Journal of Parkinson's Disease 2 (2012) 47–55
DOI 10.3233/JPD-2012-11071
IOS Press

Automated Assessment of Bradykinesia and Dyskinesia in Parkinson's Disease

Robert I. Griffiths[a], Katya Kotschet[a,b], Sian Arfon[a], Zheng Ming Xu[d], William Johnson[b], John Drago[a,b,c], Andrew Evans[e], Peter Kempster[f], Sanjay Raghav[f,g] and Malcolm K. Horne[a,b,c,*]
[a]*Florey Neuroscience Institutes, University of Melbourne, Parkville, Victoria, Australia*
[b]*St Vincent's Hospital, Fitzroy, Victoria, Australia*
[c]*Centre for Neurosciences, University of Melbourne, Parkville, Victoria, Australia*
[d]*Compumedics, Abbotsford, Victoria Australia*
[e]*Royal Melbourne Hospital, Parkville, Victoria, Australia*
[f]*Southern Health, Clayton, Victoria, Australia*
[g]*Monash University, Clayton, Victoria, Australia*

Abstract. There is a need for objective measures of dyskinesia and bradykinesia of Parkinson's disease (PD) that are continuous throughout the day and related to levodopa dosing. The output of an algorithm that calculates dyskinesia and bradykinesia scores every two minutes over 10 days (PKG: Global Kinetics Corporation) was compared with conventional rating scales for PD in PD subjects. The algorithm recognises bradykinesia as movements made with lower acceleration and amplitude and with longer intervals between movement. Similarly the algorithm recognises dyskinesia as having movements of normal amplitude and acceleration but with shorter periods without movement. The distribution of the bradykinesia and dyskinesia scores from PD subjects differed from that of normal subjects. The algorithm predicted the clinical dyskinesia rating scale AIMS with a 95% margin of error of 3.2 units compared with the inter-rater 95% limits of agreement from 3 neurologists of −3.4 to +4.3 units. Similarly the algorithm predicted the UPDRS III score with a margin of error similar to the inter-rater limits of agreement. Improvement in scores in response to changes in medication could be assessed statistically in individual patients. This algorithm provides objective, continuous and automated assessment of the clinical features of bradykinesia and dyskinesia in PD.

Keywords: Parkinson's disease, bradykinesia, dyskinesia, spectral power, automated measurement, movement disorder

INTRODUCTION

Levodopa treatment for the bradykinesia of Parkinson's Disease (PD) is initially effective, but after approximately 5 years, 50% of PD subjects develop "wearing off", motor fluctuations and dyskinesia [4, 5]. Therapeutic interventions are then directed at minimising bradykinesia, dyskinesia, "off time" and motor fluctuations. The effectiveness of these interventions are assessed clinically, with diaries or with rating scales. Clinical rating scales have their well appreciated limitations, including the inter-rater variability and limited scope for continuous monitoring. While diaries have become important in measuring PD fluctuations, patients find it difficult to report the nature, severity and timing of change in motor states. Furthermore, diaries are frequently a record of recall because many subjects defer recording their clinical states for at least a day [6]. Thus, a means of continuously and objectively monitoring each clinical state would be invaluable for assessing the effect of therapeutic interventions, both in clinical trials and in providing routine clinical care.

*Correspondence to: Prof. M.K. Horne, Florey Neuroscience Institutes, Melbourne Brain Centre, @ Genetics lane, University of Melbourne, Parkville, 3010 Victoria, Australia. Tel.: +61 9035 6589; Fax: +61 613 9035 3107; E-mail: mhorne@unimelb.edu.au.

ISSN 1877-7171/12/$27.50

Attempts have been made to describe in Newtonian terms, the differences between normal movement and bradykinesia and dyskinesia. Bradykinetic movements are characterized by slowness in initiation and execution [7, 8], and are of lower acceleration and velocity than normal movements, particularly when movement amplitude is large [9–11]. Reductions in acceleration are apparent when subjects attempt rapid alternating movements [12–16]. It is likely therefore that bradykinetic movements would manifest in a spectral analysis as a relative preponderance of low frequencies and reduced power in all frequencies. To our knowledge, bradykinesia has not previously been subject to spectral analysis.

We have developed a method that provides dyskinesia and bradykinesia scores every two minutes over 10 days (PKG: Global Kinetics Corporation). These scores are produced by an algorithm that was developed using a fuzzy logic approach, the spectral power of the low frequencies of an accelerometer trace and with iterative modelling and retesting of new subjects. The aim of this study was to assess whether this algorithm addresses some short comings of conventional clinical rating methods. It was tested against rapid alternating movements (bradykinesia), AIMS (dyskinesia) and the scores obtained from the UPDRS III and IV clinical rating scales. This algorithm was then applied to continuously monitor subjects over 10 days.

These studies suggest that neurologists use the interval between movements, and the acceleration with which movements are made, to recognise bradykinesia and dyskinesia. Control subjects are more likely to make "normally" patterned movements with greater probability than PD subjects, and prolonged recordings provide a more reliable estimate of this probability. We conclude that the algorithm can objectively quantify bradykinesia and dyskinesia in PD and measure changes in the parkinsonian state in response to therapeutic interventions.

METHODS

Subjects

This study was approved and supervised by St Vincent's Hospital and Southern Health Human Research & Ethics Committees. All subjects provided consent according to the Declaration of Helsinki [17]. Recordings and UPDRS III and IV scores were obtained from 34 PD subjects with idiopathic levodopa-responsive PD, fulfilling the UK brain criteria [18], (including 9 with predominantly unilateral disease) and from 10 age-matched controls.

Collection and analysis of data

The Parkinson's Kinetigraph (PKG; Global Kinetics Corporation) was used to measure acceleration of the wrist. This device is worn like a wrist watch, weighs 35 g and contains a rechargeable battery and a 3-axis iMEMS accelerometer (ADXL345 Analog Devices) set to record 11-bit digital measurement of acceleration with a range of ± 4 g and sampling rate of 50 samples per second using a digital micro-controller and data storage on flash memory. After recording for up to 10 days, the data was downloaded, digitally filtered to include frequencies between 0.2 Hz and 4 Hz and analysed using propriety software (PKG – Global Kinetics Corporation –Australia) to produce bradykinesia and dyskinesia scores every two minutes from waking until retiring. The device was worn on the most severely affected limb of parkinsonian subjects and on the dominant limb of control subjects. The treating neurologist provided the prescribed levodopa regimen for each PD patient. The device was programmed to vibrate to alert the subject that a dose of levodopa was due. The patient confirmed the actual time when each dose was taken by placing their thumb on a sensor zone on the device face.

Measurement of bradykinesia and dyskinesia by algorithm

An expert systems approach was used to model neurologists' recognition of bradykinesia and dyskinesia. The acceleration recordings were divided into 2 minute epoch and the inputs to the expert system included Mean Spectral Power (MSP) within bands of acceleration between 0.2 and 4 Hz, peak acceleration and the amount of time within these epochs that there was no movement. These inputs received various weightings to model bradykinesia and dyskinesia. A bradykinesia score (BKS) was produced by establishing the maximum acceleration in each 2 minute epoch of acceleration recordings and calculating the MSP surrounding this peak. The rationale was that normal subjects commonly move with higher accelerations and energy than bradykinetic subjects. The dyskinesia score (DKS) required a calculation of the mean acceleration and the MSP in a 2 min epoch. The rationale was that dyskinetic subjects would have a larger MSP, in movements slower than the mean, than would normal subjects and would spend more time than normal subjects making movements whose accel-

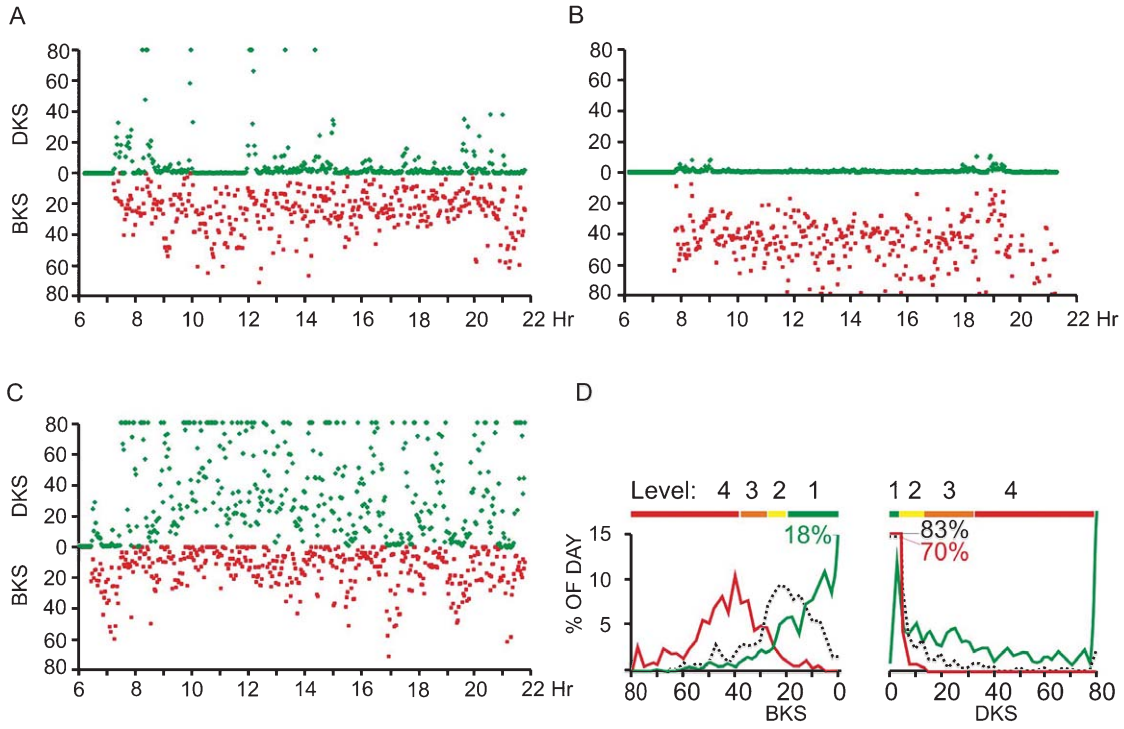

Fig. 1. BKS and DKS generated every 2 minutes from 6:00 to 22:00 throughout a single day from: A, a control; B, a subject with bradykinesia and C, a subject with dyskinesia. DKS points = Green, BKS = Red. D) A frequency histogram of BKS and DKS (x axis), normalized so as to show as a proportion rather than absolute values. Note that this is the distribution of BKS and DKS obtained for every 2 minutes in the data used to produce Fig. 1A, B and C. As each data point is measured over 2 minutes, the y axis in effect represents the percentage of time spent at a particular BKS or DKS (x axis). The y axis scale has been chopped at 15% of the day. Traces that exceed 15% have a number next to them, showing their peak. Thus the dyskinetic subject had a peak BKS of 18% and the bradykinesia subject and controls had a peak DKS of 83% and 70% respectively. In each case the peaks were close to BKS and DKS of 0. Control – Black Dotted lines; bradykinetic patient – Solid Red lines; dyskinetic patient – Solid Green lines.

eration was greater than the mean. Thus, the algorithm recognises bradykinesia as having fewer movements which, when compared with normal movements, are made with lower acceleration and amplitude and with longer intervals between movements. The algorithm similarly recognises dyskinesia as having movements of normal amplitude and acceleration but with shorter periods without movement.

Statistical analysis

The 95% limits of agreement and 95% prediction intervals were performed according to the methods of Bland and Altman [19]. Margins of error are all reported at the 95% confidence level.

RESULTS

In the algorithm, bradykinesia is recognised as epochs containing movements of lower acceleration and amplitude and with longer intervals between them. Dyskinetic epochs contain movements of normal amplitude and acceleration but with shorter intervals between them. The algorithm produced a bradykinesia score (BKS) and dyskinesia score (DKS) for each 2 minute epoch. Note that while normal subjects may have some epochs with BKS and DKS that are in the bradykinetic or dyskinetic range, PD subjects have more epochs with high BKS and DKS and this becomes clear with prolonged recordings. For example, distributions of DKS and BKS from control subjects recorded over 16 hours were markedly skewed with most scores being very low (Fig. 1A, D), but some were very high. As each BKS and DKS represents a 2 minute epoch, frequency histograms (Fig. 1D) show the proportion of the day spent at each BKS or DKS score. The average 50th, 75th and 90th percentiles of the distribution of BKS and DKS was obtained by recording from control subjects for 10 days between 9:00 and 18:00 hours: these were then used to define four levels of severity for bradykinesia and dyskinesia (Figs. 1D and 3).

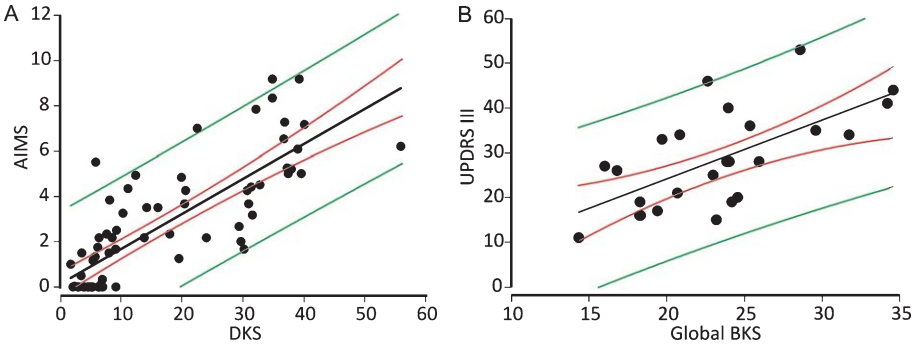

Fig. 2. A) AIMS verses 2 minute DKS. $R = 0.80$, $n = 71$, $P < 0.0001$. 95% confidence intervals (red lines) for the regression and 95% prediction intervals (green lines). B) The UPDRS III was plotted against Global BKS. The UPDRS III can be predicted from the Global BKS with a margin of error of 18 UPDRS III units.

Two PD subjects provide examples of how the distribution of BKS or DKS differs from the control subject (Fig. 1). These subjects had modest bradykinesia (Fig. 1B: UPDRS III of 21 and UPDRS IV of 0) and dyskinesia (Fig. 1C: UPDRS III of 8 and UPDRS IV of 5). The BKS of the bradykinetic subject was greater than the 50th percentile of controls (BKS = 20) for almost 90% of the time. The 95% confidence interval has the bradykinetic subject spending between 41% and 51% more time with a BKS > 20 than the control ($P < 0.0005$ ($\chi^2 = 0.274$, d.f. 32)). The DKS of the dyskinetic subject was greater than the 50th percentile of controls (DKS = 4) almost all of the time. The 95% confidence interval puts the dyskinetic patient as having between 55% and 65% more time with a DKS score >10 than the control ($P < 0.0005$ ($\chi^2 = 0.394$, d.f. 32)).

Comparisons of DKS and BKS with AIMS and UPDRS III and IV

The algorithm for DKS was initially developed within the laboratory by videoing subjects for 12 minutes while performing naturalistic tasks. The modified AIMS [20, 21], assessed by three Movement Disorder Specialists were regressed against the DKS (Fig. 2A). The correlation was highly significant (Pearson's: $p < 0.0001$: R of 0.80). The margin of error in predicting the AIMS from DKS was 3.2 AIMS units, compared with −3.4 to +4.3 AIMS units by the neurologists (with the maximum score being 20).

However, there must be sufficient 2 minute epochs to produce a reliable distribution (as in Fig. 1D) and 12 minutes of recording will only produce 6 data points. Prolonged recording of subjects going about normal activities would be ideal but it is not possible to provide AIMS over that period. Nevertheless, from a 10 day recording, the median of the total 2700 DKS recorded between 9:00 and 18:00 hours, can be reduced to a single figure: the "global median DKS". This was obtained from 25 subjects with established bilateral PD over 10 days, and compared with their UPDRS IV score (sum of Questions 32 and 33) measured in the "on" state. The Global DKS and the UPDRS IV were correlated ($p < 0.05$) and the Global DKS predicted the UPDRS IV with a margin of error of 3.9 over a possible range of 0–8. Published inter-rater variability in scoring UPDRS IV could not be found for comparison.

The algorithm for BKS was modelled on PD subjects sliding their forefinger between two large dots (diameter 30 mm) placed 300 mm apart on a piece of cardboard for 30 seconds: a variation on key press or peg board tests for assessing bradykinesia [12–15, 22] and chosen to avoid acceleration strike artefact. Bradykinesia measured by the algorithm correlated well with bradykinesia measured by the dot slide ($r = 0.63$, $p < 0.001$, $n = 79$) with a specificity of 88% and sensitivity of 95%.

UPDRS III scores were obtained from the 25 subjects with well established bilateral PD in the "on" state. A "global median BKS" was derived similarly to the Global DKS, from 10 days of recording. The UPDRS III and global median BKS were significantly correlated (Fig. 2B, $p < 0.0005$) with an R of 0.64. The slope of the line of best fit suggests that 1 UPDRS III Unit = 1.3 BKS units. The margin of error in predicting the UPDRS III from BKS was 18 UPDRS III units: note that up to 11 UPDRS III units may be due to inter-rater variation (limits of agreement for UPDRS motor scores from different raters for the same subject [2]).

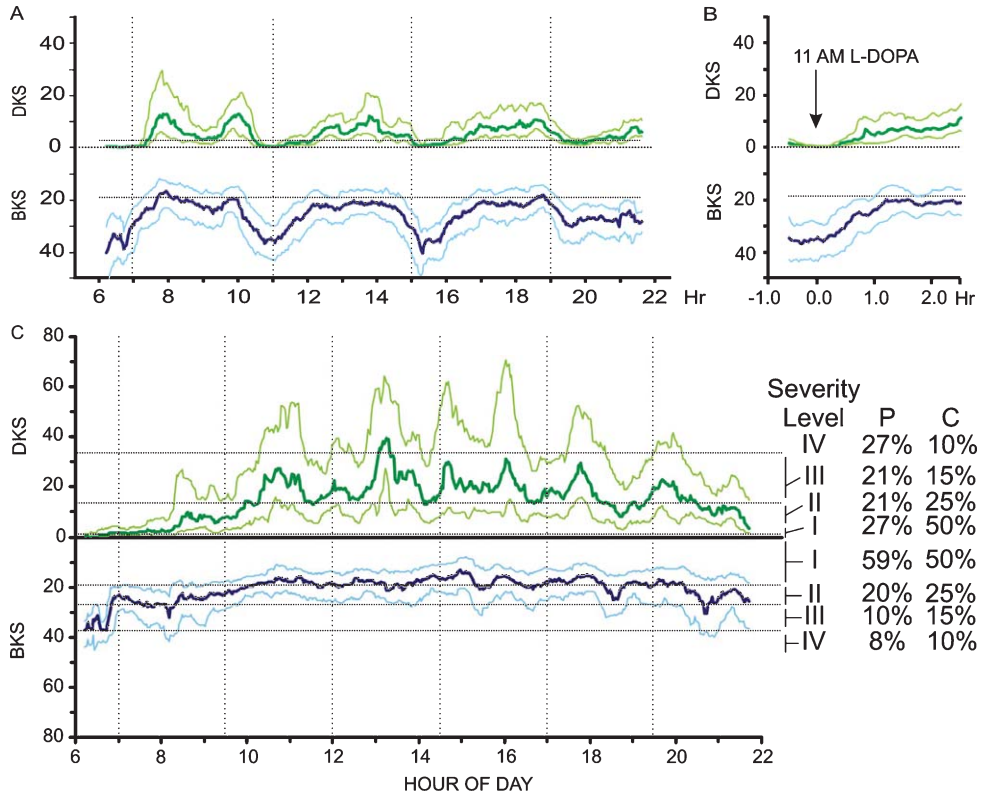

Fig. 3. Ten days recordings from a single subject showing the average moving 30 minute median of both DKS (green line: above zero) and BKS (blue line: below zero) plotted against the hour of day. The average moving 25th and 75th percentiles are plotted in light blue (BK) and light green (DK). Prescribed times for a dose of levodopa are marked with a vertical dotted black line. 3A) An example where levodopa was prescribed at 0700, 1100, 1500 and 1900 hrs. The peri dose response curve with the BKS and DKS aligned on the actual time when 1100 dose was taken is shown in (3B). 3C) Shows an example of peak dose dyskinesia. The table to the right of this graph, shows the proportion of time spent in the four levels of severity of DKS and BKS for this patient (P) compared to controls (C). The four levels were Level I: <50th percentile, Level II: 50–75th percentile, Level III: 75th −90th percentile and Level IV: >90th percentile. This subject spent 27% of the time in Level 4 compared with controls, who spend 10% of their time in this Level.

Recordings were also made from the most affected upper limb of 9 PD subjects with predominantly unilateral disease. The UPDRS III scores of these subjects were approximately half of that predicted from their Global BKS, suggesting that the algorithm scores a subject with unilateral disease as if the subject had both limbs affected equally.

Temporal variations in BK score and DK score

The temporal fluctuations in BKS and DKS over 10 days can be assessed by calculating their 25th percentile, median and 75th percentile every two minutes of the 10 day recording and smoothing (15 point moving box plot: Fig. 3A). As an example, a bradykinetic subject with a UPDRS III score of 40 and mild dyskinesia (AIMS of 5), took four doses of levodopa/day (Fig. 3A). The median BKS score ranged from about 40 (prior to medications at 7:00 and 15:00 hours) to about 15 following each of these doses. There appeared to be "wearing off", particularly obvious in the half hour prior to the 11:00 and 15:00 doses. The recordings were then aligned with the actual time that the scheduled 11:00 hours levodopa dose was taken (Fig. 3B) to produce a levodopa response curve, showing BKS and DKS for the ½ hour before and 2½ hrs after this dose. Within an hour of dosing, this subject's median BKS went from around 40, and greater than the 75th percentile (BKS 27) of control subjects, to less than 20, which is almost the median (19) of control subjects. There was a commensurate but moderate increase in DKS. The case demonstrated in Fig. 3C has peak dose dyskinesia. The table to the right of this graph shows the proportion of time spent in the four levels of severity of DKS and BKS for this patient compared to controls (Fig. 3C).

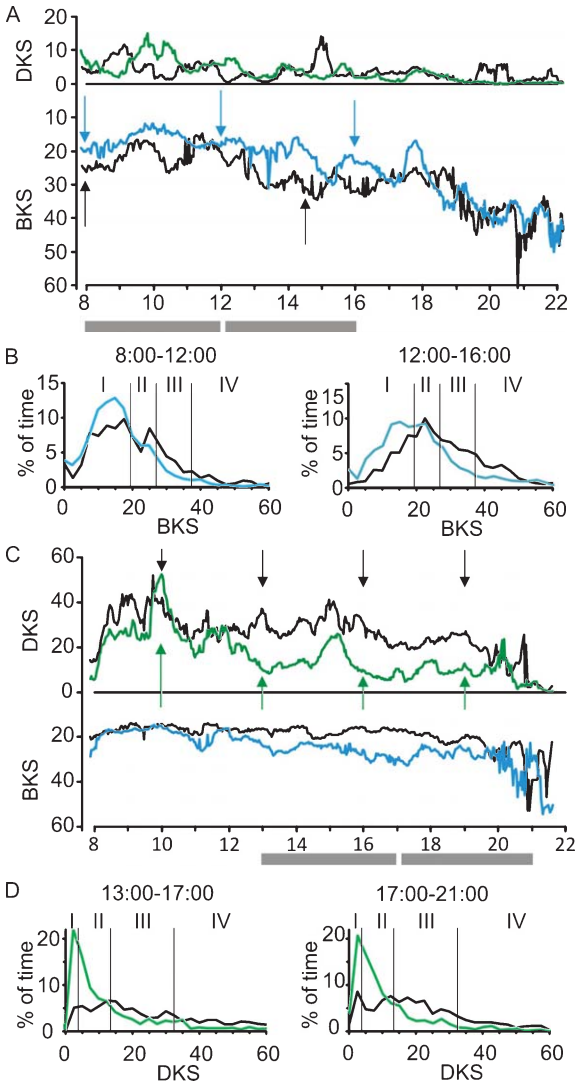

Fig. 4. A & B) Ten day recordings were made before (black lines) and after (BKS blue and DKS green lines) a change in medication to improve bradykinesia. The number of doses of levodopa was increased from two doses (black upward directed arrows) to three doses (blue downward directed arrows) a day, resulting in reduced BKS, especially between 12:00 and 17:00 hrs. 4B should be read in conjunction with Table 1. The changes in BKS in the four intervals of 8:00–12:00 and 12:00–16:00 (marked as a grey line under the x axis of A) were quantified by plotting a frequency histogram of all the 2 minute epochs that occurred over the 10 days in the respective 4 hour intervals (similar to Fig. 1D). Each plot has three vertical dotted lines dividing the BKS into the four levels of severity (as in the table of Fig. 3B). Table 1 provides a quantification of the proportion of time spent in each severity level before and after treatment. C & D) A case were the timing of levodopa doses was unchanged but the dose was lowered and amantadine was added to reduce dyskinesia. 4D should be read in conjunction with Table 1D and shows the percentage of time spent at each DKS level over a four hour period (indicated above each graph and by the grey line under the x axis of C). In both time intervals, treatment has shifted the curve in 4D to the left and into lower severity levels.

The distribution of BKS and DKS (as described in Fig. 1D) allow the effect of a therapeutic intervention to be quantified (Fig. 4). The subject in Fig. 4A had UPDRS III scores of 19 before, and 17 after, an extra levodopa dose was added. The median BKS between 8:00 till 12:00 hours improved by 3.4 units, and between 12:00 till 16:00 hours improved by 6.8 units. The percent of time spent at each BK score and in each severity level (as described above and in Fig. 1) for these time periods were plotted (Fig. 4B and Table 1). Control subjects spend 50% of time in Level I of severity, and there was significantly more time ($P < 0.001$ χ^2) in this level after the change in treatment.

The effect of treatment on DKS is illustrated by another case (Fig. 4C), whose UPDRS IV scores were 8 and 4 (respectively), before and after treatment was modified to reduce dyskinesia. The change in medication was most evident after 12:00 hours, with the global DKS reduced from 40 to 20 after treatment modification. The median DKS for the time period 13:00 till 17:00 hours improved by 16.9 units; and for 17:00 to 21:00 hours improved by 12.0 units. For the time period 13:00 till 17:00 hours, the proportion of time with DKS in Level I of severity increased from 33%, to 65% after the treatment change ($P < 0.001$ χ^2). This demonstrates that the effects of treatment can be placed under statistical scrutiny, even in an individual.

DISCUSSION

The algorithm models clinical features used by neurologists to recognise bradykinesia and dyskinesia. Two important aspects were the interval between movements and the acceleration with which movements were made: bradykinetic subjects have longer intervals between movement and when they do move it is with lower acceleration, whereas dyskinetic subjects have fewer intervals between movement and they move with a greater spectral power. Another insight was that control subjects make "normally" patterned movements with greater probability than PD subjects, rather than there being an "abnormal" signature for bradykinesia and dyskinesia. Thus prolonged recordings provide a more reliable estimate of this probability.

The algorithm differs from other studies using accelerometry in that it does not depend exclusively on mean spectral power but emphasises weighting of power in specific spectral bands, in the peak acceleration and in the duration of time where movement is absent. Importantly, it has not sought a signature

Table 1
The proportion of time spent in each BKS (patient 1) and DKS (patient 2) level of severity, before and after the change in medication. In each case there is a relative shift along the left axis of the relevant figure, shown in the table as a higher proportion of time in lower stages (and lower proportion in higher stages), after the change in medication

Patient 1 Fig. 4B					
BKS		8:00–12:00 hrs		12:00–16:00 hrs	
level	Controls	Before	After	Before	After
I	50	48	64	22	44
II	25	21	21	26	26
II	15	18	10	26	16
IV	10	13	5	16	14
Patient 2 Fig. 4D					
DKS		13:00–17:00 hrs		17:00–21:00 hrs	
level	Controls	Before	After	Before	After
I	50	8	32	12	32
II	25	21	31	24	39
II	15	31	20	40	18
IV	10	40	16	24	11

abnormality in the trace but rather has taken a probabilistic approach and shown that the distribution of DKS and BKS over time differs in controls and PD subjects, requiring that recordings are prolonged for patterns to fully emerge. The algorithm emphasises the assessment of subjects doing routine daily activities rather than in the artificial context of clinical scrutiny. Arguably the cortico-basal ganglionic circuitry damaged by PD normally produces automatic movements for the purpose of goal directed activity and these movements are best observed when attention is directed away from the movement under scrutiny.

Continuous recording allows the amount of time spent at various levels of bradykinesia and dyskinesia during the day to be quantified and correlated with the timing of medications. This permits a statistical measure of the effectiveness of a therapeutic intervention, even in an individual. However, clinical and statistical significance are not the same thing. Noting that the slope of the fitted line in Fig. 2B predicts that a UPDRS III unit equals 1.3 BKS units, then 3 UPDRS III units, which is considered a clinically important improvement in a population of several hundred patients [1], would relate to 3.9 BKS units. However, it is questionable whether a change of 3 UPDRS III units for an individual would be within the noise of intra-rater variability: even a Senior Movement Disorder Specialist has 95% intra-rater repeatability limits in the range −8.3 to 7.1 UPDRS III points [2].

The correlation between UPDRS III scores and BKS from 25 subjects with bilateral well established disease (Fig. 2A) was highly significant (Pearson's: $p < 0.0005$) although an R of 0.65 is modest. However, there are important reasons that the correlation between an algorithm and the clinical scales is unlikely to ever be much better. Rigidity, tremor and various manifestations of bradykinesia each contribute to the UPDRS III score but they may also vary independently of each other, contributing to overall variation in the UPDRS as a measure of bradykinesia [23], which is the variable of interest for the BKS. Furthermore, the variance of the algorithm's measure is small because it is repeated over several days and there is no inter-rater variance, whereas the UPDRS III is measured only once (and hence without standard error), precluding direct comparison. PD is inherently variable from day to day and a single UPDRS III score fails to capture this fluctuation. Ideally, UPDRS should be scored frequently and correlated with the BKS scores at those times. Finally, the UPDRS III scores were provided by 6 different neurologists: inter-rater variance is significant and could contribute more than half of the margin of error in predicting the UPDRS III from BKS [2, 24]. We concluded therefore that the UPDRS is making a greater contribution to low covariance and a low R than would the BKS. In the case of the DKS, the relationship with the UPDRS IV was weak. However, the UPDRS IV is a coarse scale establishing proportion of the day rather than severity and takes no account of hour to hour variation in clinical state or correlation with medications and future studies should use a broader range of rating scales and diaries for comparison. Thus, DK and BK scores do correlate with the corresponding aspects of the UPDRS, but they do measure different aspects of bradykinesia and dyskinesia so while some correlation might be expected,

close correlation would be surprising. While we have focussed on differences between the UPDRS III & IV and algorithm scores, the difficulties will be similar with other rating methods because there is no "gold standard" with which all scales should be compared. Clinical rating scales are non-linear, require training and experience and are most precise when repeated by the same clinician. These scales can only be used when a trained observer is present and thus cannot provide a continuous measure of disease fluctuation. Diaries are used in attempt to address continuous monitoring but they are reported in binary terms ("on" vs. "off"), rely on patients' self-assessment of the nature, severity and timing of change in motor states, and on their recall if subjects have deferred completion of their diary in a timely fashion. The limitations of all these measures are reported and a high correlation with these scales would simply reflect replication of their weaknesses. Furthermore, the algorithms for BKS and DKS were developed and validated against accepted measures of bradykinesia (the "dot slide" task [12–15, 22]) and dyskinesia (the AIMS score [20, 21]) and further correlation with clinical scales argues that they do have validity and justifies their consideration for long term studies. Nevertheless, this algorithm clearly requires further comparisons with diaries and other rating scales.

In summary, the algorithms described here show promise as objective, continuous, quantitative measures of the severity and proportion of time spent at various levels of bradykinesia and dyskinesia and the temporal correlation of the motor fluctuation with timing of medication. These results warrant further studies using larger numbers of subjects and a broader range of clinical scales but shows promise for measuring the effect of therapeutic interventions in clinical trials and in routine clinical practice.

ACKNOWLEDGMENTS

This research was funded by development grants from the Florey Neuroscience Institute, the Pfizer Neuroscience Research grant (to KK), by the Medical Council Research Funds and the Victorian Government's Operational Infrastructure Support Program. Authors, RG, KK, Z-MX and MH have financial interest in Global Kinetics. MH is a Medical Practioner Fellow and JD is a Senior Research Fellow of the National Health and Medical Research Council of Australia. Global Kinetics provided the PKG devices.

REFERENCES

[1] Shulman LM, Gruber-Baldini AL, Anderson KE, Fishman PS, Reich SG, & Weiner WJ (2010) The clinically important difference on the unified Parkinson's disease rating scale. *Archives of Neurology*, **67**(1), 64-70.

[2] Post B, Merkus MP, de Bie RM, de Haan RJ, & Speelman JD (2005) Unified Parkinson's disease rating scale motor examination: Are ratings of nurses, residents in neurology, and movement disorders specialists interchangeable? *Mov Disord*, **20**(12), 1577-1584.

[3] Papapetropoulos SS (2011) Patient diaries as a clinical endpoint in Parkinson's disease clinical trials. *CNS Neuroscience & Therapeutics*, doi: 10.1111/j.1755-5949.2011.00253.x

[4] Nutt JG (1990) Levodopa-induced dyskinesia: Review, observations, and speculations. *Neurology*, **40**(2), 340-345.

[5] McColl CD, Reardon KA, Shiff M, & Kempster PA (2002) Motor response to levodopa and the evolution of motor fluctuations in the first decade of treatment of Parkinson's disease. *Mov Disord*, **17**(6), 1227-1234.

[6] Stone AA, Shiffman S, Schwartz JE, Broderick JE, & Hufford MR (2002) Patient non-compliance with paper diaries. *BMJ*, [Research Support, U.S. Gov't, P.H.S.], **324**(7347), 1193-1194.

[7] Yanagisawa N, Fujimoto S, & Tamaru F (1989) Bradykinesia in Parkinson's disease: Disorders of onset and execution of fast movement. *European Neurology*, **29**(Suppl 1), 19-28.

[8] Pahwa R, & Koller WC (1995) Defining Parkinson's disease and parkinsonism. In: Ellenberg JH, Koller WC, Langston JW, Editors. Etiology of Parkinson's disease, Dekker, New York, p. 1-54.

[9] Sheridan MR, & Flowers KA (1990) Movement variability and bradykinesia in Parkinson's disease. *Brain: A Journal of Neurology*, **113**(Pt 4), 1149-1161.

[10] Marsden CD (1989) Slowness of movement in Parkinson's disease. *Mov Disord*, **4**(Suppl 1), S26-S37.

[11] Johnels B, Ingvarsson PE, Thorselius M, Valls M, & Steg G (1989) Disability profiles and objective quantitative assessment in Parkinson's disease. *Acta Neurol Scand*, **79**(3), 227-238.

[12] O'Boyle DJ, Freeman JS, & Johnels FW (1996) The accuracy and precision of timing of self-paced, repetitive movements in subjects with Parkinson's disease. *Brain: A Journal of Neurology*, **119**(Pt 1), 51-70.

[13] Ghika J, Wiegner AW, Fang JJ, Davies L, Young RR, & Growdon JH (1993) Portable system for quantifying motor abnormalities in Parkinson's disease. *IEEE Trans Biomed Eng*, **40**(3), 276-283.

[14] Eichhorn TE, Gasser T, Mai N, Marquardt C, Arnold G, Schwarz J, & Oertel WH (1996) Computational analysis of open loop handwriting movements in Parkinson's disease: A rapid method to detect dopamimetic effects. *Movement Disorders*, **11**(3), 289-297.

[15] Kraus PH, Klotz P, Fischer A, & Przuntek H (1987) Assessment of symptoms of Parkinson's disease by apparative methods. *Journal of Neural Transmission Supplementum*, **25**, 89-96.

[16] van Hilten JJ, van Eerd AA, Wagemans EA, Middelkoop HA, & Roos RA (1998) Bradykinesia and hypokinesia in Parkinson's disease: What's in a name? *J Neural Transm*, **105**(2-3), 229-237.

[17] International Committee of Medical Journal Editors (1991). Style Matters: Statements from the Vancouver Group. *British Medical Journal*, **302**, 1194.

[18] Hughes AJ, Daniel SE, Kilford L, & Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*, **55**(3), 181-184.

[19] Bland JM, & Altman DG (2003) Applying the right statistics: Analyses of measurement studies. *Ultrasound Obstet Gynecol*, **22**(1), 85-93.

[20] Marsden CD, & Schachter M (1981) Assessment of extrapyramidal disorders. *Br J Clin Pharmacol*, **11**(2), 129-151.

[21] Hughes AJ, Frankel JP, Kempster PA, Stern GM, & Lees AJ (1994) Motor response to levodopa in patients with parkinsonian motor fluctuations: A follow-up study over three years. *J Neurol Neurosurg Psychiatry*, **57**(4), 430-434.

[22] Phillips JG, Bradshaw JL, Chiu E, & Bradshaw JA (1994) Characteristics of handwriting of patients with Huntington's disease. *Mov Disord*, **9**(5), 521-530.

[23] Stochl J, Boomsma A, Ruzicka E, Brozova H, & Blahus P (2008) On the structure of motor symptoms of Parkinson's disease. *Mov Disord*, **23**(9), 1307-1312.

[24] Schuurman PR, Bruins J, Merkus MP, Bosch DA, & Speelman JD (2002) A comparison of neuropsychological effects of thalamotomy and thalamic stimulation. *Neurology*, **59**(8), 1232-1239.