



The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia?

DOI:

[10.1016/j.jns.2020.117003](https://doi.org/10.1016/j.jns.2020.117003)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Williams, S., Zhao, Z., Hafeez, A., Wong, D., Relton, S. D., Fang, H., & Alty, J. E. (2020). The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia? *Journal of Neurological Sciences.*, 416. <https://doi.org/10.1016/j.jns.2020.117003>

Published in:

Journal of Neurological Sciences.

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



Title Page

The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia ?

Authors: Stefan Williams, MBBS, BSc; Zhibin Zhao, BEng; Awais Hafeez, MS, BSc; David C Wong, PhD; Samuel D Relton, PhD; Hui Fang, PhD; Jane E Alty MD, MA, MBBChir.

Stefan Williams, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Zhibin Zhao, (1) University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK; (2) Xi'an Jiatong University, School of Mechanical Engineering, Xi'an, China.

Awais Hafeez, (1) University of Leeds, School of Mechanical Engineering, Leeds, UK; (2) University of Engineering and Technology Lahore, Department of Mechatronics and Control Engineering, Lahore, Pakistan

David C Wong, University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK

Samuel D Relton, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Hui Fang, Loughborough University, Department of Computer Science, Loughborough, UK

Jane E Alty, (1) University of Tasmania, Wicking Dementia Research & Education Centre, Hobart, Australia (2) Leeds Teaching Hospitals NHS Trust UK.

Word count: **2758**

Number of Tables: **3**

Number of Figures: **3**

Number of Videos: **1**

Corresponding Author:

Stefan Williams

Leeds Institute of Health Science, 10th Floor, Worsley Building

University of Leeds, Clarendon Way, Leeds, UK LS2 9LU

Phone: 0113 343 0839

umswi@leeds.ac.uk

ORCID: 0000-0002-4786-1356

Keywords:

Parkinson's disease; Parkinsonism; Bradykinesia; Finger tapping; Computer Vision; Artificial Intelligence

ABSTRACT

Objective

The worldwide prevalence of Parkinson's disease is increasing. There is urgent need for new tools to objectively measure the condition. Existing methods to record the cardinal motor feature of the condition, bradykinesia, using wearable sensors or smartphone apps have not reached large-scale, routine use. We evaluate new computer vision (artificial intelligence) technology, DeepLabCut, as a contactless method to quantify measures related to Parkinson's bradykinesia from smartphone videos of finger tapping.

Methods

Standard smartphone video recordings of 133 hands performing finger tapping (39 idiopathic Parkinson's patients and 30 controls) were tracked on a frame-by-frame basis with DeepLabCut. Objective computer measures of tapping speed, amplitude and rhythm were correlated with clinical ratings made by 22 movement disorder neurologists using the Modified Bradykinesia Rating Scale (MBRS) and Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS).

Results

DeepLabCut reliably tracked and measured finger tapping in standard smartphone video. Computer measures correlated well with clinical ratings of bradykinesia (Spearman coefficients): -0.74 speed, 0.66 amplitude, -0.65 rhythm for MBRS; -0.56 speed, 0.61 amplitude, -0.50 rhythm for MDS-UPDRS; 0.69 combined for MDS-UPDRS. All $p < 0.001$.

Conclusion

New computer vision software, DeepLabCut, can quantify three measures related to Parkinson's bradykinesia from smartphone videos of finger tapping. Objective 'contactless' measures of standard clinical examinations were not previously possible with wearable sensors (accelerometers, gyroscopes, infrared markers). DeepLabCut requires only conventional video recording of clinical examination and is entirely 'contactless'. This next generation technology holds potential for Parkinson's and other neurological disorders with altered movements.

INTRODUCTION

The worldwide prevalence of Parkinson's is increasing, with an estimated 10 million people already affected [1]. There is urgent need to find new tools to objectively measure the condition, to assist diagnosis, monitoring, and research metrics [2]. This is especially pressing in the era of precision medicine and with a global shortage of neurologists [3].

The cornerstone of Parkinson's assessment is the clinician's visual judgement of bradykinesia: one of the cardinal motor features of the condition, which is defined as slowness of movement and decrement in size or speed, or progressive hesitations/halts, as movements are continued [4]. There are several clinical examination methods to assess bradykinesia, but in a brief clinic room assessment, one of the most common is to observe the patient repeatedly tapping their index finger against thumb 'as quickly and as big as possible' [5]. However, clinicians' visual interpretation is inherently subjective [6,7], with considerable inter-rater variability [5,9-11]. Previous reported methods to measure bradykinesia on finger tapping have required wearable equipment such as gyroscopes [9], electromagnetic sensors [12; 13], and infrared camera markers [14], or patient interaction with an app-based task [15], and this has hindered the pathway to large-scale, routine clinical use.

Recent developments in computer interpretation of video (computer vision), clear the pathway to 'contactless' methods of quantifying clinical assessments [16], with potential widespread global application because of the availability of inexpensive smartphone cameras [17]. DeepLabCut is recently published open source computer vision software, without a requirement to write new computer code, that has been demonstrated to track and measure mice and fruit fly body parts [18,19]. We apply it to clinical neurology for the first time and assess whether the technology can be used to track and measure bradykinesia in standard smartphone video of conventional finger tapping examination.

METHODS

Ethical review

The study was approved by the North of Scotland Research Ethics Committee, United Kingdom Health Research Authority (IRAS project ID 256116).

Participants and video recording

Video recordings of 138 hands (including left and right) were obtained from 39 patients with idiopathic Parkinson's disease and 30 healthy controls, who gave written consent. One video was rejected because the hand moved outside the video frame, making 137 videos: 77 Parkinson's hands and 60 control hands.

All Parkinson's patients had previously been diagnosed by a movement disorder specialist neurologist at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society clinical diagnostic criteria [4], and were subjectively and objectively in the 'on' state at the time of video recording (no medication was withheld prior to recording). We did not exclude patients with postural hand tremor or dystonia. Controls were recruited from the companions of patients, or hospital/university staff, and had no history of Parkinson's disease or other neurological diagnosis.

Participants rested their elbow on a chair arm with the forearm lifted to a 45 degree angle. The hand was free to move as per the protocol of the Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 Finger Tapping [5]. Only the hand/forearm was within the video frame. The distance from camera to hand was approximately 1m, but not tightly defined. Digits 1 and 2 were closest to the camera. No specific instructions were given

1
2 for the position of digits 3 to 5, and participants were free to position these digits as they preferred,
3
4 although the researcher gave a brief demonstration in which digits 3 to 5 were extended.
5

6 A smartphone, placed on a tripod, was used to record standard video (60 frames per second,
7
8 1920x1080 pixel resolution), with only ambient lighting. Participants were instructed to tap their
9
10 index finger and thumb together “as quickly and as big as possible” for at least 10 seconds [5]. Each
11
12 video was edited to an 11 second clip: 1 second prior to tapping onset and 10 seconds of finger
13
14 tapping.
15
16

17 18 19 20 Clinical rating

21
22
23
24 22 consultant neurologists specialising in movement disorders (United Kingdom) were asked to rate
25
26 30 tapping videos each, selected at random from the set of 137 videos. The raters were blinded to
27
28 patient / healthy control status and each other’s scores. Each video was rated according to both the
29
30 MDS-UPDRS Item 3.4 Finger Tapping [5], and the Modified Bradykinesia Rating Scale (MBRS)
31
32 [8,9]. The MDS-UPDRS requires the rater to amalgamate judgments of finger tapping speed,
33
34 amplitude, and rhythm into a single composite score (**Table 1**). In contrast, the MBRS is comprised of
35
36 three separate scores for speed, amplitude, and rhythm (**Table 2**). The modal clinician rating for each
37
38 video was used for correlation with computer measures (see below). For each possible pair of raters,
39
40 Cohen Kappa was calculated based on all videos they had both seen. The resulting values had mean
41
42 0.28 (fair agreement) and standard deviation 0.28.
43
44
45
46
47
48

49 DeepLabCut video tracking

50
51
52
53 DeepLabCut tracks the geometrical configuration of multiple body parts in video, without a
54
55 requirement to wear markers [18,20]. It uses transfer learning, with feature detectors based on deep
56
57 neural networks that have been pretrained on ImageNet, a massive object recognition dataset [21].
58
59
60
61
62
63
64
65

1 This means that DeepLabCut is able to accurately recognise and track body parts with minimal
2 training data [18].
3
4
5

6 The set of 137 finger tapping videos was processed by DeepLabCut. We localised (labelled) six
7 distinct points on the tapping hand in 20 frames selected by k-means clustering from each 660 frame
8 video: thumb tip; index finger tip; thumb metacarpophalangeal (MCP) joint; index finger MCP joint;
9 middle finger tip; dorsal wrist/proximal dorsal hand. The deep neural network architecture of
10 DeepLabCut was then trained using these points to predict their localisation in the remaining 97% of
11 (unlabelled) video frames (1030000 training iterations, ResNet 50). This created video pixel
12 coordinates for finger tip and thumb tip throughout each video.
13
14
15
16
17
18
19
20
21
22
23

24 The accuracy of DeepLabCut to track the hand localisation points (including finger tip and thumb tip)
25 was assessed using the ‘evaluate network’ function within DeepLabCut. This function computes the
26 Euclidean error between the manual labels and the ones predicted by DeepLabCut averaged over the
27 hand locations and test images (mean absolute error, proportional to the average root mean square
28 error) [20].
29
30
31
32
33
34
35
36

37 Signal processing

38
39
40
41

42 Video pixel coordinates for the labels produced by DeepLabCut were used to calculate the pixel
43 distance between index finger tip and thumb tip. A Savitzky-Golay filter was applied to the resulting
44 time series, which removed large, sudden transient label ‘jumps’ caused by DeepLabCut mislabelling
45 (i.e. large-distance, physiologically impossible label movement across one pair of video frames). The
46 pixel number distances were standardised across all videos, by using the maximum opening distance
47 between finger and thumb tip and normalising this to a value of 1 (all values were divided by the
48 maximal value in the corresponding time series).
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Three features of the resulting finger tip to thumb tip distance time series were calculated to reflect the clinical features of finger tapping speed, amplitude and rhythm (**Figure 1**). A measure of speed was calculated as the mean rate of change of the normalised distance between finger and thumb tip over time. A measure of amplitude variability was calculated by dividing each time series into one-second windows, with maximal overlap, and then calculating the coefficient of variation of the mean difference between the maximum and minimum amplitudes in each window. A measure of rhythm regularity was calculated by undertaking Fast Fourier Transform to find the distribution of frequencies within each finger tap time series, and then measuring the power of the dominant frequency peak added to the power of the frequencies 0.2 Hz either side of it. A more regular tapping rhythm concentrates power in a narrow frequency band, increasing the power of the dominant frequency peak (and its immediate neighbours), whereas a more irregular tapping rhythm leads to a more widely spread distribution of frequency bands, reducing the power of the dominant frequency peak (and its immediate neighbours).

Spearman correlation coefficients were calculated for the each of the three DeepLabCut measures vs the modal MBRS and MDS-UPDRS clinician ratings. In addition, we calculated the Spearman correlation coefficient of the three computer scores combined (the normalised arithmetic mean of the computer speed score, amplitude variability score and rhythm regularity score) with the MDS-UPDRS rating.

RESULTS

The participant and hand video characteristics are given in **Table 3**.

Tracking accuracy

DeepLabCut reliably tracked the finger tapping movements across 133 videoed clinical examinations.

Figure 2 shows example frames from a video labelled by DeepLabCut. The mean absolute error of DeepLabCut labelling was 8.39 pixels, i.e. the average distance between manual (human) labels and those predicted by DeepLabCut was 8.39 pixels within a 1920x1080 pixel video frame. **Video 1** shows examples of four videos labelled by DeepLabCut, including particularly challenging cases with tremor and dystonia also present.

Correlation with clinical bradykinesia ratings

The quantitative scores derived from DeepLabCut measurements correlated well with the group neurologists' modal ratings of bradykinesia; **Figure 3**. The computer vision measures of tapping speed, amplitude variability and rhythm regularity had good correlations with the respective MBRS clinical ratings for speed (-0.70 , $p < 0.001$), amplitude (0.65 , $p < 0.001$) and rhythm (-0.61 , $p < 0.001$). There were also good correlations between the computer measures of tapping speed (-0.66 , $p < 0.001$),

1 amplitude variability (0.56, $p < 0.001$) and rhythm regularity (-0.50, $p < 0.001$) and the MDS-UPDRS
2 clinical rating. The MDS-UPDRS is a composite clinical rating, combining the separate components
3 of bradykinesia, and the three computer measures combined also correlated with MDS-UPDRS (-0.69,
4 $p < 0.001$).
5
6
7
8
9

10 11 12 13 14 15 16 17 **DISCUSSION**

18
19
20
21
22 We have shown that new computer vision software, DeepLabCut, can track finger tapping in a
23 standard smartphone video recording of a conventional clinical examination, and quantify three
24 measures related to Parkinson's bradykinesia: speed, amplitude variation and rhythm regularity. We
25 found good correlations between the objective computer measures of finger tapping and clinical
26 ratings of bradykinesia by 22 movement disorder neurologists using two gold standard clinical scales
27 (MBRS and MDS-UPDRS). Thus, smartphone video measures using DeepLabCut exhibit convergent
28 validity with conventional clinician ratings.
29
30
31
32
33
34
35
36
37
38
39

40 These results are new and clinically significant. This is the first study to apply new open source
41 DeepLabCut technology, developed for animal behaviour studies, to the field of human movement
42 disorders, or indeed any aspect of clinical neurology. We demonstrate new objective measures related
43 to bradykinesia by bringing together two ubiquitous ingredients: the smartphone video camera and
44 visual artificial intelligence software. With the worldwide ownership of smartphones so high (e.g. 82%
45 in the UK, 77% in the US) [17], there are no hardware barriers to further use of this approach.
46
47
48
49
50
51
52
53
54

55 Neurologists recognize that quantification of bradykinesia will improve Parkinson's management
56 through 'precision medicine' principles: it has the potential to aid early accurate diagnosis and remote
57 monitoring, and provide fine grain regular measurement of clinical trial outcomes. Previous methods
58
59
60
61
62
63
64
65

1 to quantify bradykinesia have focussed on wearable sensors measuring finger tapping [9,12-14,23] but
2 it has been challenging to incorporate these into widespread use because equipment needs to be
3 physically attached to the patient. Furthermore, the necessity for sensors to stay fixed on a specific
4 body part has often required the exclusion of tremulous and dyskinetic patients [8,9,12,14,23]. In
5 contrast, the DeepLabCut test presented here is a contactless method that simply requires a short 10
6 second video recording of the standard clinical examination. This means it can potentially be used in
7 any setting and any patient.
8
9

10
11
12
13
14
15
16
17 The 'real life' versatility of DeepLabCut in ordinary clinical settings was evaluated in this study
18 because no special efforts were made to optimize participant positioning or filming conditions. The
19 tracking and measures were robust despite variations in smartphone camera distance, positioning of
20 the hand, skin color, flexion / extension postures of digits 3 to 5, ambient lighting conditions, and
21 inclusion of a small number of patients with tremor or dystonia during tapping. **Video 1** illustrates this,
22 with accurate labelling of finger and thumb tip despite superimposed tremor or dystonia. This is
23 particularly encouraging for the relevance of DeepLabCut aiding assessment of Parkinson's in
24 conventional clinical settings.
25
26
27
28
29
30
31
32
33
34
35
36
37

38 Two further strengths of our study are the collection of clinical ratings from a large group of
39 movement disorder neurologists, and the use of two rating scales: the MBRS and MDS-UPDRS, such
40 that the clinical ratings were representative and reliable. Clinical rating scales are vulnerable to inter-
41 rater variability but using modal values derived from a large number of blinded clinical raters, and
42 two different gold-standard rating scales, provided a more robust 'ground truth' of clinical
43 bradykinesia to evaluate the new test against. As such we were able to demonstrate that clinician and
44 DeepLabCut measures correlate for both individual components of bradykinesia (MBRS) and an
45 overall combination of those components (MDS-UPDRS).
46
47
48
49
50
51
52
53
54
55
56
57

58 Only three previous studies have objectively measured finger tapping using contactless, standard
59 video analysis. One method extracted finger tapping information from 13 Parkinson's patients with
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

advanced disease [24], but it required inclusion of the patient's face in the video, limiting the practical utility. The second report, from our own group, described a video method to measure movement using an optical flow field [25] that interprets movements of the whole hand. The work here improves on this by explicitly extracting the salient thumb and finger tip points, and validating in a larger cohort. A third study of 60 Parkinson's patients (no controls) found strong correlation between MDS-UPRDS finger tap rating and video measures of tap interval (frequency) ($r=0.91$), frequency variation ($r=0.82$) and amplitude ($r=-0.94$) but not amplitude variation ($r=0.39$) [26]. The 'ground truth' of the clinical comparison was less robust, based on just two clinical raters, and there were no MBRS ratings; furthermore, some images were blurred in the 24 frame per second videos. A major advantage of our method in comparison is that DeepLabCut is open source software, that can be downloaded and used without the need to write computer code, making it available now to neurologists with only limited computing knowledge.

The present study has some limitations. Like human vision, a simple camera lacks an absolute measure of distance, and our method cannot capture 3D movement. Rotation of the thumb and finger into a horizontal plane would falsely alter amplitude measurement. However, in practice such movement is rare during tapping, and our normalized measure of finger-thumb distance based on 2D maximum opening distance appeared to capture amplitude, speed and rhythm abnormalities well, given the good correlations with clinical ratings.

We have measured amplitude variation, but not amplitude decrement (the sequence effect). Previous approaches have measured decrement as a straight line linear regression [23], but decrement would also proportionately affect a measure of amplitude variation, such that the variations in decrement should be captured within our measure. Furthermore, there are several other previous reports of good correlation between non-decrement amplitude measures using sensors and clinician ratings of finger tap bradykinesia [9].

1 Relatively few videos were rated as grade 3 or 4 bradykinesia, but this would not be expected to
2 greatly affect our findings, as computer analysis of milder bradykinesia could be considered a more
3 challenging task. There was a degree of error in DeepLabCut labelling of finger and thumb tips, but
4 this was minimal (mean of 8.39 pixels within a 1920x1080 pixel video frame) and unlikely to affect
5 individual measures. We mitigated against this by applying a filter prior to feature extraction. Full
6 flexion of digits 3 to 5 might improve labelling accuracy, but would also limit the amplitude of finger
7 tapping [23], potentially masking more subtle deficits. The labelling error provided by DeepLabCut's
8 'evaluate network' function only gives an accuracy figure for the entire set of videos but not for
9 individual videos, meaning that we have not reported accuracy in subsets of videos, such as those with
10 visible tremor.
11
12
13
14
15
16
17
18
19
20
21
22
23

24 Finally, we have demonstrated measurements that show convergent validity with clinical ratings,
25 rather than a test to discriminate patients from controls. In our view, it makes little sense to pursue
26 automated diagnostic classification using finger tapping video alone because clinical Parkinson's
27 diagnosis is by definition a broader assessment than isolated finger tapping performance [4].
28
29
30
31
32
33
34

35 This study has evaluated the clinical application of DeepLabCut in a medical condition, specifically a
36 neurological disorder. The methodology described here provides a new objective measure of
37 Parkinson's bradykinesia in conventional clinical examination. DeepLabCut provides a contactless
38 means for tracking and measuring bradykinesia in Parkinson's, without any special equipment other
39 than the ubiquitous smartphone. This research therefore provides a new tool to quantify bradykinesia.
40 It could potentially be used to support diagnosis and monitoring of Parkinson's, both clinically and as
41 an outcome measure in research studies. Beyond Parkinson's and other movement disorders, this
42 same technology could be applied to many other neurological conditions that require measurement of
43 movements or function; for example multiple sclerosis, neuropathies, epilepsy and stroke.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

REFERENCES

1. Dorsey ER, Bloem BR. The Parkinson Pandemic - A Call to Action. *JAMA Neurol.* 2018; 75(1): 9–10.
2. Delenclos M, Jones DR, McLean PJ, Uitti RJ. Biomarkers in Parkinson’s disease: Advances and strategies. *Park Relat Disord.* 2016; 22: S106-S110.
3. Burton A. How do we fix the shortage of neurologists?. *The Lancet Neurology.* 2018 Jun 1;17(6):502-3.
4. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson’s disease. *Mov Disord.* 2015;30(12):1591-1601. doi:10.1002/mds.26424
5. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* 2008; 23(15): 2129-70.
6. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology.* 2014; 83(5): 406-412.
7. Bajaj NPS, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study *Journal of Neurology, Neurosurgery & Psychiatry* 2010; 81: 1223-1228.
8. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson’s disease: Differential effects on bradykinesia. *Mov Disord.* 2007; 22(3): 328-333.
9. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson’s disease: Reliability and comparison with kinematic measures. *Mov Disord.* 2011; 26(10): 1859-1863.
10. Martinez-Martin P, Gil-Nagel A, Gracia LM, et al. Unified Parkinson’s disease rating scale characteristics and structure. *Mov Disord.* 1994; 9(1): 76-83.

- 1 11. Goetz CG, Stebbins GT. Assuring interrater reliability for the UPDRS motor section: Utility of the
2 UPDRS teaching tape. *Mov Disord.* 2004; 19(12): 1453-1456.
3
- 4 12. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and
5 amplitude of movement in Parkinson's disease: A pilot study. *Mov Disord.* 2009; 24(7): 1001-1008.
6
- 7 13. Gao C, Smith S, Lones M, et al. Objective assessment of bradykinesia in Parkinson's disease
8 using evolutionary algorithms: Clinical validation. *Transl Neurodegener.* 2018; 7(1): 1-8.
9
- 10 14. Růžička E, Krupička R, Zárubová K, Ruzs J, Jech R, Szabó Z. Tests of manual dexterity and
11 speed in Parkinson's disease: Not all measure the same. *Park Relat Disord.* 2016; 28: 118-123.
12
- 13 15. Lee CY, Kang SJ, Hong S-K, Ma H-I, Lee U, Kim YJ. A Validation Study of a Smartphone-
14 Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson's
15 Disease. *PLoS One.* 2016;11(7).
16
- 17 16. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of
18 blepharospasm severity. *Neurology.* 2016;87(20):2146-2153.
19
- 20 17. "Top Countries/Markets by Smartphone Penetration & Users." Newzoo. Available from:
21 <https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users/>. [Accessed
22 5 November 2019.]
23
- 24 18. Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: markerless pose estimation of user-
25 defined body parts with deep learning. *Nat Neurosci.* 2018; 21(9): 1281-1289.
26
- 27 19. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-
28 Dimensional Reconstruction of Paw and Digit Trajectories. *Journal of visualized experiments: JoVE.*
29 2019; Jul(149).
30
- 31 20. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D
32 markerless pose estimation across species and behaviors. *Nature protocols.* 2019; 14(7): 2152-76.
33
- 34 21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the*
35 *IEEE conference on computer vision and pattern recognition* 2016; 770-778.
36
- 37 22. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the
38 Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* 2004; 19(9): 1020-8.
39
- 40 23. Ling H, Massey LA, Lees AJ, Brown P, Day BL. Hypokinesia without decrement distinguishes
41 progressive supranuclear palsy from Parkinson's disease. *Brain.* 2012; 135(4): 1141-1153.
42
- 43 24. Khan T, Nyholm D, Westin J, Dougherty M. A computer vision framework for finger-tapping
44 evaluation in Parkinson's disease. *Artif. Intell. Med.* 2014; 60(1): 27-40.
45
- 46 25. Wong DC, Relton SD, Fang H, et al. Supervised Classification of Bradykinesia for Parkinson's
47 Disease Diagnosis from Smartphone Videos. In *IEEE 32nd International Symposium on Computer-
48 Based Medical Systems (CBMS).* 2019; 32-37.
49
- 50 26. Liu Y, Chen J, Hu C, et al. Vision-Based Method for Automatic Quantification of Parkinsonian
51 Bradykinesia. *IEEE Trans Neural Syst Rehabil Eng.* 2019;27(10):1952-1961.
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

FIGURE LEGENDS

Figure 1. Illustration of the three parameters derived from the DeepLabCut finger tip and thumb tip coordinates to measure finger tapping speed, amplitude and rhythm.

Figure 2. Example video frames taken from smartphone video labelled by DeepLabCut, showing the six localization (labelling) points: thumb tip (dark blue), thumb metacarpophalangeal (MCP) joint (light blue), index finger tip (cyan), index finger MCP joint (yellow), middle finger tip (orange), dorsal wrist (red).

Figure 3. Computer measures of finger tapping speed, amplitude variation and rhythm, derived from DeepLabCut smartphone video tracking, correlate with clinician mode ratings for the same videos. Pearson correlation coefficients shown. MBRS: Modified Bradykinesia Rating Scale. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

TABLES

Table 1. The Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 (Finger Tapping).

MDS-UPDRS Grade	Description
0: Normal	No problems
1: Slight	Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.
2: Mild	Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.
3: Moderate	Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.
4: Severe	Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb "as quickly and as big as possible". The first 10 taps are evaluated by the clinician [5].

Table 2. The Modified Bradykinesia Rating Scale (MBRS).

Score	Speed	Amplitude	Rhythm
0	Normal	Normal	Regular, no arrests or pauses in ongoing movement
1	Mild Slowing	Mild reduction in amplitude in later performance, most movements close to normal	Mild impairment, up to two brief arrests in the 10 seconds, none lasting > 1 second
2	Moderate Slowing	Moderate, reduction in amplitude visible early in performance but continues to maintain 50% amplitude through most of the tasks	Moderate, 3 to 4 arrests in 10 seconds; OR 1 or 2 lasting > 1second
3	Severe slowing	Severe, less than 50% amplitude through most of the task	Severe, 5 or more arrests/10 seconds; OR more than 2 lasting > 1 second
4	Can barely perform the task	Can barely perform the task	Can barely perform the task

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 seconds of tapping are evaluated by the clinician to give three MBRS ratings (one for each of speed, amplitude, rhythm) [6,7].

Table 3. Study participant (hand video) characteristics.

	Patients	Controls
Age (Std. Dev.) yrs	68 (9.6)	59 (19.4)
Male/Female	47/26	22/38
Median years since diagnosis	4	n/a
Median H&Y [IQR]	2 [1,3]	n/a
H&Y = 1	32	
H&Y = 1.5	2	
H&Y = 2	12	
H&Y = 2.5	4	
H&Y = 3	19	
H&Y = 4	4	
H&Y = 5	0	
Median MDS-UPDRS [IQR]	2 [1,3]	1 [0,1]
MDS-UPDRS = 0	9	23
MDS-UPDRS = 1	20	23

MDS-UPDRS = 2	19	12
MDS-UPDRS = 3	20	2
MDS-UPDRS = 4	5	0
Visible tremor in video	11	0
Visible dystonia in video	2	0

Data is split by Parkinson's hands (n=73) and control hands (n=60). H&Y: modified Hoehn and Yahr scale [22]. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale, Item 3.4 (Finger Tapping). IQR: Interquartile Range.

Acknowledgements

We would like to thank the following UK movement disorder neurologists for providing clinical rating of videos: Dr Jeremy Cosgrove, Dr Rachel Newby, Dr Stephen Butterworth, Dr Alistair Lansbury, Dr Michael Bonello, Dr Sundus Alusi, Dr Peter Moore, Dr Richard Ellis, Dr Camille Carroll, Dr Chris Kobylecki, Dr Paul Goldsmith, Dr Naomi Warren, Dr Edward Newman, Dr Philip Buttery, Dr Jonathan Evans, Dr Gillian Sare, Dr Donald Grosset, Dr Caroline Williams-Gray, Dr Monty Silverdale, Dr Louise Wiblin, Dr Adam Cassidy

We would also like to thank Dr Parisa Patel and Dr Paschal O’Gorman for data collection.

Author contributions

Name	Location	Role	Contribution
Stefan Williams, MBBS, BSc	University of Leeds, UK	Author	Design and conceptualized study; data acquisition and interpretation;

				drafted the manuscript for intellectual content
1				
2				
3				
4	Zhibin Zhao, BEng	(1) University of Manchester, UK; (2) Xi'an Jiatong University, China	Author	Analyzed the data (major role)
5				
6				
7				
8				
9	Awais Hafeez, MA, BSc	(1) University of Leeds, UK; (2) University of Engineering and Technology Lahore, Pakistan	Author	Design and conceptualized the study; major role in acquisition of data
10				
11				
12				
13				
14				
15				
16	David C Wong, PhD	University of Manchester, UK	Author	Design and conceptualized study; analyzed the data; data interpretation; revised the manuscript for intellectual content
17				
18				
19				
20				
21				
22				
23				
24				
25				
26	Samuel D Relton, PhD	University of Leeds, UK	Author	Design and conceptualized study; data interpretation; revised the manuscript for intellectual content
27				
28				
29				
30				
31				
32				
33				
34	Hui Fang, PhD	Loughborough University, UK	Author	Design and conceptualized study (major role)
35				
36				
37				
38	Jane E Alty, MD, MA, MBBChir	(1) University of Tasmania, Australia; (2) Leeds Teaching Hospitals NHS Trust, UK	Author	Design and conceptualized study; data interpretation; drafted the manuscript for intellectual content
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				

Title Page

The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia ?

Authors: Stefan Williams, MBBS, BSc; Zhibin Zhao, BEng; Awais Hafeez, MS, BSc; David C Wong, PhD; Samuel D Relton, PhD; Hui Fang, PhD; Jane E Alty MD, MA, MBBChir.

Stefan Williams, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Zhibin Zhao, (1) University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK; (2) Xi'an Jiatong University, School of Mechanical Engineering, Xi'an, China.

Awais Hafeez, (1) University of Leeds, School of Mechanical Engineering, Leeds, UK; (2) University of Engineering and Technology Lahore, Department of Mechatronics and Control Engineering, Lahore, Pakistan

David C Wong, University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK

Samuel D Relton, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Hui Fang, Loughborough University, Department of Computer Science, Loughborough, UK

Jane E Alty, (1) University of Tasmania, Wicking Dementia Research & Education Centre, Hobart, Australia (2) Leeds Teaching Hospitals NHS Trust UK.

Word count: **3101**

Number of Tables: **3**

Number of Figures: **3**

Number of Videos: **1**

Corresponding Author:

Stefan Williams

Leeds Institute of Health Science, 10th Floor, Worsley Building

University of Leeds, Clarendon Way, Leeds, UK LS2 9LU

Phone: 0113 343 0839

umswi@leeds.ac.uk

ORCID: 0000-0002-4786-1356

Keywords:

Parkinson's disease; Parkinsonism; Bradykinesia; Finger tapping; Computer Vision; Artificial Intelligence; DeepLabCut

1 **ABSTRACT**

6 **Objective**

8 The worldwide prevalence of Parkinson's disease is increasing. There is urgent need for new tools to objectively
9 measure the condition. Existing methods to record the cardinal motor feature of the condition, bradykinesia, using
10 wearable sensors or smartphone apps have not reached large-scale, routine use. We evaluate new computer vision
11 (artificial intelligence) technology, DeepLabCut, as a contactless method to quantify measures related to
12 Parkinson's bradykinesia from smartphone videos of finger tapping.
13
14
15
16
17
18
19

20 **Methods**

21 Standard smartphone video recordings of 133 hands performing finger tapping (39 idiopathic Parkinson's patients
22 and 30 controls) were tracked on a frame-by-frame basis with DeepLabCut. Objective computer measures of
23 tapping speed, amplitude and rhythm were correlated with clinical ratings made by 22 movement disorder
24 neurologists using the Modified Bradykinesia Rating Scale (MBRS) and Movement Disorder Society revision of
25 the Unified Parkinson's Disease Rating Scale (MDS-UPDRS).
26
27
28
29
30
31
32
33

34 **Results**

35 DeepLabCut reliably tracked and measured finger tapping in standard smartphone video. Computer measures
36 correlated well with clinical ratings of bradykinesia (Spearman coefficients): -0.74 speed, 0.66 amplitude, -0.65
37 rhythm for MBRS; -0.56 speed, 0.61 amplitude, -0.50 rhythm for MDS-UPDRS; 0.69 combined for MDS-
38 UPDRS. All $p < 0.001$.
39
40
41
42
43
44
45
46

47 **Conclusion**

48 New computer vision software, DeepLabCut, can quantify three measures related to Parkinson's bradykinesia
49 from smartphone videos of finger tapping. Objective 'contactless' measures of standard clinical examinations
50 were not previously possible with wearable sensors (accelerometers, gyroscopes, infrared markers). DeepLabCut
51 requires only conventional video recording of clinical examination and is entirely 'contactless'. This next
52 generation technology holds potential for Parkinson's and other neurological disorders with altered movements.
53
54
55
56
57
58
59
60
61
62
63
64
65

INTRODUCTION

1
2
3
4 The worldwide prevalence of Parkinson's is increasing, with an estimated 10 million people already
5 affected [1]. There is urgent need to find new tools to objectively measure the condition, to assist
6 diagnosis, monitoring, and research metrics [2]. This is especially pressing in the era of precision
7 medicine and with a global shortage of neurologists [3].
8
9

10
11
12
13
14
15 The cornerstone of Parkinson's assessment is the clinician's visual judgement of bradykinesia: one of
16 the cardinal motor features of the condition, which is defined as slowness of movement and decrement
17 in size or speed, or progressive hesitations/halts, as movements are continued [4]. There are several
18 clinical examination methods to assess bradykinesia, but in a brief clinic room assessment, one of the
19 most common is to observe the patient repeatedly tapping their index finger against thumb 'as quickly
20 and as big as possible' [5]. However, clinicians' visual interpretation is inherently subjective [6,7], with
21 considerable inter-rater variability [5,9-11]. Previous reported methods to measure bradykinesia on
22 finger tapping have required wearable equipment such as gyroscopes [9], electromagnetic sensors [12;
23 13], and infrared camera markers [14], or patient interaction with an app-based task [15], and this has
24 hindered the pathway to large-scale, routine clinical use.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Recent developments in computer interpretation of video (computer vision), clear the pathway to
41 'contactless' methods of quantifying clinical assessments [16], with potential widespread global
42 application because of the availability of inexpensive smartphone cameras [17]. DeepLabCut is recently
43 published open source computer vision software, without a requirement to write new computer code,
44 that has been demonstrated to track and measure mice and fruit fly body parts [18,19]. We apply it to
45 clinical neurology for the first time and assess whether the technology can be used to track and measure
46 bradykinesia in standard smartphone video of conventional finger tapping examination.
47
48
49
50
51
52
53
54
55
56
57
58
59

METHODS

1
2 Ethical review
3
4
5

6 The study was approved by the North of Scotland Research Ethics Committee, United Kingdom Health
7
8 Research Authority (IRAS project ID 256116).
9

10
11
12
13 Participants and video recording
14
15

16
17 Video recordings of 138 hands (including left and right) were obtained from 39 patients with idiopathic
18 Parkinson's disease and 30 healthy controls, who gave written consent. One video was rejected because
19
20 the hand moved outside the video frame, making 137 videos: 77 Parkinson's hands and 60 control
21
22 hands.
23
24
25

26
27
28 All Parkinson's patients had previously been diagnosed by a movement disorder specialist neurologist
29 at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society
30 clinical diagnostic criteria [4], and were subjectively and objectively in the 'on' state at the time of
31 video recording (no medication was withheld prior to recording). We did not exclude patients with
32
33 postural hand tremor or dystonia. Controls were recruited from the companions of patients, or
34
35 hospital/university staff, and had no history of Parkinson's disease or other neurological diagnosis.
36
37
38
39
40
41
42

43
44 Participants rested their elbow on a chair arm with the forearm lifted to a 45 degree angle. The hand
45 was free to move as per the protocol of the Movement Disorder Society revision of the Unified
46 Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 Finger Tapping [5]. Only the hand/forearm
47
48 was within the video frame. The distance from camera to hand was approximately 1m, but not tightly
49
50 defined. Digits 1 and 2 were closest to the camera. No specific instructions were given for the position
51
52 of digits 3 to 5, and participants were free to position these digits as they preferred, although the
53
54 researcher gave a brief demonstration in which digits 3 to 5 were extended.
55
56
57
58
59
60
61
62

1 A smartphone, placed on a tripod, was used to record standard video (60 frames per second, 1920x1080
2 pixel resolution), with only ambient lighting. Participants were instructed to tap their index finger and
3 thumb together “as quickly and as big as possible” for at least 10 seconds [5]. Each video was edited to
4 an 11 second clip: 1 second prior to tapping onset and 10 seconds of finger tapping.
5
6
7
8
9

10 Clinical rating

11
12
13
14
15 22 consultant neurologists specialising in movement disorders (United Kingdom) were asked to rate 30
16 tapping videos each, selected at random from the set of 137 videos. The median number of raters per
17 video was 5 (range 1 to 12, interquartile range 3 to 7). The raters were blinded to patient / healthy
18 control status and each other’s scores. Each video was rated according to both the MDS-UPDRS Item
19 3.4 Finger Tapping [5], and the Modified Bradykinesia Rating Scale (MBRS) [8,9]. The MDS-UPDRS
20 requires the rater to amalgamate judgments of finger tapping speed, amplitude, and rhythm into a single
21 composite score (**Table 1**). In contrast, the MBRS is comprised of three separate scores for speed,
22 amplitude, and rhythm (**Table 2**). The modal clinician rating for each video was used for correlation
23 with computer measures (see below). For each possible pair of raters, Cohen Kappa was calculated
24 based on all videos they had both seen. The resulting values had mean 0.28 (fair agreement) and
25 standard deviation 0.28.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 DeepLabCut video tracking

43
44
45
46 DeepLabCut tracks the geometrical configuration of multiple body parts in video, without a requirement
47 to wear markers [18,20]. It uses transfer learning, with feature detectors based on deep neural networks
48 that have been pretrained on ImageNet, a massive object recognition dataset [21]. This means that
49 DeepLabCut is able to accurately recognise and track body parts with minimal training data [18].
50
51
52
53
54
55
56
57

58 The set of 137 finger tapping videos was processed by DeepLabCut. We localised (labelled) six distinct
59 points on the tapping hand in 20 frames selected by k-means clustering from each 660 frame video:
60
61
62
63

1 thumb tip; index finger tip; thumb metacarpophalangeal (MCP) joint; index finger MCP joint; middle
2 finger tip; dorsal wrist/proximal dorsal hand. The deep neural network architecture of DeepLabCut was
3
4 then trained using these points to predict their localisation in the remaining 97% of (unlabelled) video
5
6 frames (1030000 training iterations, ResNet 50). This created video pixel coordinates for finger tip and
7
8 thumb tip throughout each video.
9

10
11
12
13 The accuracy of DeepLabCut to track the hand localisation points (including finger tip and thumb tip)
14
15 was assessed using the ‘evaluate network’ function within DeepLabCut. This function computes the
16
17 Euclidean error between the manual labels and the ones predicted by DeepLabCut averaged over the
18
19 hand locations and test images (mean absolute error, proportional to the average root mean square error)
20
21 [20].
22
23
24
25

26 Signal processing

27
28
29
30
31 Video pixel coordinates for the labels produced by DeepLabCut were used to calculate the pixel distance
32
33 between index finger tip and thumb tip. A Savitzky-Golay filter was applied to the resulting time series,
34
35 which removed large, sudden transient label ‘jumps’ caused by DeepLabCut mislabelling (i.e. large-
36
37 distance, physiologically impossible label movement across one pair of video frames). The pixel
38
39 number distances were standardised across all videos, by using the maximum opening distance between
40
41 finger and thumb tip and normalising this to a value of 1 (all values were divided by the maximal value
42
43 in the corresponding time series).
44
45
46
47
48

49 Three features of the resulting finger tip to thumb tip distance time series were calculated to reflect the
50
51 clinical features of finger tapping speed, amplitude and rhythm (**Figure 1**). A measure of speed was
52
53 calculated as the mean rate of change of the normalised distance between finger and thumb tip over
54
55 time. A measure of amplitude variability was calculated by dividing each time series into one-second
56
57 windows, with maximal overlap, and then calculating the coefficient of variation of the mean difference
58
59 between the maximum and minimum amplitudes in each window. A measure of rhythm regularity was
60
61
62
63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

calculated by undertaking Fast Fourier Transform to find the distribution of frequencies within each finger tap time series, and then measuring the power of the dominant frequency peak added to the power of the frequencies 0.2 Hz either side of it. A more regular tapping rhythm concentrates power in a narrow frequency band, increasing the power of the dominant frequency peak (and its immediate neighbours), whereas a more irregular tapping rhythm leads to a more widely spread distribution of frequency bands, reducing the power of the dominant frequency peak (and its immediate neighbours).

Spearman correlation coefficients were calculated for the each of the three DeepLabCut measures vs the modal MBRS and MDS-UPDRS clinician ratings. In addition, we calculated the Spearman correlation coefficient of the three computer scores combined (the normalised arithmetic mean of the computer speed score, amplitude variability score and rhythm regularity score) with the MDS-UPDRS rating.

RESULTS

The participant and hand video characteristics are given in **Table 3**.

Tracking accuracy

DeepLabCut reliably tracked the finger tapping movements across 133 videoed clinical examinations. **Figure 2** shows example frames from a video labelled by DeepLabCut. The mean absolute error of DeepLabCut labelling was 8.39 pixels, i.e. the average distance between manual (human) labels and those predicted by DeepLabCut was 8.39 pixels within a 1920x1080 pixel video frame. **Video 1** shows examples of four videos labelled by DeepLabCut, including particularly challenging cases with tremor and dystonia also present.

Correlation with clinical bradykinesia ratings

1
2 The quantitative scores derived from DeepLabCut measurements correlated well with the group
3
4 neurologists' modal ratings of bradykinesia; **Figure 3**. The computer vision measures of tapping speed,
5
6 amplitude variability and rhythm regularity had good correlations with the respective MBRS clinical
7
8 ratings for speed (-0.70, $p < 0.001$), amplitude (0.65, $p < 0.001$) and rhythm (-0.61, $p < 0.001$). There
9
10 were also good correlations between the computer measures of tapping speed (-0.66, $p < 0.001$),
11
12 amplitude variability (0.56, $p < 0.001$) and rhythm regularity (-0.50, $p < 0.001$) and the MDS-UPDRS
13
14 clinical rating. The MDS-UPDRS is a composite clinical rating, combining the separate components of
15
16 bradykinesia, and the three computer measures combined also correlated with MDS-UPDRS (-0.69, p
17
18 < 0.001).
19
20
21
22
23
24
25

26 **DISCUSSION**

27
28
29
30
31 We have shown that new computer vision software, DeepLabCut, can track finger tapping in a standard
32
33 smartphone video recording of a conventional clinical examination, and quantify three measures related
34
35 to Parkinson's bradykinesia: speed, amplitude variation and rhythm regularity. We found good
36
37 correlations between the objective computer measures of finger tapping and clinical ratings of
38
39 bradykinesia by 22 movement disorder neurologists using two gold standard clinical scales (MBRS and
40
41 MDS-UPDRS). Thus, smartphone video measures using DeepLabCut exhibit convergent validity with
42
43 conventional clinician ratings.
44
45
46
47
48

49 These results are new and clinically significant. This is the first study to apply new open source
50
51 DeepLabCut technology, developed for animal behaviour studies, to the field of human movement
52
53 disorders, or indeed any aspect of clinical neurology. We demonstrate new objective measures related
54
55 to bradykinesia by bringing together two ubiquitous ingredients: the smartphone video camera and
56
57 visual artificial intelligence software. With the worldwide ownership of smartphones so high (e.g. 82%
58
59 in the UK, 77% in the US) [17], there are no hardware barriers to further use of this approach.
60
61
62
63

1
2 Neurologists recognize that quantification of bradykinesia will improve Parkinson's management
3
4 through 'precision medicine' principles: it has the potential to aid early accurate diagnosis and remote
5
6 monitoring, and provide fine grain regular measurement of clinical trial outcomes. Previous methods to
7
8 quantify bradykinesia have focussed on wearable sensors measuring finger tapping [9,12-14,23] but it
9
10 has been challenging to incorporate these into widespread use because equipment needs to be physically
11
12 attached to the patient. Furthermore, the necessity for sensors to stay fixed on a specific body part has
13
14 often required the exclusion of tremulous and dyskinetic patients [8,9,12,14,23]. In contrast, the
15
16 DeepLabCut test presented here is a contactless method that simply requires a short 10 second video
17
18 recording of the standard clinical examination. This means it can potentially be used in any setting and
19
20 any patient.
21
22
23
24
25

26
27 The 'real life' versatility of DeepLabCut in ordinary clinical settings was evaluated in this study because
28
29 no special efforts were made to optimize participant positioning or filming conditions. The tracking and
30
31 measures were robust despite variations in smartphone camera distance, positioning of the hand, skin
32
33 color, flexion / extension postures of digits 3 to 5, ambient lighting conditions, and inclusion of a small
34
35 number of patients with tremor or dystonia during tapping. **Video 1** illustrates this, with accurate
36
37 labelling of finger and thumb tip despite superimposed tremor or dystonia. This is particularly
38
39 encouraging for the relevance of DeepLabCut aiding assessment of Parkinson's in conventional clinical
40
41 settings.
42
43
44
45

46
47 Two further strengths of our study are the collection of clinical ratings from a large group of movement
48
49 disorder neurologists, and the use of two rating scales: the MBRS and MDS-UPDRS, such that the
50
51 clinical ratings were representative and reliable. Clinical rating scales are vulnerable to inter-rater
52
53 variability but using modal values derived from a large number of blinded clinical raters, and two
54
55 different gold-standard rating scales, provided a more robust 'ground truth' of clinical bradykinesia to
56
57 evaluate the new test against. As such we were able to demonstrate that clinician and DeepLabCut
58
59 measures correlate for both individual components of bradykinesia (MBRS) and an overall combination
60
61
62
63
64
65

1 of those components (MDS-UPDRS). An interesting future application with a larger participant group
2 would be to examine the relative contribution of the three video measures to individual grades of
3
4 UPDRS bradykinesia.
5
6
7

8 Another advantage of the approach described here is that it is not limited by the physical constraints of
9 wearable or screen/keyboard tapping methods, and could be applied to detect and measure other motor
10 features of Parkinson's in a contactless manner such as tremor, gait, or posture. Computer vision is
11 general has potential for application to these signs [24,25].
12
13
14
15
16
17

18 Only three previous studies have objectively measured finger tapping using contactless, standard video
19 analysis. One method extracted finger tapping information from 13 Parkinson's patients with advanced
20 disease [26], but it required inclusion of the patient's face in the video, limiting the practical utility. The
21 second report, from our own group, described a video method to measure movement using an optical
22 flow field [27] that interprets movements of the whole hand. The work here improves on this by
23 explicitly extracting the salient thumb and finger tip points, and validating in a larger cohort. A third
24 study of 60 Parkinson's patients (no controls) found strong correlation between MDS-UPDRS finger
25 tap rating and video measures of tap interval (frequency) ($r=0.91$), frequency variation ($r=0.82$) and
26 amplitude ($r=-0.94$) but not amplitude variation ($r=0.39$) [28]. The 'ground truth' of the clinical
27 comparison was less robust, based on just two clinical raters, and there were no MBRS ratings;
28 furthermore, some images were blurred in the 24 frame per second videos. A major advantage of our
29 method in comparison is that DeepLabCut is open source software, that can be downloaded and used
30 without the need to write computer code, making it available now to neurologists with only limited
31 computing knowledge.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 A well-studied, non-camera method to measure finger tap bradykinesia is the 'Bradykinesia-Akinesia
53 Incoordination (BRAIN) Test', which involves participants tapping a standard computer keyboard for
54 30 seconds [29]. This shares advantages with our method in that it involves no special equipment, and
55 could be employed without clinician-patient contact. However, it requires patient motivation to engage
56
57
58
59
60
61
62
63

1
2 with a keyboard tapping task, rather than a simple recording of existing standard clinical examination
3 (and similar requirements for patient motivation apply to patient smartphone apps [15]). Furthermore,
4 ‘BRAIN’ correlations with MDS-UPDRS finger tapping ratings of 0.44, 0.20, 0.33, 0.03 [29], are
5 weaker than those we report (0.66, 0.56, 0.50, 0.69). One reason for this may be that a computer
6 keyboard cannot record any measure of tap amplitude, and the speed measure is limited to tap
7 frequency, so that some core aspects of bradykinesia cannot be captured by ‘BRAIN’.
8
9
10
11
12
13
14

15 The present study has some limitations. Like human vision, a simple camera lacks an absolute measure
16 of distance, and our method cannot capture 3D movement. Rotation of the thumb and finger into a
17 horizontal plane would falsely alter amplitude measurement. However, in practice such movement is
18 rare during tapping, and our normalized measure of finger-thumb distance based on 2D maximum
19 opening distance appeared to capture amplitude, speed and rhythm abnormalities well, given the good
20 correlations with clinical ratings.
21
22
23
24
25
26
27
28
29
30

31 We have measured amplitude variation, but not amplitude decrement (the sequence effect). Previous
32 approaches have measured decrement as a straight line linear regression [23], but decrement would also
33 proportionately affect a measure of amplitude variation, such that the variations in decrement should be
34 captured within our measure. Furthermore, there are several other previous reports of good correlation
35 between non-decrement amplitude measures using sensors and clinician ratings of finger tap
36 bradykinesia [9].
37
38
39
40
41
42
43
44
45
46

47 Relatively few videos were rated as grade 3 or 4 bradykinesia, and it is possible that a greater number
48 of higher grade videos might change the strength of correlations. Even with accurate tracking, all forms
49 of rating or measuring bradykinesia (including clinician) may potentially be confounded by
50 comorbidities such as joint deformities, pain, dystonia etc. However, with a larger dataset, machine
51 learning techniques are well-suited to separate such additional contributions from underlying
52 bradykinesia. There was a degree of error in DeepLabCut labelling of finger and thumb tips, but this
53 was minimal (mean of 8.39 pixels within a 1920x1080 pixel video frame) and unlikely to affect
54
55
56
57
58
59
60
61
62
63

1 individual measures. We mitigated against this by applying a filter prior to feature extraction. Full
2 flexion of digits 3 to 5 might improve labelling accuracy, but would also limit the amplitude of finger
3 tapping [23], potentially masking more subtle deficits. The labelling error provided by DeepLabCut's
4 'evaluate network' function only gives an accuracy figure for the entire set of videos but not for
5 individual videos, meaning that we have not reported accuracy in subsets of videos, such as those with
6 visible tremor. This study was exploratory, aimed at proof of concept that bradykinesia related measures
7 could be derived from smartphone video. **The protocol was thus relatively simple. The test-retest
8 reliability of this technique has not been addressed in the study but we hope to do so in future. Similarly,
9 we did not measure 'off' and 'on' medication states (sensitivity to change).**

10
11
12
13
14
15
16
17
18
19
20
21
22 Finally, we have demonstrated measurements that show convergent validity with clinical ratings, rather
23 than a test to discriminate patients from controls. In our view, it makes little sense to pursue automated
24 diagnostic classification using finger tapping video alone because clinical Parkinson's diagnosis is by
25 definition a broader assessment than isolated finger tapping performance [4].

26
27
28
29
30
31
32
33 This study has evaluated the clinical application of DeepLabCut in a medical condition, specifically a
34 neurological disorder. The methodology described here provides a new objective measure of
35 Parkinson's bradykinesia in conventional clinical examination. DeepLabCut provides a contactless
36 means for tracking and measuring bradykinesia in Parkinson's, without any special equipment other
37 than the ubiquitous smartphone. This research therefore provides a new tool to quantify bradykinesia.
38 It could potentially be used to support diagnosis and monitoring of Parkinson's, both clinically and as
39 an outcome measure in research studies. Beyond Parkinson's and other movement disorders, this same
40 technology could be applied to many other neurological conditions that require measurement of
41 movements or function; for example multiple sclerosis, neuropathies, epilepsy and stroke.

42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 **REFERENCES**

1. Dorsey ER, Bloem BR. The Parkinson Pandemic - A Call to Action. *JAMA Neurol.* 2018; 75(1): 9–10.
2. Delenclos M, Jones DR, McLean PJ, Uitti RJ. Biomarkers in Parkinson’s disease: Advances and strategies. *Park Relat Disord.* 2016; 22: S106-S110.
3. Burton A. How do we fix the shortage of neurologists?. *The Lancet Neurology.* 2018 Jun 1;17(6):502-3.
4. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson’s disease. *Mov Disord.* 2015;30(12):1591-1601. doi:10.1002/mds.26424
5. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* 2008; 23(15): 2129-70.
6. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology.* 2014; 83(5): 406-412.
7. Bajaj NPS, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study *Journal of Neurology, Neurosurgery & Psychiatry* 2010; 81: 1223-1228.
8. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson’s disease: Differential effects on bradykinesia. *Mov Disord.* 2007; 22(3): 328-333.
9. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson’s disease: Reliability and comparison with kinematic measures. *Mov Disord.* 2011; 26(10): 1859-1863.
10. Martinez-Martin P, Gil-Nagel A, Gracia LM, et al. Unified Parkinson’s disease rating scale characteristics and structure. *Mov Disord.* 1994; 9(1): 76-83.
11. Goetz CG, Stebbins GT. Assuring interrater reliability for the UPDRS motor section: Utility of the UPDRS teaching tape. *Mov Disord.* 2004; 19(12): 1453-1456.
12. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and amplitude of movement in Parkinson’s disease: A pilot study. *Mov Disord.* 2009; 24(7): 1001-1008.
13. Gao C, Smith S, Lones M, et al. Objective assessment of bradykinesia in Parkinson’s disease using evolutionary algorithms: Clinical validation. *Transl Neurodegener.* 2018; 7(1): 1-8.
14. Růžička E, Krupička R, Zárubová K, Ruzs J, Jech R, Szabó Z. Tests of manual dexterity and speed in Parkinson’s disease: Not all measure the same. *Park Relat Disord.* 2016; 28: 118-123.
15. Lee CY, Kang SJ, Hong S-K, Ma H-I, Lee U, Kim YJ. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson’s Disease. *PLoS One.* 2016;11(7).
16. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of blepharospasm severity. *Neurology.* 2016;87(20):2146-2153.
17. “Top Countries/Markets by Smartphone Penetration & Users.” Newzoo. Available from: <https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users/>. [Accessed 5 November 2019.]

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
18. Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci.* 2018; 21(9): 1281-1289.
 19. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-Dimensional Reconstruction of Paw and Digit Trajectories. *Journal of visualized experiments: JoVE.* 2019; Jul(149).
 20. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols.* 2019; 14(7): 2152-76.
 21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2016*; 770-778.
 22. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* 2004; 19(9): 1020-8.
 23. Ling H, Massey LA, Lees AJ, Brown P, Day BL. Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain.* 2012; 135(4): 1141-1153.
 24. Pinteá SL, Zheng J, Li X, Bank PJ, van Hilten JJ, van Gemert JC. Hand-tremor frequency estimation in videos. In *Proceedings of the European Conference on Computer Vision (ECCV) 2018*
 25. Khan T, Westin J, Dougherty M. Motion cue analysis for parkinsonian gait recognition. *The Open Biomedical Engineering Journal.* 2013; 7: 1.
 26. Khan T, Nyholm D, Westin J, Dougherty M. A computer vision framework for finger-tapping evaluation in Parkinson's disease. *Artif. Intell. Med.* 2014; 60(1): 27-40.
 27. Wong DC, Relton SD, Fang H, et al. Supervised Classification of Bradykinesia for Parkinson's Disease Diagnosis from Smartphone Videos. In *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS).* 2019; 32-37.
 28. Liu Y, Chen J, Hu C, et al. Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia. *IEEE Trans Neural Syst Rehabil Eng.* 2019;27(10):1952-1961.
 29. Noyce AJ, Nagy A, Acharya S, Hadavi S, Bestwick JP, Fearnley J, Lees AJ, Giovannoni G. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLoS One.* 2014; 9 (4).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

FIGURE LEGENDS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1. Illustration of the three parameters derived from the DeepLabCut finger tip and thumb tip coordinates to measure finger tapping speed, amplitude and rhythm.

Figure 2. Example video frames taken from smartphone video labelled by DeepLabCut, showing the six localization (labelling) points: thumb tip (dark blue), thumb metacarpophalangeal (MCP) joint (light blue), index finger tip (cyan), index finger MCP joint (yellow), middle finger tip (orange), dorsal wrist (red).

Figure 3. Computer measures of finger tapping speed, amplitude variation and rhythm, derived from DeepLabCut smartphone video tracking, correlate with clinician mode ratings for the same videos. Pearson correlation coefficients shown. MBRS: Modified Bradykinesia Rating Scale. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale.

TABLES

Table 1. The Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) Item 3.4 (Finger Tapping).

MDS-UPDRS Grade	Description
0: Normal	No problems
1: Slight	Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.
2: Mild	Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.
3: Moderate	Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.
4: Severe	Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 taps are evaluated by the clinician [5].

Table 2. The Modified Bradykinesia Rating Scale (MBRS).

Score	Speed	Amplitude	Rhythm
0	Normal	Normal	Regular, no arrests or pauses in ongoing movement
1	Mild Slowing	Mild reduction in amplitude in later performance, most movements close to normal	Mild impairment, up to two brief arrests in the 10 seconds, none lasting > 1 second
2	Moderate Slowing	Moderate, reduction in amplitude visible early in performance but continues to maintain 50% amplitude through most of the tasks	Moderate, 3 to 4 arrests in 10 seconds; OR 1 or 2 lasting > 1second
3	Severe slowing	Severe, less than 50% amplitude through most of the task	Severe, 5 or more arrests/10 seconds; OR more than 2 lasting > 1 second
4	Can barely perform the task	Can barely perform the task	Can barely perform the task

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 seconds of tapping are evaluated by the clinician to give three MBRS ratings (one for each of speed, amplitude, rhythm) [6,7].

Table 3. Study participant (hand video) characteristics.

	Patients	Controls
Age (Std. Dev.) yrs	68 (9.6)	59 (19.4)
Male/Female	47/26	22/38
Median years since diagnosis	4	n/a
Median H&Y [IQR]	2 [1,3]	n/a
H&Y = 1	32	
H&Y = 1.5	2	
H&Y = 2	12	
H&Y = 2.5	4	
H&Y = 3	19	
H&Y = 4	4	

H&Y = 5	0	
Median MDS-UPDRS [IQR]	2 [1,3]	1 [0,1]
MDS-UPDRS = 0	9	23
MDS-UPDRS = 1	20	23
MDS-UPDRS = 2	19	12
MDS-UPDRS = 3	20	2
MDS-UPDRS = 4	5	0
Visible tremor in video	11	0
Visible dystonia in video	2	0

Data is split by Parkinson's hands (n=73) and control hands (n=60). H&Y: modified Hoehn and Yahr scale [22]. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale, Item 3.4 (Finger Tapping). IQR: Interquartile Range.

Acknowledgements

We would like to thank the following UK movement disorder neurologists for providing clinical rating of videos: Dr Jane Alty, Dr Sundus Alusi, Dr Michael Bonello, Dr Stephen Butterworth, Dr Philip Buttery, Dr Camille Carroll, Dr Adam Cassidy, Dr Jeremy Cosgrove, Dr Richard Ellis, Dr Jonathan Evans, Dr Paul Goldsmith, Dr Donald Grosset, Dr Christopher Kobylecki, Dr Alistair Lansbury, Dr Peter Moore, Dr Rachel Newby, Dr Edward Newman, Dr Gillian Sare, Dr Monty Silverdale, Dr Naomi Warren, Dr Louise Wiblin, Dr Caroline Williams-Gray.

We would also like to thank Dr Parisa Patel and Dr Paschal O’Gorman for data collection.

Author contributions

Name	Location	Role	Contribution
-------------	-----------------	-------------	---------------------

Stefan Williams, MBBS, BSc	University of Leeds, UK	Author	Design and conceptualized study; data acquisition and interpretation; drafted the manuscript for intellectual content
Zhibin Zhao, BEng	(1) University of Manchester, UK; (2) Xi'an Jiatong University, China	Author	Analyzed the data (major role)
Awais Hafeez, MA, BSc	(1) University of Leeds, UK; (2) University of Engineering and Technology Lahore, Pakistan	Author	Design and conceptualized the study; major role in acquisition of data
David C Wong, PhD	University of Manchester, UK	Author	Design and conceptualized study; analyzed the data; data interpretation; revised the manuscript for intellectual content
Samuel D Relton, PhD	University of Leeds, UK	Author	Design and conceptualized study; data interpretation; revised the manuscript for intellectual content
Hui Fang, PhD	Loughborough University, UK	Author	Design and conceptualized study (major role)
Jane E Alty, MD, MA, MBBChir	(1) University of Tasmania, Australia; (2) Leeds Teaching Hospitals NHS Trust, UK	Author	Design and conceptualized study; data interpretation; drafted the manuscript for intellectual content

Title Page

The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia ?

Authors: Stefan Williams, MBBS, BSc; Zhibin Zhao, BEng; Awais Hafeez, MS, BSc; David C Wong, PhD; Samuel D Relton, PhD; Hui Fang, PhD; Jane E Alty MD, MA, MBBChir.

Stefan Williams, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Zhibin Zhao, (1) University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK; (2) Xi'an Jiatong University, School of Mechanical Engineering, Xi'an, China.

Awais Hafeez, (1) University of Leeds, School of Mechanical Engineering, Leeds, UK; (2) University of Engineering and Technology Lahore, Department of Mechatronics and Control Engineering, Lahore, Pakistan

David C Wong, University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK

Samuel D Relton, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Hui Fang, Loughborough University, Department of Computer Science, Loughborough, UK

Jane E Alty, (1) University of Tasmania, Wicking Dementia Research & Education Centre, Hobart, Australia (2) Leeds Teaching Hospitals NHS Trust UK.

Word count: **3099**

Number of Tables: **3**

Number of Figures: **3**

Number of Videos: **1**

Corresponding Author:

Stefan Williams

Leeds Institute of Health Science, 10th Floor, Worsley Building

University of Leeds, Clarendon Way, Leeds, UK LS2 9LU

Phone: 0113 343 0839

umswi@leeds.ac.uk

ORCID: 0000-0002-4786-1356

Keywords:

Parkinson's disease; Parkinsonism; Bradykinesia; Finger tapping; Computer Vision; Artificial Intelligence; **DeepLabCut**

1 **ABSTRACT**

6 **Objective**

8 The worldwide prevalence of Parkinson's disease is increasing. There is urgent need for new tools to objectively
9 measure the condition. Existing methods to record the cardinal motor feature of the condition, bradykinesia, using
10 wearable sensors or smartphone apps have not reached large-scale, routine use. We evaluate new computer vision
11 (artificial intelligence) technology, DeepLabCut, as a contactless method to quantify measures related to
12 Parkinson's bradykinesia from smartphone videos of finger tapping.
13
14
15
16
17
18
19

20 **Methods**

21 Standard smartphone video recordings of 133 hands performing finger tapping (39 idiopathic Parkinson's patients
22 and 30 controls) were tracked on a frame-by-frame basis with DeepLabCut. Objective computer measures of
23 tapping speed, amplitude and rhythm were correlated with clinical ratings made by 22 movement disorder
24 neurologists using the Modified Bradykinesia Rating Scale (MBRS) and Movement Disorder Society revision of
25 the Unified Parkinson's Disease Rating Scale (MDS-UPDRS).
26
27
28
29
30
31
32
33

34 **Results**

35 DeepLabCut reliably tracked and measured finger tapping in standard smartphone video. Computer measures
36 correlated well with clinical ratings of bradykinesia (Spearman coefficients): -0.74 speed, 0.66 amplitude, -0.65
37 rhythm for MBRS; -0.56 speed, 0.61 amplitude, -0.50 rhythm for MDS-UPDRS; 0.69 combined for MDS-
38 UPDRS. All $p < 0.001$.
39
40
41
42
43
44
45
46

47 **Conclusion**

48 New computer vision software, DeepLabCut, can quantify three measures related to Parkinson's bradykinesia
49 from smartphone videos of finger tapping. Objective 'contactless' measures of standard clinical examinations
50 were not previously possible with wearable sensors (accelerometers, gyroscopes, infrared markers). DeepLabCut
51 requires only conventional video recording of clinical examination and is entirely 'contactless'. This next
52 generation technology holds potential for Parkinson's and other neurological disorders with altered movements.
53
54
55
56
57
58
59
60
61
62
63
64
65

INTRODUCTION

1
2
3
4 The worldwide prevalence of Parkinson's is increasing, with an estimated 10 million people already
5 affected [1]. There is urgent need to find new tools to objectively measure the condition, to assist
6 diagnosis, monitoring, and research metrics [2]. This is especially pressing in the era of precision
7 medicine and with a global shortage of neurologists [3].
8
9

10
11
12
13
14
15 The cornerstone of Parkinson's assessment is the clinician's visual judgement of bradykinesia: one of
16 the cardinal motor features of the condition, which is defined as slowness of movement and decrement
17 in size or speed, or progressive hesitations/halts, as movements are continued [4]. There are several
18 clinical examination methods to assess bradykinesia, but in a brief clinic room assessment, one of the
19 most common is to observe the patient repeatedly tapping their index finger against thumb 'as quickly
20 and as big as possible' [5]. However, clinicians' visual interpretation is inherently subjective [6,7], with
21 considerable inter-rater variability [5,9-11]. Previous reported methods to measure bradykinesia on
22 finger tapping have required wearable equipment such as gyroscopes [9], electromagnetic sensors [12;
23 13], and infrared camera markers [14], or patient interaction with an app-based task [15], and this has
24 hindered the pathway to large-scale, routine clinical use.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Recent developments in computer interpretation of video (computer vision), clear the pathway to
41 'contactless' methods of quantifying clinical assessments [16], with potential widespread global
42 application because of the availability of inexpensive smartphone cameras [17]. DeepLabCut is recently
43 published open source computer vision software, without a requirement to write new computer code,
44 that has been demonstrated to track and measure mice and fruit fly body parts [18,19]. We apply it to
45 clinical neurology for the first time and assess whether the technology can be used to track and measure
46 bradykinesia in standard smartphone video of conventional finger tapping examination.
47
48
49
50
51
52
53
54
55
56
57
58
59

METHODS

1
2 Ethical review
3
4
5

6 The study was approved by the North of Scotland Research Ethics Committee, United Kingdom Health
7
8 Research Authority (IRAS project ID 256116).
9

10
11
12
13 Participants and video recording
14
15
16

17 Video recordings of 138 hands (including left and right) were obtained from 39 patients with idiopathic
18 Parkinson's disease and 30 healthy controls, who gave written consent. One video was rejected because
19
20 the hand moved outside the video frame, making 137 videos: 77 Parkinson's hands and 60 control
21
22 hands.
23
24
25

26
27
28 All Parkinson's patients had previously been diagnosed by a movement disorder specialist neurologist
29 at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society
30 clinical diagnostic criteria [4], and were subjectively and objectively in the 'on' state at the time of
31 video recording (no medication was withheld prior to recording). We did not exclude patients with
32
33 postural hand tremor or dystonia. Controls were recruited from the companions of patients, or
34
35 hospital/university staff, and had no history of Parkinson's disease or other neurological diagnosis.
36
37
38
39
40
41
42

43
44 Participants rested their elbow on a chair arm with the forearm lifted to a 45 degree angle. The hand
45 was free to move as per the protocol of the Movement Disorder Society revision of the Unified
46 Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 Finger Tapping [5]. Only the hand/forearm
47
48 was within the video frame. The distance from camera to hand was approximately 1m, but not tightly
49
50 defined. Digits 1 and 2 were closest to the camera. No specific instructions were given for the position
51
52 of digits 3 to 5, and participants were free to position these digits as they preferred, although the
53
54 researcher gave a brief demonstration in which digits 3 to 5 were extended.
55
56
57
58
59
60
61
62

1 A smartphone, placed on a tripod, was used to record standard video (60 frames per second, 1920x1080
2 pixel resolution), with only ambient lighting. Participants were instructed to tap their index finger and
3 thumb together “as quickly and as big as possible” for at least 10 seconds [5]. Each video was edited to
4 an 11 second clip: 1 second prior to tapping onset and 10 seconds of finger tapping.
5
6
7
8
9

10 Clinical rating

11
12
13
14
15 22 consultant neurologists specialising in movement disorders (United Kingdom) were asked to rate 30
16 tapping videos each, selected at random from the set of 137 videos. **The median number of raters per**
17 **video was 5 (range 1 to 12, interquartile range 3 to 7).** The raters were blinded to patient / healthy
18 control status and each other’s scores. Each video was rated according to both the MDS-UPDRS Item
19 3.4 Finger Tapping [5], and the Modified Bradykinesia Rating Scale (MBRS) [8,9]. The MDS-UPDRS
20 requires the rater to amalgamate judgments of finger tapping speed, amplitude, and rhythm into a single
21 composite score (**Table 1**). In contrast, the MBRS is comprised of three separate scores for speed,
22 amplitude, and rhythm (**Table 2**). The modal clinician rating for each video was used for correlation
23 with computer measures (see below). For each possible pair of raters, Cohen Kappa was calculated
24 based on all videos they had both seen. The resulting values had mean 0.28 (fair agreement) and
25 standard deviation 0.28.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 DeepLabCut video tracking

43
44
45
46 DeepLabCut tracks the geometrical configuration of multiple body parts in video, without a requirement
47 to wear markers [18,20]. It uses transfer learning, with feature detectors based on deep neural networks
48 that have been pretrained on ImageNet, a massive object recognition dataset [21]. This means that
49 DeepLabCut is able to accurately recognise and track body parts with minimal training data [18].
50
51
52
53
54
55
56

57 The set of 137 finger tapping videos was processed by DeepLabCut. We localised (labelled) six distinct
58 points on the tapping hand in 20 frames selected by k-means clustering from each 660 frame video:
59
60
61
62
63

1 thumb tip; index finger tip; thumb metacarpophalangeal (MCP) joint; index finger MCP joint; middle
2 finger tip; dorsal wrist/proximal dorsal hand. The deep neural network architecture of DeepLabCut was
3
4 then trained using these points to predict their localisation in the remaining 97% of (unlabelled) video
5
6 frames (1030000 training iterations, ResNet 50). This created video pixel coordinates for finger tip and
7
8 thumb tip throughout each video.
9

10
11
12
13 The accuracy of DeepLabCut to track the hand localisation points (including finger tip and thumb tip)
14
15 was assessed using the ‘evaluate network’ function within DeepLabCut. This function computes the
16
17 Euclidean error between the manual labels and the ones predicted by DeepLabCut averaged over the
18
19 hand locations and test images (mean absolute error, proportional to the average root mean square error)
20
21 [20].
22
23
24
25

26 Signal processing

27
28
29
30
31 Video pixel coordinates for the labels produced by DeepLabCut were used to calculate the pixel distance
32
33 between index finger tip and thumb tip. A Savitzky-Golay filter was applied to the resulting time series,
34
35 which removed large, sudden transient label ‘jumps’ caused by DeepLabCut mislabelling (i.e. large-
36
37 distance, physiologically impossible label movement across one pair of video frames). The pixel
38
39 number distances were standardised across all videos, by using the maximum opening distance between
40
41 finger and thumb tip and normalising this to a value of 1 (all values were divided by the maximal value
42
43 in the corresponding time series).
44
45
46
47
48

49 Three features of the resulting finger tip to thumb tip distance time series were calculated to reflect the
50
51 clinical features of finger tapping speed, amplitude and rhythm (**Figure 1**). A measure of speed was
52
53 calculated as the mean rate of change of the normalised distance between finger and thumb tip over
54
55 time. A measure of amplitude variability was calculated by dividing each time series into one-second
56
57 windows, with maximal overlap, and then calculating the coefficient of variation of the mean difference
58
59 between the maximum and minimum amplitudes in each window. A measure of rhythm regularity was
60
61
62
63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

calculated by undertaking Fast Fourier Transform to find the distribution of frequencies within each finger tap time series, and then measuring the power of the dominant frequency peak added to the power of the frequencies 0.2 Hz either side of it. A more regular tapping rhythm concentrates power in a narrow frequency band, increasing the power of the dominant frequency peak (and its immediate neighbours), whereas a more irregular tapping rhythm leads to a more widely spread distribution of frequency bands, reducing the power of the dominant frequency peak (and its immediate neighbours).

Spearman correlation coefficients were calculated for the each of the three DeepLabCut measures vs the modal MBRS and MDS-UPDRS clinician ratings. In addition, we calculated the Spearman correlation coefficient of the three computer scores combined (the normalised arithmetic mean of the computer speed score, amplitude variability score and rhythm regularity score) with the MDS-UPDRS rating.

RESULTS

The participant and hand video characteristics are given in **Table 3**.

Tracking accuracy

DeepLabCut reliably tracked the finger tapping movements across 133 videoed clinical examinations. **Figure 2** shows example frames from a video labelled by DeepLabCut. The mean absolute error of DeepLabCut labelling was 8.39 pixels, i.e. the average distance between manual (human) labels and those predicted by DeepLabCut was 8.39 pixels within a 1920x1080 pixel video frame. **Video 1** shows examples of four videos labelled by DeepLabCut, including particularly challenging cases with tremor and dystonia also present.

Correlation with clinical bradykinesia ratings

1
2 The quantitative scores derived from DeepLabCut measurements correlated well with the group
3
4 neurologists' modal ratings of bradykinesia; **Figure 3**. The computer vision measures of tapping speed,
5
6 amplitude variability and rhythm regularity had good correlations with the respective MBRS clinical
7
8 ratings for speed (-0.70, $p < 0.001$), amplitude (0.65, $p < 0.001$) and rhythm (-0.61, $p < 0.001$). There
9
10 were also good correlations between the computer measures of tapping speed (-0.66, $p < 0.001$),
11
12 amplitude variability (0.56, $p < 0.001$) and rhythm regularity (-0.50, $p < 0.001$) and the MDS-UPDRS
13
14 clinical rating. The MDS-UPDRS is a composite clinical rating, combining the separate components of
15
16 bradykinesia, and the three computer measures combined also correlated with MDS-UPDRS (-0.69, p
17
18 < 0.001).

26 **DISCUSSION**

27
28
29
30
31 We have shown that new computer vision software, DeepLabCut, can track finger tapping in a standard
32
33 smartphone video recording of a conventional clinical examination, and quantify three measures related
34
35 to Parkinson's bradykinesia: speed, amplitude variation and rhythm regularity. We found good
36
37 correlations between the objective computer measures of finger tapping and clinical ratings of
38
39 bradykinesia by 22 movement disorder neurologists using two gold standard clinical scales (MBRS and
40
41 MDS-UPDRS). Thus, smartphone video measures using DeepLabCut exhibit convergent validity with
42
43 conventional clinician ratings.
44
45

46
47
48
49 These results are new and clinically significant. This is the first study to apply new open source
50
51 DeepLabCut technology, developed for animal behaviour studies, to the field of human movement
52
53 disorders, or indeed any aspect of clinical neurology. We demonstrate new objective measures related
54
55 to bradykinesia by bringing together two ubiquitous ingredients: the smartphone video camera and
56
57 visual artificial intelligence software. With the worldwide ownership of smartphones so high (e.g. 82%
58
59 in the UK, 77% in the US) [17], there are no hardware barriers to further use of this approach.
60
61
62

1
2 Neurologists recognize that quantification of bradykinesia will improve Parkinson's management
3
4 through 'precision medicine' principles: it has the potential to aid early accurate diagnosis and remote
5
6 monitoring, and provide fine grain regular measurement of clinical trial outcomes. Previous methods to
7
8 quantify bradykinesia have focussed on wearable sensors measuring finger tapping [9,12-14,23] but it
9
10 has been challenging to incorporate these into widespread use because equipment needs to be physically
11
12 attached to the patient. Furthermore, the necessity for sensors to stay fixed on a specific body part has
13
14 often required the exclusion of tremulous and dyskinetic patients [8,9,12,14,23]. In contrast, the
15
16 DeepLabCut test presented here is a contactless method that simply requires a short 10 second video
17
18 recording of the standard clinical examination. This means it can potentially be used in any setting and
19
20 any patient.
21
22
23
24
25

26
27 The 'real life' versatility of DeepLabCut in ordinary clinical settings was evaluated in this study because
28
29 no special efforts were made to optimize participant positioning or filming conditions. The tracking and
30
31 measures were robust despite variations in smartphone camera distance, positioning of the hand, skin
32
33 color, flexion / extension postures of digits 3 to 5, ambient lighting conditions, and inclusion of a small
34
35 number of patients with tremor or dystonia during tapping. **Video 1** illustrates this, with accurate
36
37 labelling of finger and thumb tip despite superimposed tremor or dystonia. This is particularly
38
39 encouraging for the relevance of DeepLabCut aiding assessment of Parkinson's in conventional clinical
40
41 settings.
42
43
44
45

46
47 Two further strengths of our study are the collection of clinical ratings from a large group of movement
48
49 disorder neurologists, and the use of two rating scales: the MBRS and MDS-UPDRS, such that the
50
51 clinical ratings were representative and reliable. Clinical rating scales are vulnerable to inter-rater
52
53 variability but using modal values derived from a large number of blinded clinical raters, and two
54
55 different gold-standard rating scales, provided a more robust 'ground truth' of clinical bradykinesia to
56
57 evaluate the new test against. As such we were able to demonstrate that clinician and DeepLabCut
58
59 measures correlate for both individual components of bradykinesia (MBRS) and an overall combination
60
61
62
63

1 of those components (MDS-UPDRS). An interesting future application with a larger participant group
2 would be to examine the relative contribution of the three video measures to individual grades of
3
4 UPDRS bradykinesia.
5

6
7
8 Another advantage of the approach described here is that it is not limited by the physical constraints of
9 wearable or screen/keyboard tapping methods, and could be applied to detect and measure other motor
10 features of Parkinson's in a contactless manner such as tremor, gait, or posture. Computer vision is
11 general has potential for application to these signs [24,25].
12
13
14
15
16
17

18
19
20 Only three previous studies have objectively measured finger tapping using contactless, standard video
21 analysis. One method extracted finger tapping information from 13 Parkinson's patients with advanced
22 disease [26], but it required inclusion of the patient's face in the video, limiting the practical utility. The
23 second report, from our own group, described a video method to measure movement using an optical
24 flow field [27] that interprets movements of the whole hand. The work here improves on this by
25 explicitly extracting the salient thumb and finger tip points, and validating in a larger cohort. A third
26 study of 60 Parkinson's patients (no controls) found strong correlation between MDS-UPDRS finger
27 tap rating and video measures of tap interval (frequency) ($r=0.91$), frequency variation ($r=0.82$) and
28 amplitude ($r=-0.94$) but not amplitude variation ($r=0.39$) [28]. The 'ground truth' of the clinical
29 comparison was less robust, based on just two clinical raters, and there were no MBRS ratings;
30 furthermore, some images were blurred in the 24 frame per second videos. A major advantage of our
31 method in comparison is that DeepLabCut is open source software, that can be downloaded and used
32 without the need to write computer code, making it available now to neurologists with only limited
33 computing knowledge.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53 A well-studied, non-camera method to measure finger tap bradykinesia is the 'Bradykinesia-Akinesia
54 Incoordination (BRAIN) Test', which involves participants tapping a standard computer keyboard for
55 30 seconds [29]. This shares advantages with our method in that it involves no special equipment, and
56 could be employed without clinician-patient contact. However, it requires patient motivation to engage
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

with a keyboard tapping task, rather than a simple recording of existing standard clinical examination (and similar requirements for patient motivation apply to patient smartphone apps [15]). Furthermore, ‘BRAIN’ correlations with MDS-UPDRS finger tapping ratings of 0.44, 0.20, 0.33, 0.03 [29], are weaker than those we report (0.66, 0.56, 0.50, 0.69). One reason for this may be that a computer keyboard cannot record any measure of tap amplitude, and the speed measure is limited to tap frequency, so that some core aspects of bradykinesia cannot be captured by ‘BRAIN’.

The present study has some limitations. Like human vision, a simple camera lacks an absolute measure of distance, and our method cannot capture 3D movement. Rotation of the thumb and finger into a horizontal plane would falsely alter amplitude measurement. However, in practice such movement is rare during tapping, and our normalized measure of finger-thumb distance based on 2D maximum opening distance appeared to capture amplitude, speed and rhythm abnormalities well, given the good correlations with clinical ratings.

We have measured amplitude variation, but not amplitude decrement (the sequence effect). Previous approaches have measured decrement as a straight line linear regression [23], but decrement would also proportionately affect a measure of amplitude variation, such that the variations in decrement should be captured within our measure. Furthermore, there are several other previous reports of good correlation between non-decrement amplitude measures using sensors and clinician ratings of finger tap bradykinesia [9].

Relatively few videos were rated as grade 3 or 4 bradykinesia, and it is possible that a greater number of higher grade videos might change the strength of correlations. Even with accurate tracking, all forms of rating or measuring bradykinesia (including clinician) may potentially be confounded by comorbidities such as joint deformities, pain, dystonia etc. However, with a larger dataset, machine learning techniques are well-suited to separate such additional contributions from underlying bradykinesia. There was a degree of error in DeepLabCut labelling of finger and thumb tips, but this was minimal (mean of 8.39 pixels within a 1920x1080 pixel video frame) and unlikely to affect

1 individual measures. We mitigated against this by applying a filter prior to feature extraction. Full
2 flexion of digits 3 to 5 might improve labelling accuracy, but would also limit the amplitude of finger
3 tapping [23], potentially masking more subtle deficits. The labelling error provided by DeepLabCut's
4 'evaluate network' function only gives an accuracy figure for the entire set of videos but not for
5 individual videos, meaning that we have not reported accuracy in subsets of videos, such as those with
6 visible tremor. This study was exploratory, aimed at proof of concept that bradykinesia related measures
7 could be derived from smartphone video. The protocol was thus simple, and repeat measures such as
8 test-retest in controls, or 'off' and 'on' medication states, were not included, meaning that reliability
9 and sensitivity to change have not been characterised. Future work will assess this.
10
11
12
13
14
15
16
17
18
19
20
21

22 Finally, we have demonstrated measurements that show convergent validity with clinical ratings, rather
23 than a test to discriminate patients from controls. In our view, it makes little sense to pursue automated
24 diagnostic classification using finger tapping video alone because clinical Parkinson's diagnosis is by
25 definition a broader assessment than isolated finger tapping performance [4].
26
27
28
29
30
31

32
33 This study has evaluated the clinical application of DeepLabCut in a medical condition, specifically a
34 neurological disorder. The methodology described here provides a new objective measure of
35 Parkinson's bradykinesia in conventional clinical examination. DeepLabCut provides a contactless
36 means for tracking and measuring bradykinesia in Parkinson's, without any special equipment other
37 than the ubiquitous smartphone. This research therefore provides a new tool to quantify bradykinesia.
38 It could potentially be used to support diagnosis and monitoring of Parkinson's, both clinically and as
39 an outcome measure in research studies. Beyond Parkinson's and other movement disorders, this same
40 technology could be applied to many other neurological conditions that require measurement of
41 movements or function; for example multiple sclerosis, neuropathies, epilepsy and stroke.
42
43
44
45
46
47
48
49
50
51
52
53

54 REFERENCES

55
56
57
58
59
60
61
62
63
64
65

1. Dorsey ER, Bloem BR. The Parkinson Pandemic - A Call to Action. *JAMA Neurol.* 2018; 75(1): 9–10.
2. Delenclos M, Jones DR, McLean PJ, Uitti RJ. Biomarkers in Parkinson’s disease: Advances and strategies. *Park Relat Disord.* 2016; 22: S106-S110.
3. Burton A. How do we fix the shortage of neurologists?. *The Lancet Neurology.* 2018 Jun 1;17(6):502-3.
4. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson’s disease. *Mov Disord.* 2015;30(12):1591-1601. doi:10.1002/mds.26424
5. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* 2008; 23(15): 2129-70.
6. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology.* 2014; 83(5): 406-412.
7. Bajaj NPS, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study *Journal of Neurology, Neurosurgery & Psychiatry* 2010; 81: 1223-1228.
8. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson’s disease: Differential effects on bradykinesia. *Mov Disord.* 2007; 22(3): 328-333.
9. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson’s disease: Reliability and comparison with kinematic measures. *Mov Disord.* 2011; 26(10): 1859-1863.
10. Martinez-Martin P, Gil-Nagel A, Gracia LM, et al. Unified Parkinson’s disease rating scale characteristics and structure. *Mov Disord.* 1994; 9(1): 76-83.
11. Goetz CG, Stebbins GT. Assuring interrater reliability for the UPDRS motor section: Utility of the UPDRS teaching tape. *Mov Disord.* 2004; 19(12): 1453-1456.
12. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and amplitude of movement in Parkinson’s disease: A pilot study. *Mov Disord.* 2009; 24(7): 1001-1008.
13. Gao C, Smith S, Lones M, et al. Objective assessment of bradykinesia in Parkinson’s disease using evolutionary algorithms: Clinical validation. *Transl Neurodegener.* 2018; 7(1): 1-8.
14. Růžička E, Krupička R, Zárubová K, Ruzs J, Jech R, Szabó Z. Tests of manual dexterity and speed in Parkinson’s disease: Not all measure the same. *Park Relat Disord.* 2016; 28: 118-123.
15. Lee CY, Kang SJ, Hong S-K, Ma H-I, Lee U, Kim YJ. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson’s Disease. *PLoS One.* 2016;11(7).
16. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of blepharospasm severity. *Neurology.* 2016;87(20):2146-2153.
17. “Top Countries/Markets by Smartphone Penetration & Users.” Newzoo. Available from: <https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users/>. [Accessed 5 November 2019.]

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
18. Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci.* 2018; 21(9): 1281-1289.
 19. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-Dimensional Reconstruction of Paw and Digit Trajectories. *Journal of visualized experiments: JoVE.* 2019; Jul(149).
 20. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols.* 2019; 14(7): 2152-76.
 21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2016*; 770-778.
 22. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* 2004; 19(9): 1020-8.
 23. Ling H, Massey LA, Lees AJ, Brown P, Day BL. Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain.* 2012; 135(4): 1141-1153.
 24. Pinteá SL, Zheng J, Li X, Bank PJ, van Hilten JJ, van Gemert JC. Hand-tremor frequency estimation in videos. In *Proceedings of the European Conference on Computer Vision (ECCV) 2018*
 25. Khan T, Westin J, Dougherty M. Motion cue analysis for parkinsonian gait recognition. *The Open Biomedical Engineering Journal.* 2013; 7: 1.
 26. Khan T, Nyholm D, Westin J, Dougherty M. A computer vision framework for finger-tapping evaluation in Parkinson's disease. *Artif. Intell. Med.* 2014; 60(1): 27-40.
 27. Wong DC, Relton SD, Fang H, et al. Supervised Classification of Bradykinesia for Parkinson's Disease Diagnosis from Smartphone Videos. In *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS).* 2019; 32-37.
 28. Liu Y, Chen J, Hu C, et al. Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia. *IEEE Trans Neural Syst Rehabil Eng.* 2019;27(10):1952-1961.
 29. Noyce AJ, Nagy A, Acharya S, Hadavi S, Bestwick JP, Fearnley J, Lees AJ, Giovannoni G. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLoS One.* 2014; 9 (4).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

FIGURE LEGENDS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1. Illustration of the three parameters derived from the DeepLabCut finger tip and thumb tip coordinates to measure finger tapping speed, amplitude and rhythm.

Figure 2. Example video frames taken from smartphone video labelled by DeepLabCut, showing the six localization (labelling) points: thumb tip (dark blue), thumb metacarpophalangeal (MCP) joint (light blue), index finger tip (cyan), index finger MCP joint (yellow), middle finger tip (orange), dorsal wrist (red).

Figure 3. Computer measures of finger tapping speed, amplitude variation and rhythm, derived from DeepLabCut smartphone video tracking, correlate with clinician mode ratings for the same videos. Pearson correlation coefficients shown. MBRS: Modified Bradykinesia Rating Scale. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale.

TABLES

Table 1. The Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) Item 3.4 (Finger Tapping).

MDS-UPDRS Grade	Description
0: Normal	No problems
1: Slight	Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.
2: Mild	Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.
3: Moderate	Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.
4: Severe	Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 taps are evaluated by the clinician [5].

Table 2. The Modified Bradykinesia Rating Scale (MBRS).

Score	Speed	Amplitude	Rhythm
0	Normal	Normal	Regular, no arrests or pauses in ongoing movement
1	Mild Slowing	Mild reduction in amplitude in later performance, most movements close to normal	Mild impairment, up to two brief arrests in the 10 seconds, none lasting > 1 second
2	Moderate Slowing	Moderate, reduction in amplitude visible early in performance but continues to maintain 50% amplitude through most of the tasks	Moderate, 3 to 4 arrests in 10 seconds; OR 1 or 2 lasting > 1second
3	Severe slowing	Severe, less than 50% amplitude through most of the task	Severe, 5 or more arrests/10 seconds; OR more than 2 lasting > 1 second
4	Can barely perform the task	Can barely perform the task	Can barely perform the task

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 seconds of tapping are evaluated by the clinician to give three MBRS ratings (one for each of speed, amplitude, rhythm) [6,7].

Table 3. Study participant (hand video) characteristics.

	Patients	Controls
Age (Std. Dev.) yrs	68 (9.6)	59 (19.4)
Male/Female	47/26	22/38
Median years since diagnosis	4	n/a
Median H&Y [IQR]	2 [1,3]	n/a
H&Y = 1	32	
H&Y = 1.5	2	
H&Y = 2	12	
H&Y = 2.5	4	
H&Y = 3	19	
H&Y = 4	4	

H&Y = 5	0	
Median MDS-UPDRS [IQR]	2 [1,3]	1 [0,1]
MDS-UPDRS = 0	9	23
MDS-UPDRS = 1	20	23
MDS-UPDRS = 2	19	12
MDS-UPDRS = 3	20	2
MDS-UPDRS = 4	5	0
Visible tremor in video	11	0
Visible dystonia in video	2	0

Data is split by Parkinson's hands (n=73) and control hands (n=60). H&Y: modified Hoehn and Yahr scale [22]. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale, Item 3.4 (Finger Tapping). IQR: Interquartile Range.

Acknowledgements

We would like to thank the following UK movement disorder neurologists for providing clinical rating of videos: Dr Jane Alty, Dr Sundus Alusi, Dr Michael Bonello, Dr Stephen Butterworth, Dr Philip Buttery, Dr Camille Carroll, Dr Adam Cassidy, Dr Jeremy Cosgrove, Dr Richard Ellis, Dr Jonathan Evans, Dr Paul Goldsmith, Dr Donald Grosset, Dr Christopher Kobylecki, Dr Alistair Lansbury, Dr Peter Moore, Dr Rachel Newby, Dr Edward Newman, Dr Gillian Sare, Dr Monty Silverdale, Dr Naomi Warren, Dr Louise Wiblin, Dr Caroline Williams-Gray.

We would also like to thank Dr Parisa Patel and Dr Paschal O’Gorman for data collection.

Author contributions

Name	Location	Role	Contribution
-------------	-----------------	-------------	---------------------

Stefan Williams, MBBS, BSc	University of Leeds, UK	Author	Design and conceptualized study; data acquisition and interpretation; drafted the manuscript for intellectual content
Zhibin Zhao, BEng	(1) University of Manchester, UK; (2) Xi'an Jiatong University, China	Author	Analyzed the data (major role)
Awais Hafeez, MA, BSc	(1) University of Leeds, UK; (2) University of Engineering and Technology Lahore, Pakistan	Author	Design and conceptualized the study; major role in acquisition of data
David C Wong, PhD	University of Manchester, UK	Author	Design and conceptualized study; analyzed the data; data interpretation; revised the manuscript for intellectual content
Samuel D Relton, PhD	University of Leeds, UK	Author	Design and conceptualized study; data interpretation; revised the manuscript for intellectual content
Hui Fang, PhD	Loughborough University, UK	Author	Design and conceptualized study (major role)
Jane E Alty, MD, MA, MBBChir	(1) University of Tasmania, Australia; (2) Leeds Teaching Hospitals NHS Trust, UK	Author	Design and conceptualized study; data interpretation; drafted the manuscript for intellectual content

Title Page

The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia ?

Authors: Stefan Williams, MBBS, BSc; Zhibin Zhao, BEng; Awais Hafeez, MS, BSc; David C Wong, PhD; Samuel D Relton, PhD; Hui Fang, PhD; Jane E Alty MD, MA, MBBChir.

Stefan Williams, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Zhibin Zhao, (1) University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK; (2) Xi'an Jiatong University, School of Mechanical Engineering, Xi'an, China.

Awais Hafeez, (1) University of Leeds, School of Mechanical Engineering, Leeds, UK; (2) University of Engineering and Technology Lahore, Department of Mechatronics and Control Engineering, Lahore, Pakistan

David C Wong, University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK

Samuel D Relton, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Hui Fang, Loughborough University, Department of Computer Science, Loughborough, UK

Jane E Alty, (1) University of Tasmania, Wicking Dementia Research & Education Centre, Hobart, Australia (2) Leeds Teaching Hospitals NHS Trust UK.

Word count: **3101**

Number of Tables: **3**

Number of Figures: **3**

Number of Videos: **1**

Corresponding Author:

Stefan Williams

Leeds Institute of Health Science, 10th Floor, Worsley Building

University of Leeds, Clarendon Way, Leeds, UK LS2 9LU

Phone: 0113 343 0839

umswi@leeds.ac.uk

ORCID: 0000-0002-4786-1356

Keywords:

Parkinson's disease; Parkinsonism; Bradykinesia; Finger tapping; Computer Vision; Artificial Intelligence; DeepLabCut

1 **ABSTRACT**

6 **Objective**

8 The worldwide prevalence of Parkinson's disease is increasing. There is urgent need for new tools to objectively
9 measure the condition. Existing methods to record the cardinal motor feature of the condition, bradykinesia, using
10 wearable sensors or smartphone apps have not reached large-scale, routine use. We evaluate new computer vision
11 (artificial intelligence) technology, DeepLabCut, as a contactless method to quantify measures related to
12 Parkinson's bradykinesia from smartphone videos of finger tapping.
13
14
15
16
17
18
19

20 **Methods**

21 Standard smartphone video recordings of 133 hands performing finger tapping (39 idiopathic Parkinson's patients
22 and 30 controls) were tracked on a frame-by-frame basis with DeepLabCut. Objective computer measures of
23 tapping speed, amplitude and rhythm were correlated with clinical ratings made by 22 movement disorder
24 neurologists using the Modified Bradykinesia Rating Scale (MBRS) and Movement Disorder Society revision of
25 the Unified Parkinson's Disease Rating Scale (MDS-UPDRS).
26
27
28
29
30
31
32
33

34 **Results**

35 DeepLabCut reliably tracked and measured finger tapping in standard smartphone video. Computer measures
36 correlated well with clinical ratings of bradykinesia (Spearman coefficients): -0.74 speed, 0.66 amplitude, -0.65
37 rhythm for MBRS; -0.56 speed, 0.61 amplitude, -0.50 rhythm for MDS-UPDRS; 0.69 combined for MDS-
38 UPDRS. All $p < 0.001$.
39
40
41
42
43
44
45
46

47 **Conclusion**

48 New computer vision software, DeepLabCut, can quantify three measures related to Parkinson's bradykinesia
49 from smartphone videos of finger tapping. Objective 'contactless' measures of standard clinical examinations
50 were not previously possible with wearable sensors (accelerometers, gyroscopes, infrared markers). DeepLabCut
51 requires only conventional video recording of clinical examination and is entirely 'contactless'. This next
52 generation technology holds potential for Parkinson's and other neurological disorders with altered movements.
53
54
55
56
57
58
59
60
61
62
63
64
65

INTRODUCTION

1
2
3
4 The worldwide prevalence of Parkinson's is increasing, with an estimated 10 million people already
5 affected [1]. There is urgent need to find new tools to objectively measure the condition, to assist
6 diagnosis, monitoring, and research metrics [2]. This is especially pressing in the era of precision
7 medicine and with a global shortage of neurologists [3].
8
9

10
11
12
13
14
15 The cornerstone of Parkinson's assessment is the clinician's visual judgement of bradykinesia: one of
16 the cardinal motor features of the condition, which is defined as slowness of movement and decrement
17 in size or speed, or progressive hesitations/halts, as movements are continued [4]. There are several
18 clinical examination methods to assess bradykinesia, but in a brief clinic room assessment, one of the
19 most common is to observe the patient repeatedly tapping their index finger against thumb 'as quickly
20 and as big as possible' [5]. However, clinicians' visual interpretation is inherently subjective [6,7], with
21 considerable inter-rater variability [5,9-11]. Previous reported methods to measure bradykinesia on
22 finger tapping have required wearable equipment such as gyroscopes [9], electromagnetic sensors [12;
23 13], and infrared camera markers [14], or patient interaction with an app-based task [15], and this has
24 hindered the pathway to large-scale, routine clinical use.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Recent developments in computer interpretation of video (computer vision), clear the pathway to
41 'contactless' methods of quantifying clinical assessments [16], with potential widespread global
42 application because of the availability of inexpensive smartphone cameras [17]. DeepLabCut is recently
43 published open source computer vision software, without a requirement to write new computer code,
44 that has been demonstrated to track and measure mice and fruit fly body parts [18,19]. We apply it to
45 clinical neurology for the first time and assess whether the technology can be used to track and measure
46 bradykinesia in standard smartphone video of conventional finger tapping examination.
47
48
49
50
51
52
53
54
55
56
57
58
59

METHODS

1
2 Ethical review
3
4
5

6 The study was approved by the North of Scotland Research Ethics Committee, United Kingdom Health
7
8 Research Authority (IRAS project ID 256116).
9

10
11
12
13 Participants and video recording
14
15

16
17 Video recordings of 138 hands (including left and right) were obtained from 39 patients with idiopathic
18 Parkinson's disease and 30 healthy controls, who gave written consent. One video was rejected because
19
20 the hand moved outside the video frame, making 137 videos: 77 Parkinson's hands and 60 control
21
22 hands.
23
24
25

26
27
28 All Parkinson's patients had previously been diagnosed by a movement disorder specialist neurologist
29 at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society
30 clinical diagnostic criteria [4], and were subjectively and objectively in the 'on' state at the time of
31 video recording (no medication was withheld prior to recording). We did not exclude patients with
32
33 postural hand tremor or dystonia. Controls were recruited from the companions of patients, or
34
35 hospital/university staff, and had no history of Parkinson's disease or other neurological diagnosis.
36
37
38
39
40
41
42

43
44 Participants rested their elbow on a chair arm with the forearm lifted to a 45 degree angle. The hand
45 was free to move as per the protocol of the Movement Disorder Society revision of the Unified
46 Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 Finger Tapping [5]. Only the hand/forearm
47
48 was within the video frame. The distance from camera to hand was approximately 1m, but not tightly
49
50 defined. Digits 1 and 2 were closest to the camera. No specific instructions were given for the position
51
52 of digits 3 to 5, and participants were free to position these digits as they preferred, although the
53
54 researcher gave a brief demonstration in which digits 3 to 5 were extended.
55
56
57
58
59
60
61
62

1 A smartphone, placed on a tripod, was used to record standard video (60 frames per second, 1920x1080
2 pixel resolution), with only ambient lighting. Participants were instructed to tap their index finger and
3 thumb together “as quickly and as big as possible” for at least 10 seconds [5]. Each video was edited to
4 an 11 second clip: 1 second prior to tapping onset and 10 seconds of finger tapping.
5
6
7
8
9

10 Clinical rating

11
12
13
14
15 22 consultant neurologists specialising in movement disorders (United Kingdom) were asked to rate 30
16 tapping videos each, selected at random from the set of 137 videos. The median number of raters per
17 video was 5 (range 1 to 12, interquartile range 3 to 7). The raters were blinded to patient / healthy
18 control status and each other’s scores. Each video was rated according to both the MDS-UPDRS Item
19 3.4 Finger Tapping [5], and the Modified Bradykinesia Rating Scale (MBRS) [8,9]. The MDS-UPDRS
20 requires the rater to amalgamate judgments of finger tapping speed, amplitude, and rhythm into a single
21 composite score (**Table 1**). In contrast, the MBRS is comprised of three separate scores for speed,
22 amplitude, and rhythm (**Table 2**). The modal clinician rating for each video was used for correlation
23 with computer measures (see below). For each possible pair of raters, Cohen Kappa was calculated
24 based on all videos they had both seen. The resulting values had mean 0.28 (fair agreement) and
25 standard deviation 0.28.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 DeepLabCut video tracking

43
44
45
46 DeepLabCut tracks the geometrical configuration of multiple body parts in video, without a requirement
47 to wear markers [18,20]. It uses transfer learning, with feature detectors based on deep neural networks
48 that have been pretrained on ImageNet, a massive object recognition dataset [21]. This means that
49 DeepLabCut is able to accurately recognise and track body parts with minimal training data [18].
50
51
52
53
54
55
56

57 The set of 137 finger tapping videos was processed by DeepLabCut. We localised (labelled) six distinct
58 points on the tapping hand in 20 frames selected by k-means clustering from each 660 frame video:
59
60
61
62
63

1 thumb tip; index finger tip; thumb metacarpophalangeal (MCP) joint; index finger MCP joint; middle
2 finger tip; dorsal wrist/proximal dorsal hand. The deep neural network architecture of DeepLabCut was
3
4 then trained using these points to predict their localisation in the remaining 97% of (unlabelled) video
5
6 frames (1030000 training iterations, ResNet 50). This created video pixel coordinates for finger tip and
7
8 thumb tip throughout each video.
9

10
11
12
13 The accuracy of DeepLabCut to track the hand localisation points (including finger tip and thumb tip)
14
15 was assessed using the ‘evaluate network’ function within DeepLabCut. This function computes the
16
17 Euclidean error between the manual labels and the ones predicted by DeepLabCut averaged over the
18
19 hand locations and test images (mean absolute error, proportional to the average root mean square error)
20
21 [20].
22
23
24
25

26 Signal processing

27
28
29
30
31 Video pixel coordinates for the labels produced by DeepLabCut were used to calculate the pixel distance
32
33 between index finger tip and thumb tip. A Savitzky-Golay filter was applied to the resulting time series,
34
35 which removed large, sudden transient label ‘jumps’ caused by DeepLabCut mislabelling (i.e. large-
36
37 distance, physiologically impossible label movement across one pair of video frames). The pixel
38
39 number distances were standardised across all videos, by using the maximum opening distance between
40
41 finger and thumb tip and normalising this to a value of 1 (all values were divided by the maximal value
42
43 in the corresponding time series).
44
45
46
47
48

49 Three features of the resulting finger tip to thumb tip distance time series were calculated to reflect the
50
51 clinical features of finger tapping speed, amplitude and rhythm (**Figure 1**). A measure of speed was
52
53 calculated as the mean rate of change of the normalised distance between finger and thumb tip over
54
55 time. A measure of amplitude variability was calculated by dividing each time series into one-second
56
57 windows, with maximal overlap, and then calculating the coefficient of variation of the mean difference
58
59 between the maximum and minimum amplitudes in each window. A measure of rhythm regularity was
60
61
62
63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

calculated by undertaking Fast Fourier Transform to find the distribution of frequencies within each finger tap time series, and then measuring the power of the dominant frequency peak added to the power of the frequencies 0.2 Hz either side of it. A more regular tapping rhythm concentrates power in a narrow frequency band, increasing the power of the dominant frequency peak (and its immediate neighbours), whereas a more irregular tapping rhythm leads to a more widely spread distribution of frequency bands, reducing the power of the dominant frequency peak (and its immediate neighbours).

Spearman correlation coefficients were calculated for the each of the three DeepLabCut measures vs the modal MBRS and MDS-UPDRS clinician ratings. In addition, we calculated the Spearman correlation coefficient of the three computer scores combined (the normalised arithmetic mean of the computer speed score, amplitude variability score and rhythm regularity score) with the MDS-UPDRS rating.

RESULTS

The participant and hand video characteristics are given in **Table 3**.

Tracking accuracy

DeepLabCut reliably tracked the finger tapping movements across 133 videoed clinical examinations. **Figure 2** shows example frames from a video labelled by DeepLabCut. The mean absolute error of DeepLabCut labelling was 8.39 pixels, i.e. the average distance between manual (human) labels and those predicted by DeepLabCut was 8.39 pixels within a 1920x1080 pixel video frame. **Video 1** shows examples of four videos labelled by DeepLabCut, including particularly challenging cases with tremor and dystonia also present.

Correlation with clinical bradykinesia ratings

1
2 The quantitative scores derived from DeepLabCut measurements correlated well with the group
3
4 neurologists' modal ratings of bradykinesia; **Figure 3**. The computer vision measures of tapping speed,
5
6 amplitude variability and rhythm regularity had good correlations with the respective MBRS clinical
7
8 ratings for speed (-0.70, $p < 0.001$), amplitude (0.65, $p < 0.001$) and rhythm (-0.61, $p < 0.001$). There
9
10 were also good correlations between the computer measures of tapping speed (-0.66, $p < 0.001$),
11
12 amplitude variability (0.56, $p < 0.001$) and rhythm regularity (-0.50, $p < 0.001$) and the MDS-UPDRS
13
14 clinical rating. The MDS-UPDRS is a composite clinical rating, combining the separate components of
15
16 bradykinesia, and the three computer measures combined also correlated with MDS-UPDRS (-0.69, p
17
18 < 0.001).

26 **DISCUSSION**

27
28
29
30
31 We have shown that new computer vision software, DeepLabCut, can track finger tapping in a standard
32
33 smartphone video recording of a conventional clinical examination, and quantify three measures related
34
35 to Parkinson's bradykinesia: speed, amplitude variation and rhythm regularity. We found good
36
37 correlations between the objective computer measures of finger tapping and clinical ratings of
38
39 bradykinesia by 22 movement disorder neurologists using two gold standard clinical scales (MBRS and
40
41 MDS-UPDRS). Thus, smartphone video measures using DeepLabCut exhibit convergent validity with
42
43 conventional clinician ratings.
44
45

46
47
48
49 These results are new and clinically significant. This is the first study to apply new open source
50
51 DeepLabCut technology, developed for animal behaviour studies, to the field of human movement
52
53 disorders, or indeed any aspect of clinical neurology. We demonstrate new objective measures related
54
55 to bradykinesia by bringing together two ubiquitous ingredients: the smartphone video camera and
56
57 visual artificial intelligence software. With the worldwide ownership of smartphones so high (e.g. 82%
58
59 in the UK, 77% in the US) [17], there are no hardware barriers to further use of this approach.
60
61
62

1
2 Neurologists recognize that quantification of bradykinesia will improve Parkinson's management
3
4 through 'precision medicine' principles: it has the potential to aid early accurate diagnosis and remote
5
6 monitoring, and provide fine grain regular measurement of clinical trial outcomes. Previous methods to
7
8 quantify bradykinesia have focussed on wearable sensors measuring finger tapping [9,12-14,23] but it
9
10 has been challenging to incorporate these into widespread use because equipment needs to be physically
11
12 attached to the patient. Furthermore, the necessity for sensors to stay fixed on a specific body part has
13
14 often required the exclusion of tremulous and dyskinetic patients [8,9,12,14,23]. In contrast, the
15
16 DeepLabCut test presented here is a contactless method that simply requires a short 10 second video
17
18 recording of the standard clinical examination. This means it can potentially be used in any setting and
19
20 any patient.
21
22
23
24
25

26
27 The 'real life' versatility of DeepLabCut in ordinary clinical settings was evaluated in this study because
28
29 no special efforts were made to optimize participant positioning or filming conditions. The tracking and
30
31 measures were robust despite variations in smartphone camera distance, positioning of the hand, skin
32
33 color, flexion / extension postures of digits 3 to 5, ambient lighting conditions, and inclusion of a small
34
35 number of patients with tremor or dystonia during tapping. **Video 1** illustrates this, with accurate
36
37 labelling of finger and thumb tip despite superimposed tremor or dystonia. This is particularly
38
39 encouraging for the relevance of DeepLabCut aiding assessment of Parkinson's in conventional clinical
40
41 settings.
42
43
44
45

46
47 Two further strengths of our study are the collection of clinical ratings from a large group of movement
48
49 disorder neurologists, and the use of two rating scales: the MBRS and MDS-UPDRS, such that the
50
51 clinical ratings were representative and reliable. Clinical rating scales are vulnerable to inter-rater
52
53 variability but using modal values derived from a large number of blinded clinical raters, and two
54
55 different gold-standard rating scales, provided a more robust 'ground truth' of clinical bradykinesia to
56
57 evaluate the new test against. As such we were able to demonstrate that clinician and DeepLabCut
58
59 measures correlate for both individual components of bradykinesia (MBRS) and an overall combination
60
61
62
63

1 of those components (MDS-UPDRS). An interesting future application with a larger participant group
2 would be to examine the relative contribution of the three video measures to individual grades of
3
4 UPDRS bradykinesia.
5

6
7
8 Another advantage of the approach described here is that it is not limited by the physical constraints of
9 wearable or screen/keyboard tapping methods, and could be applied to detect and measure other motor
10 features of Parkinson's in a contactless manner such as tremor, gait, or posture. Computer vision is
11 general has potential for application to these signs [24,25].
12
13
14
15
16
17

18
19 Only three previous studies have objectively measured finger tapping using contactless, standard video
20 analysis. One method extracted finger tapping information from 13 Parkinson's patients with advanced
21 disease [26], but it required inclusion of the patient's face in the video, limiting the practical utility. The
22 second report, from our own group, described a video method to measure movement using an optical
23 flow field [27] that interprets movements of the whole hand. The work here improves on this by
24 explicitly extracting the salient thumb and finger tip points, and validating in a larger cohort. A third
25 study of 60 Parkinson's patients (no controls) found strong correlation between MDS-UPDRS finger
26 tap rating and video measures of tap interval (frequency) ($r=0.91$), frequency variation ($r=0.82$) and
27 amplitude ($r=-0.94$) but not amplitude variation ($r=0.39$) [28]. The 'ground truth' of the clinical
28 comparison was less robust, based on just two clinical raters, and there were no MBRS ratings;
29 furthermore, some images were blurred in the 24 frame per second videos. A major advantage of our
30 method in comparison is that DeepLabCut is open source software, that can be downloaded and used
31 without the need to write computer code, making it available now to neurologists with only limited
32 computing knowledge.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52
53 A well-studied, non-camera method to measure finger tap bradykinesia is the 'Bradykinesia-Akinesia
54 Incoordination (BRAIN) Test', which involves participants tapping a standard computer keyboard for
55 30 seconds [29]. This shares advantages with our method in that it involves no special equipment, and
56 could be employed without clinician-patient contact. However, it requires patient motivation to engage
57
58
59
60
61
62
63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

with a keyboard tapping task, rather than a simple recording of existing standard clinical examination (and similar requirements for patient motivation apply to patient smartphone apps [15]). Furthermore, ‘BRAIN’ correlations with MDS-UPDRS finger tapping ratings of 0.44, 0.20, 0.33, 0.03 [29], are weaker than those we report (0.66, 0.56, 0.50, 0.69). One reason for this may be that a computer keyboard cannot record any measure of tap amplitude, and the speed measure is limited to tap frequency, so that some core aspects of bradykinesia cannot be captured by ‘BRAIN’.

The present study has some limitations. Like human vision, a simple camera lacks an absolute measure of distance, and our method cannot capture 3D movement. Rotation of the thumb and finger into a horizontal plane would falsely alter amplitude measurement. However, in practice such movement is rare during tapping, and our normalized measure of finger-thumb distance based on 2D maximum opening distance appeared to capture amplitude, speed and rhythm abnormalities well, given the good correlations with clinical ratings.

We have measured amplitude variation, but not amplitude decrement (the sequence effect). Previous approaches have measured decrement as a straight line linear regression [23], but decrement would also proportionately affect a measure of amplitude variation, such that the variations in decrement should be captured within our measure. Furthermore, there are several other previous reports of good correlation between non-decrement amplitude measures using sensors and clinician ratings of finger tap bradykinesia [9].

Relatively few videos were rated as grade 3 or 4 bradykinesia, and it is possible that a greater number of higher grade videos might change the strength of correlations. Even with accurate tracking, all forms of rating or measuring bradykinesia (including clinician) may potentially be confounded by comorbidities such as joint deformities, pain, dystonia etc. However, with a larger dataset, machine learning techniques are well-suited to separate such additional contributions from underlying bradykinesia. There was a degree of error in DeepLabCut labelling of finger and thumb tips, but this was minimal (mean of 8.39 pixels within a 1920x1080 pixel video frame) and unlikely to affect

1 individual measures. We mitigated against this by applying a filter prior to feature extraction. Full
2 flexion of digits 3 to 5 might improve labelling accuracy, but would also limit the amplitude of finger
3 tapping [23], potentially masking more subtle deficits. The labelling error provided by DeepLabCut's
4 'evaluate network' function only gives an accuracy figure for the entire set of videos but not for
5 individual videos, meaning that we have not reported accuracy in subsets of videos, such as those with
6 visible tremor. This study was exploratory, aimed at proof of concept that bradykinesia related measures
7 could be derived from smartphone video. The protocol was thus relatively simple. The test-retest
8 reliability of this technique has not been addressed in the study but we hope to do so in future. Similarly,
9 we did not measure 'off' and 'on' medication states (sensitivity to change).

10
11
12
13
14
15
16
17
18
19
20
21
22 Finally, we have demonstrated measurements that show convergent validity with clinical ratings, rather
23 than a test to discriminate patients from controls. In our view, it makes little sense to pursue automated
24 diagnostic classification using finger tapping video alone because clinical Parkinson's diagnosis is by
25 definition a broader assessment than isolated finger tapping performance [4].

26
27
28
29
30
31
32
33 This study has evaluated the clinical application of DeepLabCut in a medical condition, specifically a
34 neurological disorder. The methodology described here provides a new objective measure of
35 Parkinson's bradykinesia in conventional clinical examination. DeepLabCut provides a contactless
36 means for tracking and measuring bradykinesia in Parkinson's, without any special equipment other
37 than the ubiquitous smartphone. This research therefore provides a new tool to quantify bradykinesia.
38 It could potentially be used to support diagnosis and monitoring of Parkinson's, both clinically and as
39 an outcome measure in research studies. Beyond Parkinson's and other movement disorders, this same
40 technology could be applied to many other neurological conditions that require measurement of
41 movements or function; for example multiple sclerosis, neuropathies, epilepsy and stroke.

42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 **REFERENCES**

1. Dorsey ER, Bloem BR. The Parkinson Pandemic - A Call to Action. *JAMA Neurol.* 2018; 75(1): 9–10.
2. Delenclos M, Jones DR, McLean PJ, Uitti RJ. Biomarkers in Parkinson’s disease: Advances and strategies. *Park Relat Disord.* 2016; 22: S106-S110.
3. Burton A. How do we fix the shortage of neurologists?. *The Lancet Neurology.* 2018 Jun 1;17(6):502-3.
4. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson’s disease. *Mov Disord.* 2015;30(12):1591-1601. doi:10.1002/mds.26424
5. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* 2008; 23(15): 2129-70.
6. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology.* 2014; 83(5): 406-412.
7. Bajaj NPS, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study *Journal of Neurology, Neurosurgery & Psychiatry* 2010; 81: 1223-1228.
8. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson’s disease: Differential effects on bradykinesia. *Mov Disord.* 2007; 22(3): 328-333.
9. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson’s disease: Reliability and comparison with kinematic measures. *Mov Disord.* 2011; 26(10): 1859-1863.
10. Martinez-Martin P, Gil-Nagel A, Gracia LM, et al. Unified Parkinson’s disease rating scale characteristics and structure. *Mov Disord.* 1994; 9(1): 76-83.
11. Goetz CG, Stebbins GT. Assuring interrater reliability for the UPDRS motor section: Utility of the UPDRS teaching tape. *Mov Disord.* 2004; 19(12): 1453-1456.
12. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and amplitude of movement in Parkinson’s disease: A pilot study. *Mov Disord.* 2009; 24(7): 1001-1008.
13. Gao C, Smith S, Lones M, et al. Objective assessment of bradykinesia in Parkinson’s disease using evolutionary algorithms: Clinical validation. *Transl Neurodegener.* 2018; 7(1): 1-8.
14. Růžička E, Krupička R, Zárubová K, Ruzs J, Jech R, Szabó Z. Tests of manual dexterity and speed in Parkinson’s disease: Not all measure the same. *Park Relat Disord.* 2016; 28: 118-123.
15. Lee CY, Kang SJ, Hong S-K, Ma H-I, Lee U, Kim YJ. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson’s Disease. *PLoS One.* 2016;11(7).
16. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of blepharospasm severity. *Neurology.* 2016;87(20):2146-2153.
17. “Top Countries/Markets by Smartphone Penetration & Users.” Newzoo. Available from: <https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users/>. [Accessed 5 November 2019.]

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
18. Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci*. 2018; 21(9): 1281-1289.
 19. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-Dimensional Reconstruction of Paw and Digit Trajectories. *Journal of visualized experiments: JoVE*. 2019; Jul(149).
 20. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols*. 2019; 14(7): 2152-76.
 21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2016*; 770-778.
 22. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord*. 2004; 19(9): 1020-8.
 23. Ling H, Massey LA, Lees AJ, Brown P, Day BL. Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain*. 2012; 135(4): 1141-1153.
 24. Pinteá SL, Zheng J, Li X, Bank PJ, van Hilten JJ, van Gemert JC. Hand-tremor frequency estimation in videos. In *Proceedings of the European Conference on Computer Vision (ECCV) 2018*
 25. Khan T, Westin J, Dougherty M. Motion cue analysis for parkinsonian gait recognition. *The Open Biomedical Engineering Journal*. 2013; 7: 1.
 26. Khan T, Nyholm D, Westin J, Dougherty M. A computer vision framework for finger-tapping evaluation in Parkinson's disease. *Artif. Intell. Med*. 2014; 60(1): 27-40.
 27. Wong DC, Relton SD, Fang H, et al. Supervised Classification of Bradykinesia for Parkinson's Disease Diagnosis from Smartphone Videos. In *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*. 2019; 32-37.
 28. Liu Y, Chen J, Hu C, et al. Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia. *IEEE Trans Neural Syst Rehabil Eng*. 2019;27(10):1952-1961.
 29. Noyce AJ, Nagy A, Acharya S, Hadavi S, Bestwick JP, Fearnley J, Lees AJ, Giovannoni G. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLoS One*. 2014; 9 (4).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

FIGURE LEGENDS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1. Illustration of the three parameters derived from the DeepLabCut finger tip and thumb tip coordinates to measure finger tapping speed, amplitude and rhythm.

Figure 2. Example video frames taken from smartphone video labelled by DeepLabCut, showing the six localization (labelling) points: thumb tip (dark blue), thumb metacarpophalangeal (MCP) joint (light blue), index finger tip (cyan), index finger MCP joint (yellow), middle finger tip (orange), dorsal wrist (red).

Figure 3. Computer measures of finger tapping speed, amplitude variation and rhythm, derived from DeepLabCut smartphone video tracking, correlate with clinician mode ratings for the same videos. Pearson correlation coefficients shown. MBRS: Modified Bradykinesia Rating Scale. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale.

TABLES

Table 1. The Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) Item 3.4 (Finger Tapping).

MDS-UPDRS Grade	Description
0: Normal	No problems
1: Slight	Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.
2: Mild	Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.
3: Moderate	Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.
4: Severe	Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 taps are evaluated by the clinician [5].

Table 2. The Modified Bradykinesia Rating Scale (MBRS).

Score	Speed	Amplitude	Rhythm
0	Normal	Normal	Regular, no arrests or pauses in ongoing movement
1	Mild Slowing	Mild reduction in amplitude in later performance, most movements close to normal	Mild impairment, up to two brief arrests in the 10 seconds, none lasting > 1 second
2	Moderate Slowing	Moderate, reduction in amplitude visible early in performance but continues to maintain 50% amplitude through most of the tasks	Moderate, 3 to 4 arrests in 10 seconds; OR 1 or 2 lasting > 1second
3	Severe slowing	Severe, less than 50% amplitude through most of the task	Severe, 5 or more arrests/10 seconds; OR more than 2 lasting > 1 second
4	Can barely perform the task	Can barely perform the task	Can barely perform the task

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 seconds of tapping are evaluated by the clinician to give three MBRS ratings (one for each of speed, amplitude, rhythm) [6,7].

Table 3. Study participant (hand video) characteristics.

	Patients	Controls
Age (Std. Dev.) yrs	68 (9.6)	59 (19.4)
Male/Female	47/26	22/38
Median years since diagnosis	4	n/a
Median H&Y [IQR]	2 [1,3]	n/a
H&Y = 1	32	
H&Y = 1.5	2	
H&Y = 2	12	
H&Y = 2.5	4	
H&Y = 3	19	
H&Y = 4	4	

H&Y = 5	0	
Median MDS-UPDRS [IQR]	2 [1,3]	1 [0,1]
MDS-UPDRS = 0	9	23
MDS-UPDRS = 1	20	23
MDS-UPDRS = 2	19	12
MDS-UPDRS = 3	20	2
MDS-UPDRS = 4	5	0
Visible tremor in video	11	0
Visible dystonia in video	2	0

Data is split by Parkinson's hands (n=73) and control hands (n=60). H&Y: modified Hoehn and Yahr scale [22]. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale, Item 3.4 (Finger Tapping). IQR: Interquartile Range.

Acknowledgements

We would like to thank the following UK movement disorder neurologists for providing clinical rating of videos: Dr Jane Alty, Dr Sundus Alusi, Dr Michael Bonello, Dr Stephen Butterworth, Dr Philip Buttery, Dr Camille Carroll, Dr Adam Cassidy, Dr Jeremy Cosgrove, Dr Richard Ellis, Dr Jonathan Evans, Dr Paul Goldsmith, Dr Donald Grosset, Dr Christopher Kobylecki, Dr Alistair Lansbury, Dr Peter Moore, Dr Rachel Newby, Dr Edward Newman, Dr Gillian Sare, Dr Monty Silverdale, Dr Naomi Warren, Dr Louise Wiblin, Dr Caroline Williams-Gray.

We would also like to thank Dr Parisa Patel and Dr Paschal O’Gorman for data collection.

Author contributions

Name	Location	Role	Contribution
-------------	-----------------	-------------	---------------------

Stefan Williams, MBBS, BSc	University of Leeds, UK	Author	Design and conceptualized study; data acquisition and interpretation; drafted the manuscript for intellectual content
Zhibin Zhao, BEng	(1) University of Manchester, UK; (2) Xi'an Jiatong University, China	Author	Analyzed the data (major role)
Awais Hafeez, MA, BSc	(1) University of Leeds, UK; (2) University of Engineering and Technology Lahore, Pakistan	Author	Design and conceptualized the study; major role in acquisition of data
David C Wong, PhD	University of Manchester, UK	Author	Design and conceptualized study; analyzed the data; data interpretation; revised the manuscript for intellectual content
Samuel D Relton, PhD	University of Leeds, UK	Author	Design and conceptualized study; data interpretation; revised the manuscript for intellectual content
Hui Fang, PhD	Loughborough University, UK	Author	Design and conceptualized study (major role)
Jane E Alty, MD, MA, MBBChir	(1) University of Tasmania, Australia; (2) Leeds Teaching Hospitals NHS Trust, UK	Author	Design and conceptualized study; data interpretation; drafted the manuscript for intellectual content

Title Page

The discerning eye of computer vision: can it measure Parkinson's finger tap bradykinesia ?

Authors: Stefan Williams, MBBS, BSc; Zhibin Zhao, BEng; Awais Hafeez, MS, BSc; David C Wong, PhD; Samuel D Relton, PhD; Hui Fang, PhD; Jane E Alty MD, MA, MBBChir.

Stefan Williams, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Zhibin Zhao, (1) University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK; (2) Xi'an Jiatong University, School of Mechanical Engineering, Xi'an, China.

Awais Hafeez, (1) University of Leeds, School of Mechanical Engineering, Leeds, UK; (2) University of Engineering and Technology Lahore, Department of Mechatronics and Control Engineering, Lahore, Pakistan

David C Wong, University of Manchester, Division of Informatics, Imaging & Data Sciences, Manchester, UK

Samuel D Relton, University of Leeds, Leeds Institute of Health Science, Leeds, UK

Hui Fang, Loughborough University, Department of Computer Science, Loughborough, UK

Jane E Alty, (1) University of Tasmania, Wicking Dementia Research & Education Centre, Hobart, Australia (2) Leeds Teaching Hospitals NHS Trust UK.

Word count: **3099**

Number of Tables: **3**

Number of Figures: **3**

Number of Videos: **1**

Corresponding Author:

Stefan Williams

Leeds Institute of Health Science, 10th Floor, Worsley Building

University of Leeds, Clarendon Way, Leeds, UK LS2 9LU

Phone: 0113 343 0839

umswi@leeds.ac.uk

ORCID: 0000-0002-4786-1356

Keywords:

Parkinson's disease; Parkinsonism; Bradykinesia; Finger tapping; Computer Vision; Artificial Intelligence; DeepLabCut

1 **ABSTRACT**

6 **Objective**

8 The worldwide prevalence of Parkinson's disease is increasing. There is urgent need for new tools to objectively
9 measure the condition. Existing methods to record the cardinal motor feature of the condition, bradykinesia, using
10 wearable sensors or smartphone apps have not reached large-scale, routine use. We evaluate new computer vision
11 (artificial intelligence) technology, DeepLabCut, as a contactless method to quantify measures related to
12 Parkinson's bradykinesia from smartphone videos of finger tapping.
13
14
15
16
17
18
19

20 **Methods**

21 Standard smartphone video recordings of 133 hands performing finger tapping (39 idiopathic Parkinson's patients
22 and 30 controls) were tracked on a frame-by-frame basis with DeepLabCut. Objective computer measures of
23 tapping speed, amplitude and rhythm were correlated with clinical ratings made by 22 movement disorder
24 neurologists using the Modified Bradykinesia Rating Scale (MBRS) and Movement Disorder Society revision of
25 the Unified Parkinson's Disease Rating Scale (MDS-UPDRS).
26
27
28
29
30
31
32
33

34 **Results**

35 DeepLabCut reliably tracked and measured finger tapping in standard smartphone video. Computer measures
36 correlated well with clinical ratings of bradykinesia (Spearman coefficients): -0.74 speed, 0.66 amplitude, -0.65
37 rhythm for MBRS; -0.56 speed, 0.61 amplitude, -0.50 rhythm for MDS-UPDRS; 0.69 combined for MDS-
38 UPDRS. All $p < 0.001$.
39
40
41
42
43
44
45
46

47 **Conclusion**

48 New computer vision software, DeepLabCut, can quantify three measures related to Parkinson's bradykinesia
49 from smartphone videos of finger tapping. Objective 'contactless' measures of standard clinical examinations
50 were not previously possible with wearable sensors (accelerometers, gyroscopes, infrared markers). DeepLabCut
51 requires only conventional video recording of clinical examination and is entirely 'contactless'. This next
52 generation technology holds potential for Parkinson's and other neurological disorders with altered movements.
53
54
55
56
57
58
59
60
61
62
63
64
65

INTRODUCTION

The worldwide prevalence of Parkinson's is increasing, with an estimated 10 million people already affected [1]. There is urgent need to find new tools to objectively measure the condition, to assist diagnosis, monitoring, and research metrics [2]. This is especially pressing in the era of precision medicine and with a global shortage of neurologists [3].

The cornerstone of Parkinson's assessment is the clinician's visual judgement of bradykinesia: one of the cardinal motor features of the condition, which is defined as slowness of movement and decrement in size or speed, or progressive hesitations/halts, as movements are continued [4]. There are several clinical examination methods to assess bradykinesia, but in a brief clinic room assessment, one of the most common is to observe the patient repeatedly tapping their index finger against thumb 'as quickly and as big as possible' [5]. However, clinicians' visual interpretation is inherently subjective [6,7], with considerable inter-rater variability [5,9-11]. Previous reported methods to measure bradykinesia on finger tapping have required wearable equipment such as gyroscopes [9], electromagnetic sensors [12; 13], and infrared camera markers [14], or patient interaction with an app-based task [15], and this has hindered the pathway to large-scale, routine clinical use.

Recent developments in computer interpretation of video (computer vision), clear the pathway to 'contactless' methods of quantifying clinical assessments [16], with potential widespread global application because of the availability of inexpensive smartphone cameras [17]. DeepLabCut is recently published open source computer vision software, without a requirement to write new computer code, that has been demonstrated to track and measure mice and fruit fly body parts [18,19]. We apply it to clinical neurology for the first time and assess whether the technology can be used to track and measure bradykinesia in standard smartphone video of conventional finger tapping examination.

METHODS

1
2 Ethical review
3
4
5

6 The study was approved by the North of Scotland Research Ethics Committee, United Kingdom Health
7
8 Research Authority (IRAS project ID 256116).
9

10
11
12
13 Participants and video recording
14
15

16
17 Video recordings of 138 hands (including left and right) were obtained from 39 patients with idiopathic
18 Parkinson's disease and 30 healthy controls, who gave written consent. One video was rejected because
19
20 the hand moved outside the video frame, making 137 videos: 77 Parkinson's hands and 60 control
21
22 hands.
23
24
25

26
27
28 All Parkinson's patients had previously been diagnosed by a movement disorder specialist neurologist
29 at Leeds Teaching Hospitals NHS Trust United Kingdom, according to Movement Disorder Society
30 clinical diagnostic criteria [4], and were subjectively and objectively in the 'on' state at the time of
31 video recording (no medication was withheld prior to recording). We did not exclude patients with
32 postural hand tremor or dystonia. Controls were recruited from the companions of patients, or
33 hospital/university staff, and had no history of Parkinson's disease or other neurological diagnosis.
34
35
36
37
38
39
40
41
42
43

44 Participants rested their elbow on a chair arm with the forearm lifted to a 45 degree angle. The hand
45 was free to move as per the protocol of the Movement Disorder Society revision of the Unified
46 Parkinson's Disease Rating Scale (MDS-UPDRS) Item 3.4 Finger Tapping [5]. Only the hand/forearm
47 was within the video frame. The distance from camera to hand was approximately 1m, but not tightly
48 defined. Digits 1 and 2 were closest to the camera. No specific instructions were given for the position
49 of digits 3 to 5, and participants were free to position these digits as they preferred, although the
50 researcher gave a brief demonstration in which digits 3 to 5 were extended.
51
52
53
54
55
56
57
58
59
60
61
62
63

1 A smartphone, placed on a tripod, was used to record standard video (60 frames per second, 1920x1080
2 pixel resolution), with only ambient lighting. Participants were instructed to tap their index finger and
3 thumb together “as quickly and as big as possible” for at least 10 seconds [5]. Each video was edited to
4 an 11 second clip: 1 second prior to tapping onset and 10 seconds of finger tapping.
5
6
7
8
9

10 Clinical rating

11
12
13
14
15 22 consultant neurologists specialising in movement disorders (United Kingdom) were asked to rate 30
16 tapping videos each, selected at random from the set of 137 videos. The median number of raters per
17 video was 5 (range 1 to 12, interquartile range 3 to 7). The raters were blinded to patient / healthy
18 control status and each other’s scores. Each video was rated according to both the MDS-UPDRS Item
19 3.4 Finger Tapping [5], and the Modified Bradykinesia Rating Scale (MBRS) [8,9]. The MDS-UPDRS
20 requires the rater to amalgamate judgments of finger tapping speed, amplitude, and rhythm into a single
21 composite score (**Table 1**). In contrast, the MBRS is comprised of three separate scores for speed,
22 amplitude, and rhythm (**Table 2**). The modal clinician rating for each video was used for correlation
23 with computer measures (see below). For each possible pair of raters, Cohen Kappa was calculated
24 based on all videos they had both seen. The resulting values had mean 0.28 (fair agreement) and
25 standard deviation 0.28.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 DeepLabCut video tracking

43
44
45
46 DeepLabCut tracks the geometrical configuration of multiple body parts in video, without a requirement
47 to wear markers [18,20]. It uses transfer learning, with feature detectors based on deep neural networks
48 that have been pretrained on ImageNet, a massive object recognition dataset [21]. This means that
49 DeepLabCut is able to accurately recognise and track body parts with minimal training data [18].
50
51
52
53
54
55
56
57

58 The set of 137 finger tapping videos was processed by DeepLabCut. We localised (labelled) six distinct
59 points on the tapping hand in 20 frames selected by k-means clustering from each 660 frame video:
60
61
62
63
64
65

1 thumb tip; index finger tip; thumb metacarpophalangeal (MCP) joint; index finger MCP joint; middle
2 finger tip; dorsal wrist/proximal dorsal hand. The deep neural network architecture of DeepLabCut was
3
4 then trained using these points to predict their localisation in the remaining 97% of (unlabelled) video
5
6 frames (1030000 training iterations, ResNet 50). This created video pixel coordinates for finger tip and
7
8 thumb tip throughout each video.
9

10
11
12
13 The accuracy of DeepLabCut to track the hand localisation points (including finger tip and thumb tip)
14
15 was assessed using the ‘evaluate network’ function within DeepLabCut. This function computes the
16
17 Euclidean error between the manual labels and the ones predicted by DeepLabCut averaged over the
18
19 hand locations and test images (mean absolute error, proportional to the average root mean square error)
20
21 [20].
22
23
24
25

26 Signal processing

27
28
29
30
31 Video pixel coordinates for the labels produced by DeepLabCut were used to calculate the pixel distance
32
33 between index finger tip and thumb tip. A Savitzky-Golay filter was applied to the resulting time series,
34
35 which removed large, sudden transient label ‘jumps’ caused by DeepLabCut mislabelling (i.e. large-
36
37 distance, physiologically impossible label movement across one pair of video frames). The pixel
38
39 number distances were standardised across all videos, by using the maximum opening distance between
40
41 finger and thumb tip and normalising this to a value of 1 (all values were divided by the maximal value
42
43 in the corresponding time series).
44
45
46
47
48

49 Three features of the resulting finger tip to thumb tip distance time series were calculated to reflect the
50
51 clinical features of finger tapping speed, amplitude and rhythm (**Figure 1**). A measure of speed was
52
53 calculated as the mean rate of change of the normalised distance between finger and thumb tip over
54
55 time. A measure of amplitude variability was calculated by dividing each time series into one-second
56
57 windows, with maximal overlap, and then calculating the coefficient of variation of the mean difference
58
59 between the maximum and minimum amplitudes in each window. A measure of rhythm regularity was
60
61
62
63

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

calculated by undertaking Fast Fourier Transform to find the distribution of frequencies within each finger tap time series, and then measuring the power of the dominant frequency peak added to the power of the frequencies 0.2 Hz either side of it. A more regular tapping rhythm concentrates power in a narrow frequency band, increasing the power of the dominant frequency peak (and its immediate neighbours), whereas a more irregular tapping rhythm leads to a more widely spread distribution of frequency bands, reducing the power of the dominant frequency peak (and its immediate neighbours).

Spearman correlation coefficients were calculated for the each of the three DeepLabCut measures vs the modal MBRS and MDS-UPDRS clinician ratings. In addition, we calculated the Spearman correlation coefficient of the three computer scores combined (the normalised arithmetic mean of the computer speed score, amplitude variability score and rhythm regularity score) with the MDS-UPDRS rating.

RESULTS

The participant and hand video characteristics are given in **Table 3**.

Tracking accuracy

DeepLabCut reliably tracked the finger tapping movements across 133 videoed clinical examinations. **Figure 2** shows example frames from a video labelled by DeepLabCut. The mean absolute error of DeepLabCut labelling was 8.39 pixels, i.e. the average distance between manual (human) labels and those predicted by DeepLabCut was 8.39 pixels within a 1920x1080 pixel video frame. **Video 1** shows examples of four videos labelled by DeepLabCut, including particularly challenging cases with tremor and dystonia also present.

Correlation with clinical bradykinesia ratings

1
2 The quantitative scores derived from DeepLabCut measurements correlated well with the group
3
4 neurologists' modal ratings of bradykinesia; **Figure 3**. The computer vision measures of tapping speed,
5
6 amplitude variability and rhythm regularity had good correlations with the respective MBRS clinical
7
8 ratings for speed (-0.70, $p < 0.001$), amplitude (0.65, $p < 0.001$) and rhythm (-0.61, $p < 0.001$). There
9
10 were also good correlations between the computer measures of tapping speed (-0.66, $p < 0.001$),
11
12 amplitude variability (0.56, $p < 0.001$) and rhythm regularity (-0.50, $p < 0.001$) and the MDS-UPDRS
13
14 clinical rating. The MDS-UPDRS is a composite clinical rating, combining the separate components of
15
16 bradykinesia, and the three computer measures combined also correlated with MDS-UPDRS (-0.69, p
17
18 < 0.001).

21 22 23 24 25 26 **DISCUSSION**

27
28
29
30
31 We have shown that new computer vision software, DeepLabCut, can track finger tapping in a standard
32
33 smartphone video recording of a conventional clinical examination, and quantify three measures related
34
35 to Parkinson's bradykinesia: speed, amplitude variation and rhythm regularity. We found good
36
37 correlations between the objective computer measures of finger tapping and clinical ratings of
38
39 bradykinesia by 22 movement disorder neurologists using two gold standard clinical scales (MBRS and
40
41 MDS-UPDRS). Thus, smartphone video measures using DeepLabCut exhibit convergent validity with
42
43 conventional clinician ratings.
44
45
46
47
48

49 These results are new and clinically significant. This is the first study to apply new open source
50
51 DeepLabCut technology, developed for animal behaviour studies, to the field of human movement
52
53 disorders, or indeed any aspect of clinical neurology. We demonstrate new objective measures related
54
55 to bradykinesia by bringing together two ubiquitous ingredients: the smartphone video camera and
56
57 visual artificial intelligence software. With the worldwide ownership of smartphones so high (e.g. 82%
58
59 in the UK, 77% in the US) [17], there are no hardware barriers to further use of this approach.
60
61
62

1
2 Neurologists recognize that quantification of bradykinesia will improve Parkinson's management
3
4 through 'precision medicine' principles: it has the potential to aid early accurate diagnosis and remote
5
6 monitoring, and provide fine grain regular measurement of clinical trial outcomes. Previous methods to
7
8 quantify bradykinesia have focussed on wearable sensors measuring finger tapping [9,12-14,23] but it
9
10 has been challenging to incorporate these into widespread use because equipment needs to be physically
11
12 attached to the patient. Furthermore, the necessity for sensors to stay fixed on a specific body part has
13
14 often required the exclusion of tremulous and dyskinetic patients [8,9,12,14,23]. In contrast, the
15
16 DeepLabCut test presented here is a contactless method that simply requires a short 10 second video
17
18 recording of the standard clinical examination. This means it can potentially be used in any setting and
19
20 any patient.
21
22
23
24
25

26
27 The 'real life' versatility of DeepLabCut in ordinary clinical settings was evaluated in this study because
28
29 no special efforts were made to optimize participant positioning or filming conditions. The tracking and
30
31 measures were robust despite variations in smartphone camera distance, positioning of the hand, skin
32
33 color, flexion / extension postures of digits 3 to 5, ambient lighting conditions, and inclusion of a small
34
35 number of patients with tremor or dystonia during tapping. **Video 1** illustrates this, with accurate
36
37 labelling of finger and thumb tip despite superimposed tremor or dystonia. This is particularly
38
39 encouraging for the relevance of DeepLabCut aiding assessment of Parkinson's in conventional clinical
40
41 settings.
42
43
44
45

46
47 Two further strengths of our study are the collection of clinical ratings from a large group of movement
48
49 disorder neurologists, and the use of two rating scales: the MBRS and MDS-UPDRS, such that the
50
51 clinical ratings were representative and reliable. Clinical rating scales are vulnerable to inter-rater
52
53 variability but using modal values derived from a large number of blinded clinical raters, and two
54
55 different gold-standard rating scales, provided a more robust 'ground truth' of clinical bradykinesia to
56
57 evaluate the new test against. As such we were able to demonstrate that clinician and DeepLabCut
58
59 measures correlate for both individual components of bradykinesia (MBRS) and an overall combination
60
61
62
63

1 of those components (MDS-UPDRS). An interesting future application with a larger participant group
2 would be to examine the relative contribution of the three video measures to individual grades of
3
4 UPDRS bradykinesia.
5
6
7

8 Another advantage of the approach described here is that it is not limited by the physical constraints of
9 wearable or screen/keyboard tapping methods, and could be applied to detect and measure other motor
10 features of Parkinson's in a contactless manner such as tremor, gait, or posture. Computer vision is
11 general has potential for application to these signs [24,25].
12
13
14
15
16
17
18
19

20 Only three previous studies have objectively measured finger tapping using contactless, standard video
21 analysis. One method extracted finger tapping information from 13 Parkinson's patients with advanced
22 disease [26], but it required inclusion of the patient's face in the video, limiting the practical utility. The
23 second report, from our own group, described a video method to measure movement using an optical
24 flow field [27] that interprets movements of the whole hand. The work here improves on this by
25 explicitly extracting the salient thumb and finger tip points, and validating in a larger cohort. A third
26 study of 60 Parkinson's patients (no controls) found strong correlation between MDS-UPDRS finger
27 tap rating and video measures of tap interval (frequency) ($r=0.91$), frequency variation ($r=0.82$) and
28 amplitude ($r=-0.94$) but not amplitude variation ($r=0.39$) [28]. The 'ground truth' of the clinical
29 comparison was less robust, based on just two clinical raters, and there were no MBRS ratings;
30 furthermore, some images were blurred in the 24 frame per second videos. A major advantage of our
31 method in comparison is that DeepLabCut is open source software, that can be downloaded and used
32 without the need to write computer code, making it available now to neurologists with only limited
33 computing knowledge.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53 A well-studied, non-camera method to measure finger tap bradykinesia is the 'Bradykinesia-Akinesia
54 Incoordination (BRAIN) Test', which involves participants tapping a standard computer keyboard for
55 30 seconds [29]. This shares advantages with our method in that it involves no special equipment, and
56 could be employed without clinician-patient contact. However, it requires patient motivation to engage
57
58
59
60
61
62
63
64
65

1
2 with a keyboard tapping task, rather than a simple recording of existing standard clinical examination
3 (and similar requirements for patient motivation apply to patient smartphone apps [15]). Furthermore,
4 'BRAIN' correlations with MDS-UPDRS finger tapping ratings of 0.44, 0.20, 0.33, 0.03 [29], are
5 weaker than those we report (0.66, 0.56, 0.50, 0.69). One reason for this may be that a computer
6 keyboard cannot record any measure of tap amplitude, and the speed measure is limited to tap
7 frequency, so that some core aspects of bradykinesia cannot be captured by 'BRAIN'.
8
9
10
11
12
13
14

15 The present study has some limitations. Like human vision, a simple camera lacks an absolute measure
16 of distance, and our method cannot capture 3D movement. Rotation of the thumb and finger into a
17 horizontal plane would falsely alter amplitude measurement. However, in practice such movement is
18 rare during tapping, and our normalized measure of finger-thumb distance based on 2D maximum
19 opening distance appeared to capture amplitude, speed and rhythm abnormalities well, given the good
20 correlations with clinical ratings.
21
22
23
24
25
26
27
28
29
30

31 We have measured amplitude variation, but not amplitude decrement (the sequence effect). Previous
32 approaches have measured decrement as a straight line linear regression [23], but decrement would also
33 proportionately affect a measure of amplitude variation, such that the variations in decrement should be
34 captured within our measure. Furthermore, there are several other previous reports of good correlation
35 between non-decrement amplitude measures using sensors and clinician ratings of finger tap
36 bradykinesia [9].
37
38
39
40
41
42
43
44
45
46

47 Relatively few videos were rated as grade 3 or 4 bradykinesia, and it is possible that a greater number
48 of higher grade videos might change the strength of correlations. Even with accurate tracking, all forms
49 of rating or measuring bradykinesia (including clinician) may potentially be confounded by
50 comorbidities such as joint deformities, pain, dystonia etc. However, with a larger dataset, machine
51 learning techniques are well-suited to separate such additional contributions from underlying
52 bradykinesia. There was a degree of error in DeepLabCut labelling of finger and thumb tips, but this
53 was minimal (mean of 8.39 pixels within a 1920x1080 pixel video frame) and unlikely to affect
54
55
56
57
58
59
60
61
62
63

1 individual measures. We mitigated against this by applying a filter prior to feature extraction. Full
2 flexion of digits 3 to 5 might improve labelling accuracy, but would also limit the amplitude of finger
3 tapping [23], potentially masking more subtle deficits. The labelling error provided by DeepLabCut's
4 'evaluate network' function only gives an accuracy figure for the entire set of videos but not for
5 individual videos, meaning that we have not reported accuracy in subsets of videos, such as those with
6 visible tremor. This study was exploratory, aimed at proof of concept that bradykinesia related measures
7 could be derived from smartphone video. The protocol was thus simple, and repeat measures such as
8 test-retest in controls, or 'off' and 'on' medication states, were not included, meaning that reliability
9 and sensitivity to change have not been characterised. Future work will assess this.
10
11
12
13
14
15
16
17
18
19
20
21

22 Finally, we have demonstrated measurements that show convergent validity with clinical ratings, rather
23 than a test to discriminate patients from controls. In our view, it makes little sense to pursue automated
24 diagnostic classification using finger tapping video alone because clinical Parkinson's diagnosis is by
25 definition a broader assessment than isolated finger tapping performance [4].
26
27
28
29
30
31

32
33 This study has evaluated the clinical application of DeepLabCut in a medical condition, specifically a
34 neurological disorder. The methodology described here provides a new objective measure of
35 Parkinson's bradykinesia in conventional clinical examination. DeepLabCut provides a contactless
36 means for tracking and measuring bradykinesia in Parkinson's, without any special equipment other
37 than the ubiquitous smartphone. This research therefore provides a new tool to quantify bradykinesia.
38 It could potentially be used to support diagnosis and monitoring of Parkinson's, both clinically and as
39 an outcome measure in research studies. Beyond Parkinson's and other movement disorders, this same
40 technology could be applied to many other neurological conditions that require measurement of
41 movements or function; for example multiple sclerosis, neuropathies, epilepsy and stroke.
42
43
44
45
46
47
48
49
50
51
52
53

54 **REFERENCES**

55
56
57
58
59
60
61
62
63
64
65

1. Dorsey ER, Bloem BR. The Parkinson Pandemic - A Call to Action. *JAMA Neurol.* 2018; 75(1): 9–10.
2. Delenclos M, Jones DR, McLean PJ, Uitti RJ. Biomarkers in Parkinson’s disease: Advances and strategies. *Park Relat Disord.* 2016; 22: S106-S110.
3. Burton A. How do we fix the shortage of neurologists?. *The Lancet Neurology.* 2018 Jun 1;17(6):502-3.
4. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson’s disease. *Mov Disord.* 2015;30(12):1591-1601. doi:10.1002/mds.26424
5. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord.* 2008; 23(15): 2129-70.
6. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology.* 2014; 83(5): 406-412.
7. Bajaj NPS, Gontu V, Birchall J, et al. Accuracy of clinical diagnosis in tremulous parkinsonian patients: a blinded video study *Journal of Neurology, Neurosurgery & Psychiatry* 2010; 81: 1223-1228.
8. Kishore A, Espay AJ, Marras C, et al. Unilateral versus bilateral tasks in early asymmetric Parkinson’s disease: Differential effects on bradykinesia. *Mov Disord.* 2007; 22(3): 328-333.
9. Heldman DA, Giuffrida JP, Chen R, et al. The modified bradykinesia rating scale for Parkinson’s disease: Reliability and comparison with kinematic measures. *Mov Disord.* 2011; 26(10): 1859-1863.
10. Martinez-Martin P, Gil-Nagel A, Gracia LM, et al. Unified Parkinson’s disease rating scale characteristics and structure. *Mov Disord.* 1994; 9(1): 76-83.
11. Goetz CG, Stebbins GT. Assuring interrater reliability for the UPDRS motor section: Utility of the UPDRS teaching tape. *Mov Disord.* 2004; 19(12): 1453-1456.
12. Espay AJ, Beaton DE, Morgante F, Gunraj CA, Lang AE, Chen R. Impairments of speed and amplitude of movement in Parkinson’s disease: A pilot study. *Mov Disord.* 2009; 24(7): 1001-1008.
13. Gao C, Smith S, Lones M, et al. Objective assessment of bradykinesia in Parkinson’s disease using evolutionary algorithms: Clinical validation. *Transl Neurodegener.* 2018; 7(1): 1-8.
14. Růžička E, Krupička R, Zárubová K, Ruzs J, Jech R, Szabó Z. Tests of manual dexterity and speed in Parkinson’s disease: Not all measure the same. *Park Relat Disord.* 2016; 28: 118-123.
15. Lee CY, Kang SJ, Hong S-K, Ma H-I, Lee U, Kim YJ. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson’s Disease. *PLoS One.* 2016;11(7).
16. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of blepharospasm severity. *Neurology.* 2016;87(20):2146-2153.
17. “Top Countries/Markets by Smartphone Penetration & Users.” Newzoo. Available from: <https://newzoo.com/insights/rankings/top-countries-by-smartphone-penetration-and-users/>. [Accessed 5 November 2019.]

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
18. Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci.* 2018; 21(9): 1281-1289.
 19. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-Dimensional Reconstruction of Paw and Digit Trajectories. *Journal of visualized experiments: JoVE.* 2019; Jul(149).
 20. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols.* 2019; 14(7): 2152-76.
 21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2016*; 770-778.
 22. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* 2004; 19(9): 1020-8.
 23. Ling H, Massey LA, Lees AJ, Brown P, Day BL. Hypokinesia without decrement distinguishes progressive supranuclear palsy from Parkinson's disease. *Brain.* 2012; 135(4): 1141-1153.
 24. Pinteá SL, Zheng J, Li X, Bank PJ, van Hilten JJ, van Gemert JC. Hand-tremor frequency estimation in videos. In *Proceedings of the European Conference on Computer Vision (ECCV) 2018*
 25. Khan T, Westin J, Dougherty M. Motion cue analysis for parkinsonian gait recognition. *The Open Biomedical Engineering Journal.* 2013; 7: 1.
 26. Khan T, Nyholm D, Westin J, Dougherty M. A computer vision framework for finger-tapping evaluation in Parkinson's disease. *Artif. Intell. Med.* 2014; 60(1): 27-40.
 27. Wong DC, Relton SD, Fang H, et al. Supervised Classification of Bradykinesia for Parkinson's Disease Diagnosis from Smartphone Videos. In *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS).* 2019; 32-37.
 28. Liu Y, Chen J, Hu C, et al. Vision-Based Method for Automatic Quantification of Parkinsonian Bradykinesia. *IEEE Trans Neural Syst Rehabil Eng.* 2019;27(10):1952-1961.
 29. Noyce AJ, Nagy A, Acharya S, Hadavi S, Bestwick JP, Fearnley J, Lees AJ, Giovannoni G. Bradykinesia-akinesia incoordination test: validating an online keyboard test of upper limb function. *PLoS One.* 2014; 9 (4).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

FIGURE LEGENDS

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1. Illustration of the three parameters derived from the DeepLabCut finger tip and thumb tip coordinates to measure finger tapping speed, amplitude and rhythm.

Figure 2. Example video frames taken from smartphone video labelled by DeepLabCut, showing the six localization (labelling) points: thumb tip (dark blue), thumb metacarpophalangeal (MCP) joint (light blue), index finger tip (cyan), index finger MCP joint (yellow), middle finger tip (orange), dorsal wrist (red).

Figure 3. Computer measures of finger tapping speed, amplitude variation and rhythm, derived from DeepLabCut smartphone video tracking, correlate with clinician mode ratings for the same videos. Pearson correlation coefficients shown. MBRS: Modified Bradykinesia Rating Scale. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale.

TABLES

Table 1. The Movement Disorder Society revision of the Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) Item 3.4 (Finger Tapping).

MDS-UPDRS Grade	Description
0: Normal	No problems
1: Slight	Any of the following: a) the regular rhythm is broken with one or two interruptions or hesitations of the tapping movement; b) slight slowing; c) the amplitude decrements near the end of the 10 taps.
2: Mild	Any of the following: a) 3 to 5 interruptions during tapping; b) mild slowing; c) the amplitude decrements midway in the 10-tap sequence.
3: Moderate	Any of the following: a) more than 5 interruptions during tapping or at least one longer arrest (freeze) in ongoing movement; b) moderate slowing; c) the amplitude decrements starting after the 1st tap.
4: Severe	Cannot or can only barely perform the task because of slowing, interruptions, or decrements.

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 taps are evaluated by the clinician [5].

Table 2. The Modified Bradykinesia Rating Scale (MBRS).

Score	Speed	Amplitude	Rhythm
0	Normal	Normal	Regular, no arrests or pauses in ongoing movement
1	Mild Slowing	Mild reduction in amplitude in later performance, most movements close to normal	Mild impairment, up to two brief arrests in the 10 seconds, none lasting > 1 second
2	Moderate Slowing	Moderate, reduction in amplitude visible early in performance but continues to maintain 50% amplitude through most of the tasks	Moderate, 3 to 4 arrests in 10 seconds; OR 1 or 2 lasting > 1second
3	Severe slowing	Severe, less than 50% amplitude through most of the task	Severe, 5 or more arrests/10 seconds; OR more than 2 lasting > 1 second
4	Can barely perform the task	Can barely perform the task	Can barely perform the task

Each hand is tested separately. The patient is instructed to tap the index finger on the thumb “as quickly and as big as possible”. The first 10 seconds of tapping are evaluated by the clinician to give three MBRS ratings (one for each of speed, amplitude, rhythm) [6,7].

Table 3. Study participant (hand video) characteristics.

	Patients	Controls
Age (Std. Dev.) yrs	68 (9.6)	59 (19.4)
Male/Female	47/26	22/38
Median years since diagnosis	4	n/a
Median H&Y [IQR]	2 [1,3]	n/a
H&Y = 1	32	
H&Y = 1.5	2	
H&Y = 2	12	
H&Y = 2.5	4	
H&Y = 3	19	
H&Y = 4	4	

H&Y = 5	0	
Median MDS-UPDRS [IQR]	2 [1,3]	1 [0,1]
MDS-UPDRS = 0	9	23
MDS-UPDRS = 1	20	23
MDS-UPDRS = 2	19	12
MDS-UPDRS = 3	20	2
MDS-UPDRS = 4	5	0
Visible tremor in video	11	0
Visible dystonia in video	2	0

Data is split by Parkinson's hands (n=73) and control hands (n=60). H&Y: modified Hoehn and Yahr scale [22]. MDS-UPDRS: Movement Disorder Society revision of the Unified Parkinson's Disease Rating Scale, Item 3.4 (Finger Tapping). IQR: Interquartile Range.

Acknowledgements

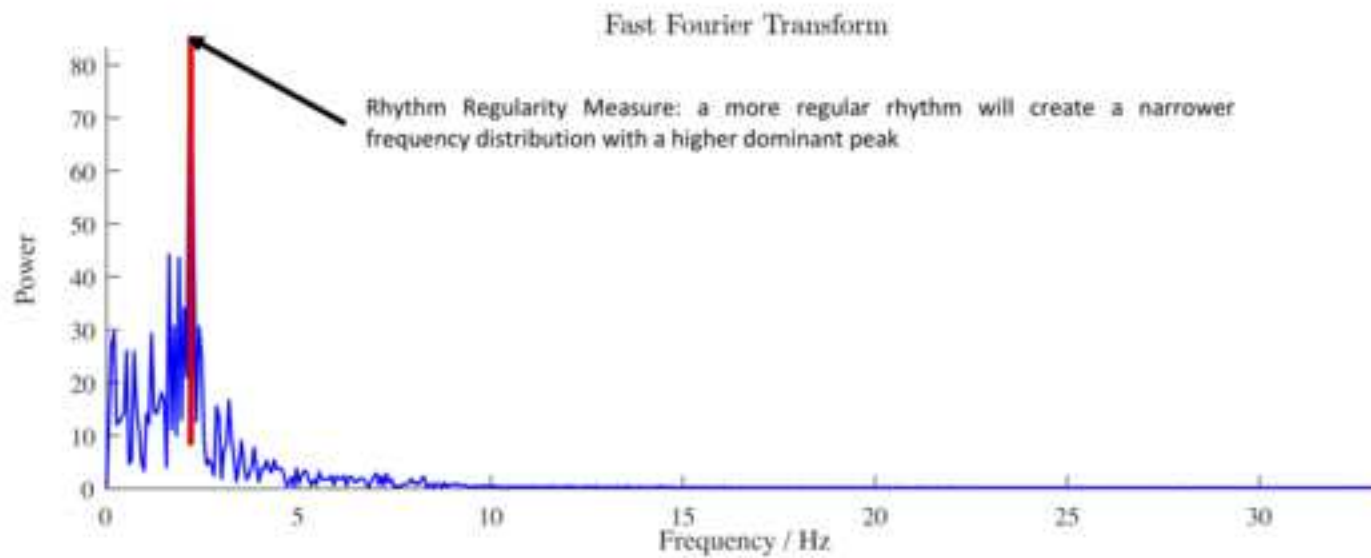
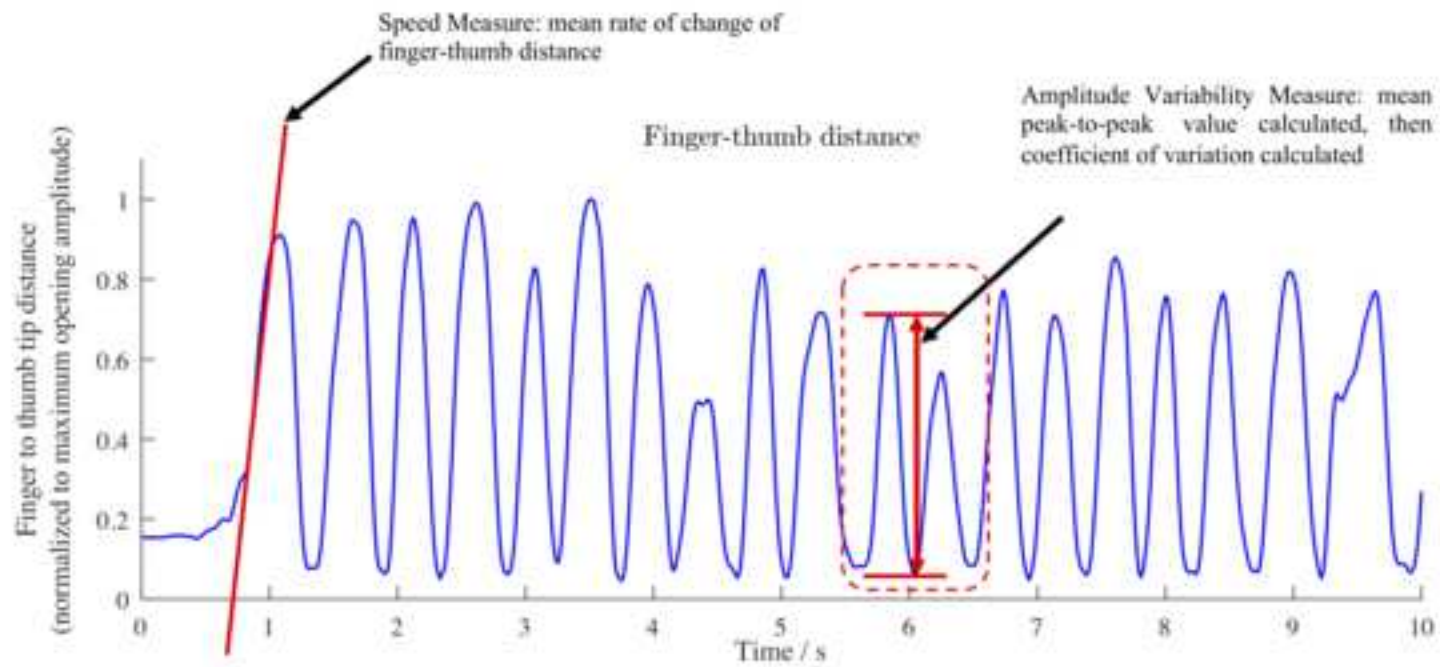
We would like to thank the following UK movement disorder neurologists for providing clinical rating of videos: Dr Jane Alty, Dr Sundus Alusi, Dr Michael Bonello, Dr Stephen Butterworth, Dr Philip Buttery, Dr Camille Carroll, Dr Adam Cassidy, Dr Jeremy Cosgrove, Dr Richard Ellis, Dr Jonathan Evans, Dr Paul Goldsmith, Dr Donald Grosset, Dr Christopher Kobylecki, Dr Alistair Lansbury, Dr Peter Moore, Dr Rachel Newby, Dr Edward Newman, Dr Gillian Sare, Dr Monty Silverdale, Dr Naomi Warren, Dr Louise Wiblin, Dr Caroline Williams-Gray.

We would also like to thank Dr Parisa Patel and Dr Paschal O’Gorman for data collection.

Author contributions

Name	Location	Role	Contribution
-------------	-----------------	-------------	---------------------

Stefan Williams, MBBS, BSc	University of Leeds, UK	Author	Design and conceptualized study; data acquisition and interpretation; drafted the manuscript for intellectual content
Zhibin Zhao, BEng	(1) University of Manchester, UK; (2) Xi'an Jiatong University, China	Author	Analyzed the data (major role)
Awais Hafeez, MA, BSc	(1) University of Leeds, UK; (2) University of Engineering and Technology Lahore, Pakistan	Author	Design and conceptualized the study; major role in acquisition of data
David C Wong, PhD	University of Manchester, UK	Author	Design and conceptualized study; analyzed the data; data interpretation; revised the manuscript for intellectual content
Samuel D Relton, PhD	University of Leeds, UK	Author	Design and conceptualized study; data interpretation; revised the manuscript for intellectual content
Hui Fang, PhD	Loughborough University, UK	Author	Design and conceptualized study (major role)
Jane E Alty, MD, MA, MBBChir	(1) University of Tasmania, Australia; (2) Leeds Teaching Hospitals NHS Trust, UK	Author	Design and conceptualized study; data interpretation; drafted the manuscript for intellectual content



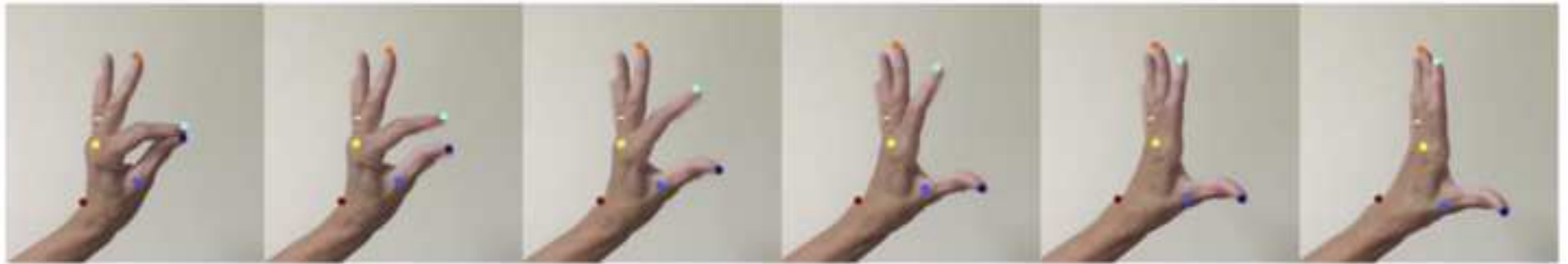
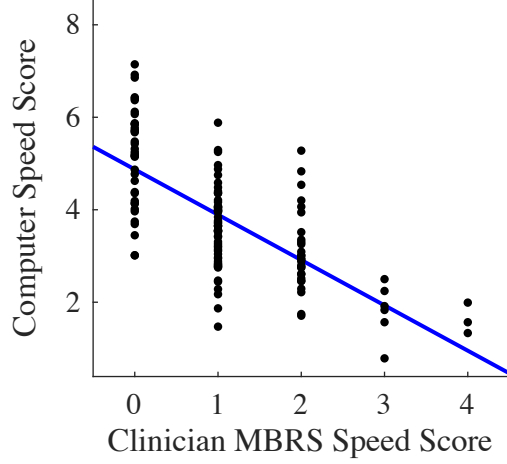
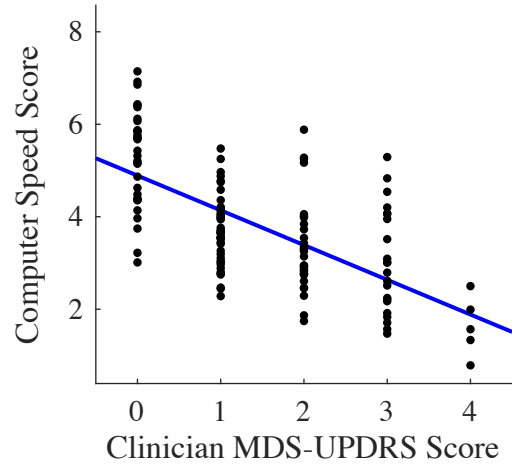


Figure 3

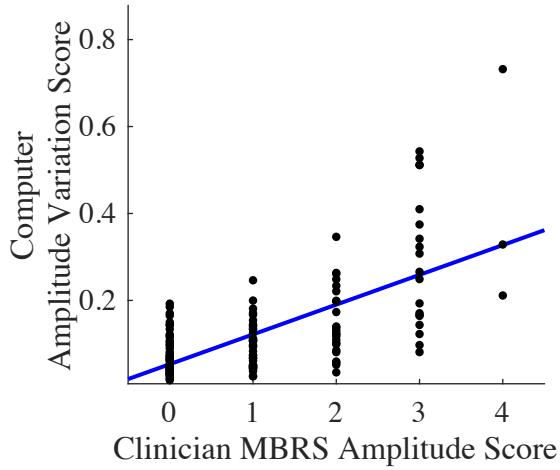
(a) $r = 0.70, p \ll 0.001$



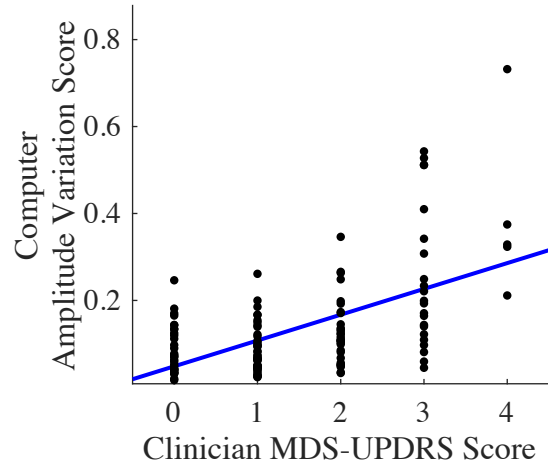
(d) $r = -0.65, p \ll 0.001$



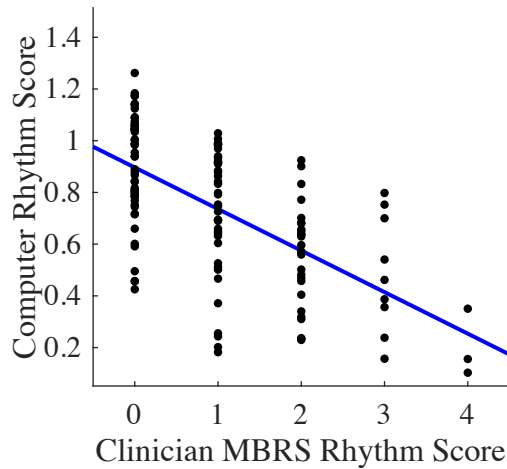
(b) $r = 0.65, p \ll 0.001$



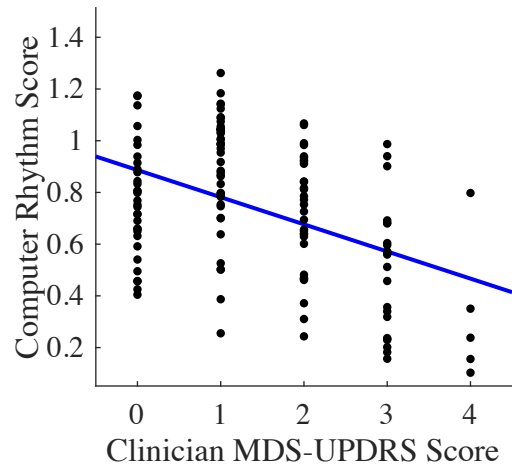
(e) $r = 0.57, p \ll 0.001$



(c) $r = -0.62, p \ll 0.001$



(f) $r = -0.44, p \ll 0.001$





Click here to access/download
e-component
Video_1.MOV