A Clinical Applicable Smartwatch Application for Measuring Hyperkinetic Movement Disorder Severity

Vochteloo, M.; Tijssen, M. A.J.; Beudel, M.

Published in:
2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2019

DOI:
10.1109/EMBC.2019.8857869

Publication date:
2019

A Clinical Applicable Smartwatch Application for Measuring Hyperkinetic Movement Disorder Severity.

M. Vochteloo, M.A.J. Tijssen, M. Beudel

Abstract **Measuring the severity of hyperkinetic movement disorders like tremor and myoclonus is challenging. Although many accelerometers are available to quantify movements, the vast majority lacks real-time analysis and an interface that makes it possible to real-time adjust therapy like deep brain stimulation (DBS). Here, we developed a smartwatch / smartphone application that is capable of real-time analysing movement disorder severity. Movement analysis was realised by integrating acceleration values, to velocity and subsequently to distance. Measured distances were compared with a validated accelerometer already applied for quantifying movement disorders. Further validation was done by quantitative assessment of simulated movement disorders in 10 healthy volunteers. Finally, the approach was tested in two patients treated with DBS to quantify the effect of different DBS settings on myoclonus and tremor severity, respectively.**
The distance measured with the application had a 96% accuracy. This was non-inferior (p = 0.76) compared to accelerometers already clinically applied. Furthermore, (simulated) movement disorder severity could be classified correctly in 93% of the cases. Finally, the method was capable of distinguishing effective from non-effective DBS parameters in two patients. In summary, with our approach we realised an instantaneous and reliable estimation of the severity of movement disorders which can assist in real time titrating therapy like DBS.

I. Introduction

Tremor is an oscillating involuntary rhythmic movement of a body part (2) and is the most frequently occurring movement disorder. In most cases tremor is mild and can be managed with medication but in severe cases, in which it interferes with activities of daily living, Deep Brain Stimulation (DBS) can be a therapeutic option (3). DBS is a treatment in which high-frequent electrical stimulation modulates pathological neural activation in deep brain nuclei. Although DBS is an established treatment for refractory tremor, there are still limitations in terms of efficacy and side effects. Furthermore, no consensus is present about optimally titrating stimulation parameters.

Myoclonus is a rare type of movement disorder in which patients experience repetitive (very) brief involuntary twitching of a muscle or a group of muscles. In severe cases, myoclonus patients can be treated with DBS (4). Although the clinical benefit of DBS is established, optimal stimulation settings are difficult to acquire and differ between patients. Furthermore, previous research has shown that the visual characterisation of myoclonus is difficult and unreliable (5). One possible explanation for this might be that the typical duration of myoclonus (for review see 6), around 100 ms, is too short to be reliably observed by human (visual) perception (7).

Since it is of crucial importance to objectively characterise the severity of tremor and myoclonus for titrating DBS stimulation settings, there is an unmet clinical need. This is especially the case since > 10.000 DBS stimulation settings can be configured are present, which can't all be clinically tested due to patient fatigue and limited healthcare resources. For this reason we developed and validated a smartwatch application that is capable of real-time analysing movement disorder severity.

II. Materials and Methods

A. Soft and Hardware

The application was written using Xcode version 8.3.3 (Apple Inc, Cupertino, California, US) and Swift 3.0 (Apple Inc, Cupertino, California, US). Experimental testing was done using the accelerometer and gyroscope in the Apple Watch series 2 with WatchOs 3.2 (Apple Inc, Cupertino, California, US) that was controlled by an iPhone 5 with iOS 10.3 (Apple Inc, Cupertino, California, US). The analysis of the clinical validation data was performed using the R software package (version 3.3.3, R Core Team, Auckland, New Zealand).

B. Data Processing Pipeline

Human upper extremity movement data was extracted from the accelerometer and gyroscope of the Apple Watch using the 'Core Motion' framework from which device motion filtered from gravitational biases could be derived. This filtering was performed using a build-in accelerometer-gyroscope fusion algorithm (8). Sensor bias due to DC-offset, calibration time and constant noise corrections have been implemented in the software design. DC-offset was corrected by subtracting the average value of each individual measurement. The calibration time differed per accelerometer axis, per hardware device and between sessions. The majority of calibration times remained below 1 second. For this reason, a calibration time of at least 1 second was used for all recordings. Constant noise was subtracted for each axis. A sample frequency of 50 Hz was used.
To quantify movements the accelerometer signal of the *x,y* and *z* axis was integrated twice to obtain a distance estimation. Although this approach has its limitations in freely moving situations, it has been validated in clinical testing (9). Integration was performed using trapezoid rule approach. To obtain a dynamic estimation of distance and reduce measurement mistakes in the double integral, a moving average approach was used. For further analyses, a 'severity score' was used as quantification. This score represents the total displacement over all three axis in a

M. Vochteloo is with the Institute of Life Sciences and Technology, Hanze University of Applied Sciences, Groningen, The Netherlands.
M.A.J. Tijssen is with the Neurology Department, University Medical Center Groningen, University of Groningen, The Netherlands
M. Beudel is with the Neurology Department, Amsterdam University Medical Centers, Amsterdam, The Netherlands (e-mail: m.beudel@amc.uva.nl).

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

certain temporal interval. Given the restrictions of the applied method mentioned above we used the value as a score and not as a true distance in centimeters.

C. Validation

The application (https://itunes.apple.com/us/app/wmdq/id1342633514?mt=80, Fig. 1) was validated in four stages. The first stage considered the estimation of the accelerometer biases and default parameter setting. The second stage considered the accuracy of linear motion detection and comparison with an established accelerometer applied in movement disorders, the third stage considered the classification of a simulated movement disorders and the fourth stage considered clinical testing of the approach in two patients treated with DBS.

1. Accelerometer Bias Corrections

A common sensor bias in accelerometer sensors is a short deviation at the start of a new measurement. It is only after a few iterations that correct acceleration values are reported. The duration of this calibration time differs between changing environmental conditions and various (sub)types of accelerometers. Due to the initially short duration of the measurements performed for accessing the severity score in this study, a modest deviation for only few iterations can have major impact on the resulting score. Because of the known high variability in calibration time for these sensors, there is no default practice for correcting for this. In order to reduce the error caused by this sensor bias, the calibration time was measured at 100 different instances divided over several days and time periods. The average calibration time was defined as the moment that accelerometers read out an acceleration below 0.005G over all axis while having the sensor on a flat horizontal surface.

The noise of the acceleration has been measured in a similar fashion as aforementioned (100 instances divided over several days and time periods). The noise was defined as the average acceleration per axis per measurement after calibration of the seniors was achieved (< 0.005G acceleration measured on flat horizontal surface).

2. Linear Movements

In order to use the severity score as a reliable measurement tool for tremor and myoclonus severity, the score needs to have a linear correlation with the traveled distance of a patients wrist. In other words, a double the increase in wrist movement is expected to results in double the severity score. In order to investigate that the accelerometer in the smartwatch is capable of doing such measurements, the accuracy of the accelerometer was tested by moving the smartwatch over a known linear distance and comparing the severity score (total displacement in centimeters) with the actual traveled distance. This test was performed by moving the smartwatch five times over six different known distances (0, 5, 10, 15, 20 and 25 centimeter) over a custom made track for improved single-axis movement. After this, the averages of each distance were plotted and the linear trend of the measurements was calculated using the R-squared approach. Furthermore, the accuracy of the smartwatch accelerometer was compared with the accuracy of another commercially available accelerometer (Shimmer, Dublin, Ireland) already applied in movement disorder research (1).

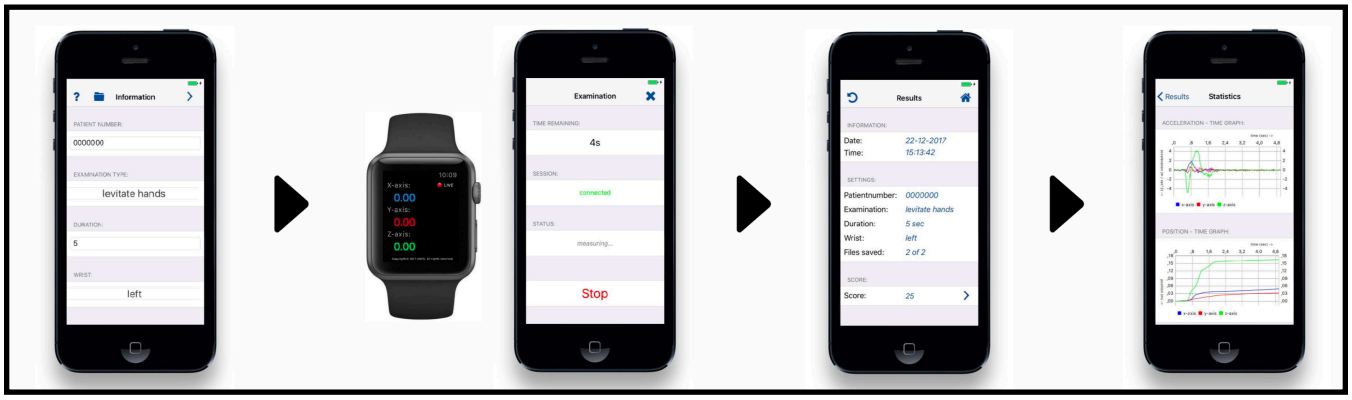
3. Movement Disorder Simulation

In the third step of the validation experiment ten volunteers were instructed to simulate myoclonus with three different severity levels. The explicit instruction was to make an up- or downward movement with both arms with either a small, medium or large amplitude as well as a null measurement. Subjects were ask to perform each condition three times (total 12 responses). Simulated myoclonus was quantified by performing a 5 second recording. After these measurements, a series of four random generated amplitudes were simulated blinded from the researcher. Each blinded measurement was classified to one of the four severities according to the severity score resulting from the application. The classification process was done by classifying the score to the closest mean of the baseline measurement of that subject.

4. Clinical Testing

In the third step of the validation, the severity score was used to assess the severity of tremor and myoclonus during the application of different DBS settings. In the tremor patient, 10 combinations of pseudo-randomised stimulation parameters (voltage, pulse-width, frequency) were applied to one DBS contact in one hemisphere. For each configuration, a measurement with the smart-watch was performed. Next to this, a subjective rating was given after testing each stimulation configuration. The experimental procedure was approved by the local medical ethical board and deemed "care as usual".

Figure 1: appearance of the combined smartphone / smartwatch app. The first screen indicates the start screen whereafter the smartwatch is activated and starts the examination (second screen). The third screen depicts the visualisation of the results including the severity score (bottom line). The last screen depicts the graphical visualisation of the movement recording.



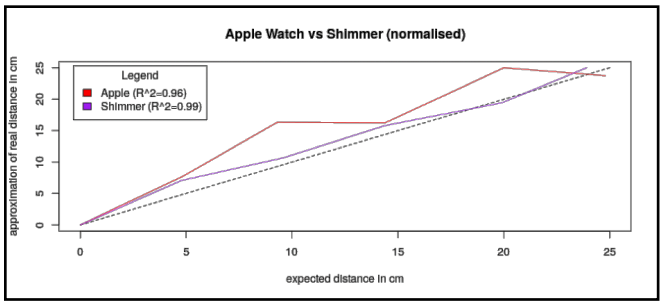


Figure 2: comparison of measured distance of the new smartwatch application with an established accelerometer applied for measuring movement disorders. Measured distances were normalised between 0-25. No significant differences were present between the two measurements. R^2 = R squared. cm = centimeter.

In the myoclonus patient each of the DBS contact points (2 * 4) in each hemisphere was individually tested with ascending voltages. This was performed in such way that myoclonus severity was assessed from 0V until 4V or until the voltage that resulted in the occurrence of side effect in steps of 0.5V. The most prominent change in positive and negative direction of the severity score induced by DBS was calculated for each electrode and averaged afterwards. Positive and negative deflections were tested for normality and compared using a dependent T test (Fig 4b).

III. Results

1. Accelerometer Bias Corrections

The average acceleration time was comparable between axis and was on average 0.9 seconds (longest time reported per axis per measurement). An observation worth mentioning here is that the calibration time spikes the first time the sensor is read out after a longer period of inactivity. Concluded from these result, the calibration time has been set on 1.0 seconds as default value.

The accelerometer noise was $3.4 \bullet 10^{-4}$ and $1.7 \bullet 10^{-3}$ cm per measurement for the x,y and z-axis respectively. This noise values for each of the axis is subtracted for each readout of each the corresponding axis as correction for this bias. Even though this correction does not add up to a substantial correction in the initial short duration of measurement conducted in this study, as the duration of the measurement increases this effect will add up in longer recordings.

2. Linear Movements

The accelerometer of the smartwatch was able to follow a linear trend with an 96% accuracy. This did not differ significantly from the shimmer accelerometer (with 99% accuracy, p = 0.74, a = 0.05, Fig 2).

3. Movement Disorder Simulation

Using solely the severity score resulting from the smartwatch application in combination with the baseline of each volunteer, the amplitude of this blinded movement was classified with 93% accuracy (Table 1).

Subj	*Null*	*Small*	*Medium*	*Large*	*Correct*
1	3.62	13.70	18.45	39.12	100.00
2	1.98	2.00	16.39	60.80	100.00
3	2.93	14.56	27.46	46.68	100.00
4	3.03	25.67	45.57	58.38	77.67
5	2.21	12.01	28.48	82.21	77.67
6	4.17	11.28	19.11	36.06	100.00
7	7.49	38.66	54.84	89.05	100.00
8	1.76	10.31	28.96	62.17	100.00
9	1.44	18.53	35.00	49.05	100.00
10	0.52	8.37	22.68	68.86	100.00
Av	2.92	15.51	29.69	59.24	93.00

Table 1: individual values of the severity scores (a.u.) measured during the simulation of myoclonus in ten volunteers (depicted by the numbers 1-10) and their averages (bottom row). Null indicates no movement, small, medium and large respectively indicate respectively a small, medium and large simulated myoclonus. In the last column, the percentage of correct classifications based on the closest group mean of the baseline classification (at subject level) are depicted.

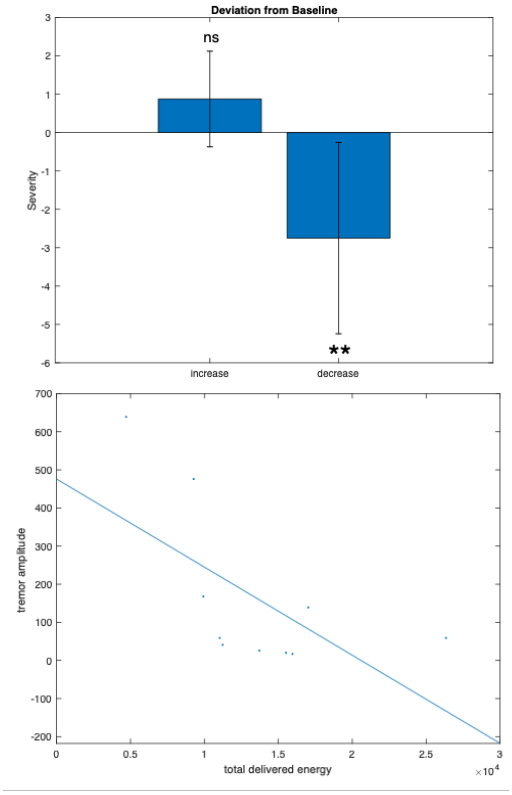


Figure 4: quantifying DBS parameters a. Depiction of myoclonus deterioration (left, increase) and improvement (right, decrease) after the application of DBS compared to baseline in a myoclonus patient expressed in severity score (a.u.). No significant deterioration of myoclonus was present, whereas significant improvement of myoclonus was present ** = $p < 0.01$Relation

between tremor severity (a.u.) and total delivered energy by DBS. An inverse and significant correlation is present.

4. Clinical Testing

In the myoclonus patient, normalised OFF stimulation severity scores were significantly higher than the optimal ON stimulation scores ($p < 0.01$, resp 0 and -2.7 ± 2.5, Fig 4a). The measured differences were subtle and did not result in a subjective reduction in myoclonus.

In the tremor patient, a significant correlation between the measured severity score and total delivered energy was present ($p < 0.05$, cc = 0.64, Fig 4b). Furthermore, severity scores of tremor ratings judged as 'better than previous ' (3/10, 20±3) differed significantly from ratings judged as 'similar or worse than previous' (7/10, 99± 50 $p < 0.05$).

IV. Discussion

In this paper we describe the design, validation and first implementation of a smartwatch application that is able to objectify the severity of hyperkinetic movement disorders. The key findings are that the smartwatch based application has a similar accuracy in detecting linear movements compared to clinically applied accelerometers, with the biggest difference that severity is calculated instantaneously and immediately visible for users. Next to this, the application was also able to accurately detect differences between severities of simulated myoclonus. Finally, the application was able to distinguish between DBS settings with more and less suppression of tremor and myoclonus. For myoclonus this difference was not noticed by the patient but in tremor there was a strong concordance between subjective and objective severity detection. This correlation between subjective and objective ratings opens avenues for closed loop approaches (e.g. 10) which would make it possible to automatically titrate DBS based on smartwatch findings. It should however be emphasised that the application has only been tested in two patients and is not yet fully validated. Furthermore, other challenges were encountered during the design process, which will be discussed in the following sections.

The hardware and approach used in this project has minor shortcomings preventing the usage of these devices to their highest potential yet. Firstly, limitations in the connectivity between smartwatch and smartphone was obstructed when one of them became dormant. Once dormant, data transfer between both devices is postponed until awaking of the device. The measurements performed in this paper were executed properly. However, longer duration measurement (>12 second) are impractical with the current setup. Newer versions of WatchOs have addressed this problem and allow wake times up to 70 seconds. Second, the sensor range of the accelerometers could be improved. The smartwatch used in this paper was equipped with a static 16G range accelerometer. Previous research has shown 6G to be sufficient for recording human movement (11). Since the required accelerometer range was exceeded in our setup, accuracy was lost in the recordings. More accurate and reliable measurement in future research can probably be achieved with a reduced accelerometer range. Finally, since the approach of calculating the integral results in an estimation of the actual distance. The current approach is not suitable for a setting in which patients are also making intentional movements for over a longer period. These systems are nowadays widely available (e.g. Griffiths et al. 2012) but lack the possibility of providing *real-time* estimation of symptom severity.

V. Conclusion

In this paper we describe the first objective, comfortable and instantaneous smartwatch application for measuring movement disorder severity. Excellent performance in classification of simulated myoclonus severity and classification between effective and less effective DBS parameters illustrates the clinical relevance of this application. Future studies in larger patient cohorts should be performed to objectify the added value of the application over care as usual.

References

1. Martinez Manzanera O, Roosma E, Beudel M, Borgemeester R, van Laar T, Maurits N. A method for automatic and objective scoring of bradykinesia using orientation sensors and classification algorithms. IEEE Trans Biomed Eng 2015; 63: 1016–1024.
2. Bhatia KP, Bain P, Bajaj N, Elble RJ, Hallett M, Louis ED, et al. Consensus Statement on the classification of tremors. from the task force on tremor of the International Parkinson and Movement Disorder Society. Movement Disorders 2018; 33: 75–87.
3. Flora ED, Perera CL, Cameron AL, Maddern GJ. Deep brain stimulation for essential tremor: a systematic review. Mov. Disord. 2010; 25: 1550–1559.
4. Anderson DG, Németh AH, Fawcett KA, Sims D, Miler J, Krause A. Deep Brain Stimulation in Three Related Cases of North Sea Progressive Myoclonic Epilepsy from South Africa. Movement Disorders Clinical Practice 2016; 15: 164–5.
5. van Zijl JC, Beudel M, Elting JWJ, de Jong BM, van der Naalt J, van den Bergh WM, et al. The Inter-rater Variability of Clinical Assessment in Post-anoxic Myoclonus. Tremor Other Hyperkinet Mov (N Y) 2017; 7: 470.
6. Zutt R, van Egmond ME, Elting JW, van Laar PJ, Brouwer OF, Sival DA, et al. A novel diagnostic approach to patients with myoclonus. Nature Publishing Group 2015; 11: 687–697.
7. Beudel M, de Geus CM, Leenders KL, de Jong BM. Acceleration bias in visually perceived velocity change and effects of Parkinson's bradykinesia. NeuroReport 2013; 24: 773–778.
8. Abyarjoo F., Barreto A., Cofino J., Ortega F.R. (2015) Implementing a Sensor Fusion Algorithm for 3D Orientation Detection with Inertial/ Magnetic Sensors. In: Sobh T., Elleithy K. (eds) Innovations and Advances in Computing, Informatics, Systems Sciences, Networking and Engineering. Lecture Notes in Electrical Engineering, vol 313. Springer, Cham
9. Jeon H, Kim SK, Jeon B, Park KS. Distance estimation from acceleration for quantitative evaluation of Parkinson tremor. Conf Proc IEEE Eng Med Biol Soc 2011; 2011: 393–396.
10. Malekmohammadi M, Herron J, Velisar A, Blumenfeld Z, Trager MH, Chizeck HJ, et al. Kinematic Adaptive Deep Brain Stimulation for Resting Tremor in Parkinson's Disease. Mov. Disord. 2016; 31: 426–428.
11. M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement." *Physiol. Meas.* vol. 25, no. 2, pp R1-R20, 2004.
12. Griffiths RI, Kotschet K, Arfon S, Xu ZM, Johnson W, Drago J, et al. Automated assessment of bradykinesia and dyskinesia in Parkinson's disease. J Parkinsons Dis 2012; 2: 47–55.