01
Right Wrist Movements	0.182	0.448	0.00168	0.329	0.634	0.113
Average Right Hip Angle	0.532	0.0537	1.75	0.029	0.0248	2.18
Average Left Hip Angle	2.43	0.00592	0.000147	0.784	1.24	0.162
Average Knee Gait	2e-12	0.123	8.23e-10	3.25e-18	0.403	1.02e-14
Left Foot Pause Length	0.982	0.964	0.0902	0.00232	2.88	0.00127
Right Foot Pause Length	2.96	0.414	0.224	0.00251	1.45	0.000518
Left foot relative travel	6.39e-06	2.01	3.17e-06	0.00522	1.91	0.000206
Right foot relative travel	2.3e-05	2.61	1.78e-06	0.395	1.23	0.0247
Foot Crossovers	0.000409	0.177	1.04e-05	2.12	0.0111	0.0365

Table 1: Bonferroni corrected p-values of t-tests comparing computed features against different groups of patients.

	Away	Towards
Left Wrist Movements	0.00356	0.0238
Right Wrist Movements	0.00826	0.105
Average Right Hip Angle	0.115	0.00636
Average Left Hip Angle	0.000166	0.19
Average Knee Gait	1.05e-18	2.39e-23
Left Foot Pause Length	0.114	0.000365
Right Foot Pause Length	0.162	2.31e-06
Left foot relative travel	2.59e-08	0.000307
Right foot relative travel	8.22e-08	0.0504
Foot Crossovers	0.00011	0.00507

Table 2: One way ANOVA test p-values for relevance of features between Away/Towards groups, for all three patient categories.

	Away	Towards
Left Wrist Movements	0.000112	2.31e-08
Right Wrist Movements	2.33e-05	9.03e-10
Average Right Hip Angle	0.0126	0.00769
Average Left Hip Angle	1.78e-05	0.00514
Average Knee Gait	1.72e-15	4.83e-18
Left Foot Pause Length	0.0608	6.79e-05
Right Foot Pause Length	0.0385	4.62e-07
Left foot relative travel	3.2e-07	0.00017
Right foot relative travel	4.44e-07	0.0129
Foot Crossovers	4.98e-05	0.0384

Table 3: Kruskal test for relevance of features between Away/Towards groups, for all three patient categories.

distinguisher of Parkinson’s disease from other groups highlights how patients with PD raise their feet higher than participants from other groups

In the videos where patients walked toward a camera, we observed higher mean values for the pause lengths for the left and right feet for the participants in the Parkinson’s group. These detected longer pauses align with the symptoms of Parkinson’s disease

and can be attributed as another marker for distinguishing it from Ataxia (an average step by a left foot was longer by 0.0634 seconds; right foot was longer by 0.1067 seconds) or a Control group (an average step by a left foot was longer by 0.0643 seconds; right foot was longer by 0.1180 seconds).

Our evaluation of how foot crossovers affected the distribution of our dataset showed how the number of participants from each movement disorder category (PD/Ataxia/Control) who had one foot cross in front of the other at least once during their diagnostic task was unequal. If we consider passing one foot in front of the other as a means to catch oneself when losing balance, it is understandable that we would have more occurrences with PD which notes impaired balance as a symptom. Another interesting finding is that the number of occurrences where a patient with Ataxia has at least no foot crossovers during their walk is higher than expected. We can generalize this result as patients who have Ataxia have a better ability to maintain balance when doing a diagnostic walking task.

6.2 Diagnostician Survey

We interviewed two doctors who diagnose patients with Movement disorders, specifically one who specializes in Ataxia and another who specializes in Parkinson’s disease, on the usefulness of our interface. Due to COVID restrictions, the interviews took place over video calls (Zoom), but we gave the interviewees screen share control to interact with the web interface. Our interviews were structured in two parts: first, we presented our tool, toured the interface, and explained how to navigate the tool to get information about the patient’s diagnostic task; Next, we reviewed the motivation for using an automated tool such as ours, asked about the usefulness of our tool in their line of work, and discussed what improvements they believed should be included in this tool and other automated systems in their discipline. After reviewing the audio-recorded feedback, we identified two main themes:

First, we consistently heard back that doctors generally do not want another system to learn or an additional device to add to their already large set of tools to learn. What would be more useful is a “check engine light” platform with no buttons to press, like ours, to make a doctor aware of high-risk patients before doing complex procedures. For example, with orthopedic surgery, there is use in showing this patient-derived data if it can highlight gait statistics

that are related to the operation such as “Five times more likely to have a fall risk. Should not do the operation”. In its current form, video perspectives such as “Ankle View” and the Openpose skeleton give “a helpful initial insight” to that. A key statement was “I’m less interested in the individual biomarkers and more interested in the interpretation; This is a patient who falls or is more likely to fall”.

Second, a key point of feedback we heard was the ease of use of the tool and its applicability for more movement-based disciplines. Being able to view the same patient with overlaid information would be “very very useful” when looking at movement disorders that have overlapping symptoms. One doctor noted they appreciated “how easy it is to get to the highlights so I don’t have to spend time reading all of the stats to see the stand-out stuff”.

6.3 Limitations

An important consideration of this work is the accuracy of the OpenPose tool when detecting human joints and measuring their associated movements. We sought to use a keypoint detection framework that was both easily available, to ensure reproducibility, and widely adopted across the academic community. We also chose to use CMU Openpose [6] as it was more consistent in accuracy [24] as opposed to alternatives such as BlazePose (Google’s MediaPipe) and PoseNet.

An additional limitation would be the perspective of the videos in our dataset. We only had access to the frontal plane of patient movement, as opposed to the sagittal plane (side view), which can provide a clearer view of the patient’s stride and gait, and therefore more information to distinguish the movement disorders that the patients may have been diagnosed with. This did not however prevent us from being able to detect the gait cycle or movements associated with it. In an extension of this work, we plan to validate the findings as outlined in previous sections from a sagittal plane and use a larger group of participants to see how our results are reflected on a larger population.

6.4 Future Work

In a future study, we see a strong benefit in exploring how this work can be applied to different disciplines. Based on the feedback from the doctors we surveyed, looking into providing natural language interpretations and highlights on the full gait cycle would be useful in multiple disciplines for identifying high-risk patients. Being able to highlight and explain the walking task as it happened could let doctors make key decisions before fully performing surgery or engaging in another invasive task.

Similarly, we would be interested in extending this quantification of movement disorders to a different plane of view and validating the results against a series of wireless sensors or motion capture tools.

7 CONCLUSION

In this work, we found metrics that can be used to distinguish patients with a movement disorder or from a control group and designed an intelligent assistive interface to show this information to doctors who might be diagnosing a patient. After building this tool we showed it to doctors and got feedback on its ease of use and

potential place in a doctor’s office. We see this as a foundational step to building interfaces that can alert doctors to abnormalities in gait and general movement. With additional information about the diagnostic task being available, we see the opportunity to add more transparency to the diagnostic process and allow patients the opportunity about what parts of their diagnostic task were most relevant. From the feedback we have received and the ability to distinguish movement disorders beyond eyesight alone, it is clear that this is the first step to assistive interfaces in the future.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Award IIS1750380, the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award number P50NS108676, the Gordon and Betty Moore Foundation, and the Google Faculty Award.

APPENDIX A

The following READISCA contributors are multiple principal investigators of NIH U01NS104326 (READISCA): Henry Paulson, MD, PhD (Neurology, University of Michigan) Gülin Öz, PhD (Radiology and Center for Magnetic Resonance Research, University of Minnesota) Thomas Klockgether, MD Alexandra Durr, MD, PhD Tetsuo Ashizawa, MD

READISCA Consortium Collaborators: Sheng-Han Kuo, MD (Neurology, Columbia University), George Wilmot, MD, PhD (Neurology, Emory University), Laura M. Scorr, MD (Neurology, Emory University), Liana Rosenthal, MD, PhD, (Neurology, Johns Hopkins University), Chiadi Onyike, MD, (Psychiatry and Behavioral Science, Johns Hopkins University), and Haris I. Sair, MD (Radiology, Johns Hopkins University), Jeremy Schmahmann, MD, PhD (Neurology, Harvard University), Puneet Opal, MD, PhD (Neurology, Northwestern University), Yasushi Kisanuki, MD (Neurology, Ohio State University), Sandra Kostyk, MD, PhD (Neurology, Ohio State University) Sharon Sha, MD, Jacinda Sampson, MD, PhD, Veronica Santini, MD (Neurology & Neurological Science, Stanford University) Talene Yacoubian, MD, PhD, Marissa N. Dean, MD (Neurology, University of Alabama Birmingham) Michael Geschwind, MD, PhD, Alexandra Nelson, MD, PhD, Cameron Dieticker, MD (Neurology, University of California San Francisco) Christopher Gomez, MD, PhD (Neurology, University of Chicago), Trevor Hawkins, MD (Neurology, University of Colorado), Drew Kern, MD (Neurology, University of Colorado), Matthew Burns, MD, PhD (Neurology, University of Florida), Peggy Nopoulos, MD (Psychiatry, University of Iowa), Annie Killoran, MD (Neurology, University of Iowa), Khalaf Bushara, MD (University of Minnesota), Peter Morrison, DO (Neurology, University of Rochester), Vikram Shakkottai, MD, PhD (University of Texas Southwestern), Ilaria Giordano (Neurology, DZNE), Marcus Grobe-Einsler (Neurology, DZNE)

Email addresses of the listed investigators are available at <https://readisca.org/readisca-team/>.

REFERENCES

- [1] Kenta Abe, Ken-Ichi Tabei, Keita Matsuura, Kazuyuki Kobayashi, and Tomoyuki Ohkubo. 2021. OpenPose-based Gait Analysis System For Parkinson’s Disease Patients From Arm Swing Data. In *2021 International Conference on Advanced*

- Mechatronic Systems (ICAMechS)*, 61–65. <https://doi.org/10.1109/ICAMechS54019.2021.9661562>
- [2] Mohammad Rafayet Ali, Javier Hernandez, E Ray Dorsey, Ehsan Hoque, and Daniel McDuff. 2020. Spatio-temporal attention and magnification for classification of Parkinson's disease from videos collected via the internet. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 207–214.
 - [3] Mohammad Rafayet Ali, Taylan Sen, Qianyi Li, Raina Langevin, Taylor Myers, E Ray Dorsey, Saloni Sharma, and Ehsan Hoque. 2021. Analyzing head pose in remotely collected videos of people with Parkinson's disease. *ACM Transactions on Computing for Healthcare* 2, 4 (2021), 1–13.
 - [4] Tetsuo Ashizawa and Guangbin Xia. 2016. Ataxia. *Continuum: Lifelong Learning in Neurology* 22, 4 Movement Disorders (2016), 1208.
 - [5] Bastiaan R Bloem, Michael S Okun, and Christine Klein. 2021. Parkinson's disease. *The Lancet* 397, 10291 (2021), 2284–2303. [https://doi.org/10.1016/S0140-6736\(21\)00218-X](https://doi.org/10.1016/S0140-6736(21)00218-X)
 - [6] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. 2019. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019).
 - [7] Claudia Ferraris, Gianluca Amprimo, Giulia Masi, Luca Vismara, Riccardo Cremascoli, Serena Sinagra, Giuseppe Pettiti, Alessandro Mauro, and Lorenzo Priano. 2022. Evaluation of Arm Swing Features and Asymmetry during Gait in Parkinson's Disease Using the Azure Kinect Sensor. *Sensors* 22, 16 (2022). <https://doi.org/10.3390/s22166282>
 - [8] Parkinson's Foundation. 2023. What is Parkinson's? <https://www.parkinson.org/understanding-parkinsons/what-is-parkinsons>
 - [9] Rocío García-Retamero and Ulrich Hoffrage. 2013. Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine* 83 (2013), 27–33.
 - [10] Christopher G Goetz, Barbara C Tilley, Stephanie R Shaftman, Glenn T Stebbins, Stanley Fahn, Pablo Martinez-Martin, Werner Poewe, Cristina Sampaio, Matthew B Stern, Richard Dodel, et al. 2008. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders: official journal of the Movement Disorder Society* 23, 15 (2008), 2129–2170.
 - [11] Winfried Ilg and Dagmar Timmann. 2013. Gait ataxia—specific cerebellar influences and their rehabilitation. *Movement disorders* 28, 11 (2013), 1566–1575.
 - [12] Md Saiful Islam, Wasifur Rahman, Abdelrahman Abdelkader, Phillip T. Yang Sangwu Lee, Jennifer Lynn Purks, Jamie Lynn Adams, Ruth B. Schneider, and Earl Ray Dorsey & Ehsan Hoque. 2023. Using AI to measure Parkinson's disease severity at home - npj Digital Medicine - nature.com.
 - [13] Samay Jain, Steven E Lo, and Elan D Louis. 2006. Common misdiagnosis of a common neurological disorder: how are we misdiagnosing essential tremor? *Archives of neurology* 63, 8 (2006), 1100–1104.
 - [14] Joseph Jankovic. 2008. Parkinson's disease: clinical features and diagnosis. *Journal of neurology, neurosurgery & psychiatry* 79, 4 (2008), 368–376.
 - [15] Raina Langevin, Mohammad Rafayet Ali, Taylan Sen, Christopher Snyder, Taylor Myers, E. Ray Dorsey, and Mohammed Ehsan Hoque. 2019. The PARK Framework for Automated Analysis of Parkinson's Disease Characteristics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 2, Article 54 (jun 2019), 22 pages. <https://doi.org/10.1145/3328925>
 - [16] Robert LeMoyné, Frederic Heerincx, Tanya Aranca, Robert De Jager, Theresa Zesiewicz, and Harry J Saal. 2016. Wearable body and wireless inertial sensors for machine learning classification of gait for people with Friedreich's ataxia. In *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 147–151.
 - [17] Mark Lew. 2007. Overview of Parkinson's Disease. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 27, 12P2 (2007), 155S–160S. <https://doi.org/10.1592/phco.27.12part2.155S> arXiv:<https://accpjournals.onlinelibrary.wiley.com/doi/pdf/10.1592/phco.27.12part2.155S>
 - [18] Florian Lipsmeier, Kirsten I. Taylor, Timothy Kilchenmann, Detlef Wolf, Alf Scotland, Jens Schjodt-Eriksen, Wei-Yi Cheng, Ignacio Fernandez-Garcia, Juliane Siebourg-Polster, Liping Jin, Jay Soto, Lynne Verselis, Frank Boess, Martin Koller, Michael Grundman, Andreas U. Monsch, Ronald B. Postuma, Anirvan Ghosh, Thomas Kremer, Christian Czech, Christian Gossens, and Michael Lindemann. 2018. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Movement Disorders* 33, 8 (2018), 1287–1297. <https://doi.org/10.1002/mds.27376> arXiv:<https://movementdisorders.onlinelibrary.wiley.com/doi/pdf/10.1002/mds.27376>
 - [19] Yingcheng Liu, Guo Zhang, Christopher G Tarolli, Rumen Hristov, Stella Jensen-Roberts, Emma M Waddell, Taylor L Myers, Meghan E Pawlik, Julia M Soto, Renee M Wilson, et al. 2022. Monitoring gait at home with radio waves in Parkinson's disease: A marker of severity, progression, and medication response. *Science Translational Medicine* 14, 663 (2022), ead9669.
 - [20] Mandy Lu, Qingyu Zhao, Kathleen L. Poston, Edith V. Sullivan, Adolf Pfefferbaum, Marian Shahid, Maya Katz, Leila Montaser Kouhsari, Kevin Schulman, Arnold Milstein, Juan Carlos Niebles, Victor W. Henderson, Li Fei-Fei, Kilian M. Pohl, and Ehsan Adeli. 2021. Quantifying Parkinson's disease motor severity under uncertainty using MDS-UPDRS videos. *Medical Image Analysis* 73 (2021), 102179. <https://doi.org/10.1016/j.media.2021.102179>
 - [21] Camillo Lugaesi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Yong, Juhyun Lee, Wan-Teh Chang, Wei Hua, Manfred Georg, and Matthias Grundmann. 2019. MediaPipe: A Framework for Perceiving and Processing Reality. In *Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR) 2019*. https://mixedreality.cs.cornell.edu/s/NewTitle_May1_MediaPipe_CVPR_CV4ARVR_Workshop_2019.pdf
 - [22] Walter Maetzler, Josefa Domingos, Karin Srujies, Joaquim J Ferreira, and Bastiaan R Bloem. 2013. Quantitative wearable sensors for objective assessment of Parkinson's disease. *Movement Disorders* 28, 12 (2013), 1628–1637.
 - [23] Johns Hopkins Medicine. 2023. Ataxia | John Hopkins Medicine. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/ataxia>
 - [24] Sarah Mroz, Natalie Baddour, Connor McGurk, Pascale Juneau, Albert Tu, Kevin Cheung, and Edward Lemaire. 2021. Comparing the Quality of Human Pose Estimation with BlazePose or OpenPose. 1–4. <https://doi.org/10.1109/BioSMART54244.2021.9677850>
 - [25] Megumi Ota, Hiroshige Tateuchi, Takaya Hashiguchi, and Noriaki Ichihashi. 2021. Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm. *Gait & Posture* 85 (2021), 290–297. <https://doi.org/10.1016/j.gaitpost.2021.02.006>
 - [26] Dung Phan, Nhan Nguyen, Pubudu N. Pathirana, Malcolm Horne, Laura Power, and David Szumlewiecz. 2020. A Random Forest Approach for Quantifying Gait Ataxia With Truncal and Peripheral Measurements Using Multiple Wearable Sensors. *IEEE Sensors Journal* 20, 2 (2020), 723–734. <https://doi.org/10.1109/JSEN.2019.2943879>
 - [27] Wasifur Rahman, Masum Hasan, Md Saiful Islam, Titilayo Olubajo, Jeet Thaker, Abdelrahman Abdelkader, Phillip Yang, Tetsuo Ashizawa, and Ehsan Hoque. 2022. Auto-Gait: Automatic Ataxia Risk Assessment with Computer Vision on Gait Task Videos. *arXiv preprint arXiv:2203.08215* (2022).
 - [28] Wasifur Rahman, Masum Hasan, Md Saiful Islam, Titilayo Olubajo, Jeet Thaker, Abdel-Rahman Abdelkader, Phillip Yang, Henry Paulson, Guln Oz, Alexandra Durr, Thomas Klockgether, Tetsuo Ashizawa, Readisca Investigators, and Ehsan Hoque. 2023. Auto-Gait: Automatic Ataxia Risk Assessment with Computer Vision from Gait Task Videos. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 1, Article 26 (mar 2023), 19 pages. <https://doi.org/10.1145/3580845>
 - [29] Wasifur Rahman, Sangwu Lee, Md Saiful Islam, Victor Nikhil Antony, Harshil Ratnu, Mohammad Rafayet Ali, Abdullah Al Mamun, Ellen Wagner, Stella Jensen-Roberts, Emma Waddell, et al. 2021. Detecting parkinson disease using a web-based speech task: Observational study. *Journal of Medical Internet Research* 23, 10 (2021), e26305.
 - [30] Jeremy D Schmammann, Raquel Gardner, Jason MacMore, and Mark G Vangel. 2009. Development of a brief ataxia rating scale (BARS) based on a modified form of the ICARS. *Movement disorders* 24, 12 (2009), 1820–1828.
 - [31] Tanja Schmitz-Hübsch, S Tezenas Du Montcel, L Baliko, J Berciano, S Boesch, Chantal Depondt, P Giunti, C Globas, J Infante, J-S Kang, et al. 2006. Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology* 66, 11 (2006), 1717–1720.
 - [32] A Schrag, Y Ben-Shlomo, and N Quinn. 2002. How valid is the clinical diagnosis of Parkinson's disease in the community? *Journal of Neurology, Neurosurgery & Psychiatry* 73, 5 (2002), 529–534. <https://doi.org/10.1136/jnnp.73.5.529> arXiv:<https://jnnp.bmj.com/content/73/5/529.full.pdf>
 - [33] Sub H Subramony. 2007. SARA—a new clinical scale for the assessment and rating of ataxia. *Nature clinical practice Neurology* 3, 3 (2007), 136–137.
 - [34] Lorenzo Verriello, Giada Pauletto, Annacarmen Nilo, Incoronata Lonigro, Elena Betto, Mariarosaria Valente, Francesco Curcio, and Gian Luigi Gigli. 2021. Epilepsy and episodic ataxia type 2: family study and review of the literature. *Journal of Neurology* 268, 11 (2021), 4296–4302.
 - [35] Emma M Waddell, Karthik Dinesh, Kelsey L Spear, Molly J Elson, Ellen Wagner, Michael J Curtis, David J Mitten, Christopher G Tarolli, Gaurav Sharma, E Dorsey, et al. 2021. GEORGE®: A Pilot Study of a Smartphone Application for Huntington's Disease. *Journal of Huntington's Disease* 10, 2 (2021), 293–301.