Title: Assessment of Fitbit Charge 4 for sleep stage and heart rate monitoring against polysomnography and during home monitoring in Huntington's disease.

Authors, Degrees: Emer P. Doheny[1], PhD, Klavs Renerts[2], MD, Andreas Braun[2], MD, Esther Werth[2], PhD, Christian Baumann[2], MD, Philipp Baumgartner[2], MD, Philippa Morgan-Jones[3,4], PhD, Monica Busse[3], PhD, Madeleine M. Lowery[1], PhD, Hans H. Jung[2], MD.

Author affiliations:

1. School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland
2. Department of Neurology, University and University Hospital Zurich, Zurich, Switzerland
3. Centre for Trials Research, Cardiff University, Cardiff, Wales, United Kingdom
4. School of Engineering, Cardiff University, Cardiff, United Kingdom

Corresponding author:

Emer Doheny, School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland. Emer.doheny@ucd.ie

Institution where work was performed: University Hospital Zurich and University College Dublin.

All authors have seen and approved this manuscript.

Author declarations:

Financial support: Research supported by EU Joint Programme for Neurodegenerative Disease Research with funding from Alzheimer's Society, Secretary of State for Health and Social Care, Health and Care Research Wales, Public Health Agency Northern Ireland, Jacques and Gloria Gossweiler Foundation, Bundesministerium für Bildung und Forschung, Narodowe Centrum Badań i Rozwoju, Swiss National Science Foundation (32ND30_185548) and Health Research Board (JPND-HSC-2018-003).

Conflict of interest: All authors declare that they have no conflict of interest.

Number of Figures: 4

Number of Tables: 1

Abstract word count: 249

Brief summary word count: 120

Manuscript word count: 3534 (Introduction-Conclusion)

Abstract

Study Objectives: Wearable devices, monitoring sleep stages and heart rate (HR), bring the potential for longitudinal sleep monitoring in patients with neurodegenerative diseases. Sleep quality reduces with disease progression in Huntington's disease (HD). However, the involuntary movements characteristic of HD may affect the accuracy of wrist-worn devices. This study compares sleep stage and heart rate data from the Fitbit Charge 4 (FB) against polysomnography (PSG) in participants with HD.

Methods: Ten participants with manifest HD wore a FB during overnight hospital-based PSG, and for nine of these participants continued to wear the FB for seven nights at home. Sleep stages (30s epochs) and minute-by-minute HR were extracted and compared against PSG data.

Results: FB-estimated total sleep and wake times, and sleep stage times were in good agreement with PSG, with intra-class correlations 0.79-0.96. However, poor agreement was observed for Wake After Sleep Onset, and the number of awakenings. FB detected wake with 68.6±15.5% sensitivity and 93.7±2.5% specificity, rapid eye movement (REM) sleep with high sensitivity and specificity (78.7±31.9%, 95.6±2.3%), and deep sleep with lower sensitivity but high specificity (56.4±28.8%, 95.0±4.8%). FB HR was strongly correlated with PSG, and the mean absolute error between FB and PSG HR data was 1.16 ± 0.42 bpm. At home, longer sleep and shorter wake times were observed compared to hospital data, while percentage sleep stage times were consistent with hospital data.

Conclusions: Results suggest the potential for long-term monitoring of sleep patterns using wrist-worn wearable devices as part of symptom management in HD.

Keywords: Wearable sensor; Home monitoring; PPG; PSG; Neurodegeneration.

Brief Summary

Current Knowledge/Study Rationale: Sleep quality reduces with disease progression in Huntington's disease (HD). Wearable devices enable longitudinal monitoring of sleep staging and heart rate (HR), with potential to support disease management following disease-specific validation.

Study Impact: Sleep stages (wake, REM, light/N1+N2, deep/N3) and HR from Fitbit Charge 4 (FB) were compared with polysomnography (PSG) in participants with manifest HD. FB accurately estimated overall sleep stage times, with good sensitivity and specificity to REM. Poor detection of awakenings, and poor sensitivity to deep/N3 sleep were observed relative to PSG. At home, participants slept for longer compared to PSG, while percentage sleep stage times were consistent with PSG. Wearable devices, such as FB, have potential to monitor overall sleep metrics in HD.

Introduction

Huntington's disease (HD) is a rare, autosomal-dominant, neurodegenerative disorder caused by an expanded CAG repeat in the Huntingtin gene. As in other neurodegenerative diseases[1], sleep dysfunction is highly prevalent in HD, with patients experiencing insomnia, increased sleep onset latency, decreased total sleep time, frequent awakenings, restlessness and excessive daytime sleepiness[2–5]. Polysomnography (PSG) has also revealed reduced time in deep sleep, more sleep spindles and more time in non-rapid eye movement (REM) sleep in HD[6,7]. These sleep issues arise from an early stage in HD, often preceding the onset of motor symptoms [8–10]. The severity and frequency of sleep disorders have also been associated with clinical and radiological disease progression[11], while the number of CAG repeats is linked to an early onset of sleep disturbance[12]. Dspite this evidence and the potential of sleep as a modifiable risk factor for neurodegeneration[13–15], sleep is not typically monitored in HD.

Overnight PSG in an AASM-accredited [16] sleep lab is the gold standard for sleep monitoring, and is necessary to diagnose many sleep disorders and to provide a complete clinical assessment of sleep. However, PSG is expensive and often not representative of natural sleep patterns due to multiple wired sensors, and bedtimes and wake-up times determined by hospital routines. Wearable sleep monitors, particularly recent dual-sensor devices featuring both inertial and photoplethysmography (PPG) sensors, offer the potential to monitor sleep stages and heart rate (HR) during sleep in a more natural environment, and to examine changes in sleep patterns over time due to disease progression, treatments or interventions [16–19]. However, the involuntary and nocturnal movements associated with HD may lead to reduced performance of wrist-worn devices in this cohort[3,20]. Therefore, disease-specific validation against PSG is necessary to establish the accuracy of wrist-worn sleep monitoring devices before home-based clinical monitoring is possible.

Few previous studies have examined sleep using single-sensor Fitbit devices, detecting sleep based on movement, in HD[21,22]. However, several studies have validated dual-sensor Fitbit devices against PSG in healthy adults [23–25]. Fitbit Charge 2 has been compared to PSG in healthy adults, with a significant overestimation of total sleep time (TST), and limitations in detection of deep sleep reported,

while no difference relative to PSG was reported for REM sleep time[23]. Fitbit Charge 3 was previously validated against PSG in healthy adolescents and adolescents with insomnia[24], with epoch-by-epoch analysis revealing sensitivities to sleep stages ranging from 59% for deep sleep to 78% for light sleep detection. Fitbit Charge 2 and Fitbit Alta HR were compared to PSG in adults with obstructive sleep pane, with both Fitbit models providing significantly different values for all sleep metrics with the exception of REM sleep time [25]. Older single-sensor Fitbit models, not featuring an optical HR sensor (Fitbit One, Fitbit Flex, Fitbit Charge), have also been validated against PSG in healthy adults [26–28] and in HD[21], however sleep stages estimated using these devices were less accurate than the more recent dual-sensor models. Sleep metrics from Fitbit One were compared with PSG in 7 patients with HD, three with manifested motor symptoms, with a >60-minute overestimation of total sleep time (TST) reported[21], however that study concluded that the device may still be sufficient to examine overall sleep-wake patterns in HD[21]. Another study used Fitbit One to monitor at-home sleep in a group of 42 HD gene carriers (20 with manifest HD), and reported an association between increased time in bed and reduced cognitive functioning in HD[22]. However, four-stage (wake, REM, light and deep) sleep staging or HR monitoring using wearable sensors, including any Fitbit device, have not previously been validated against PSG or examined in the home with participants with HD.

Heart rate, a vital sign, can be used to monitor overall health and to alert clinicians to values outside the expected range. The accuracy of heart rate data estimated using PPG sensors on the wrist may by affected by movement artifacts in the PPG signal due to involuntary movements, or chorea, associated with Huntington's disease. Signal artifacts due to involuntary movements may also influence the accuracy of sleep stage estimates, also derived from the PPG signal. Previous studies have compared HR estimated using Fitbit devices with ECG in healthy adults, showing good performance for the Fitbit Charge 4 during rest and physical activity[29], and for the Fitbit ChargeHr during wake, REM and non-REM sleep[30], and under free-living conditions[31]. However, minute-by-minute HR estimated by any Fitbit model has not previously been validated in HD or in any pathological cohort.

The aim of this study was to establish the accuracy of sleep stage and HR data provided by Fitbit Charge 4 compared to PSG in participants with manifest HD. Home-monitoring was also conducted to examine natural sleep patterns in this cohort.

Methods

Participants and protocol

Ten participants (5 female) with a genetically confirmed diagnosis of Huntington's disease (HD) participated in this study. All participants had motor manifestations of the disease, with Unified HD Rating Scale total motor scores (UHDRS-TMS) in the range 5-51 points (/124), were self-ambulatory, home-dwelling and required no specialized care at the time of the study. Approval was obtained from the Ethics Committee of the Canton of Zurich (Kantonale Ethikkommission Zürich). Each participant wore a Fitbit Charge 4 (FB) on their non-dominant wrist during overnight PSG (Embla N7000, RemLogic v3.2, ResMed, Australia) in a hospital-based sleep laboratory. Nine of the ten participants continued to wear the FB in free-living conditions for a further seven days and nights.

Polysomnography data

Overnight PSG (Embla N7000, RemLogic v3.2, ResMed, Australia) was performed in a hospital-based sleep laboratory. PSG recordings commenced at "Lights Out" and ended at "Lights On" in the sleep laboratory.

Sleep stages were scored in 30 s epochs by an expert sleep physiologist following the criteria set by AASM[16]. For comparison with the wearable device, stage N1 and N2 were merged and referred to as light sleep, and stage N3 was referred to as deep sleep.

HR data reported by the PSG software based on three lead electrocardiography (ECG recorded at 512 Hz), were extracted at 5 Hz, and down-sampled to 1 sample per 60 s for comparison with FB data.

The prevalence of sleep disorders or events associated with nocturnal movements or arousals were documented for the PSG night. Daytime sleepiness was examined using the Epworth Sleepiness Scale (ESS, scored out of 24 points), and periodic limb movement (PLMS, events/hr) and apnea hypopnea index (AHI, events/hr) were assessed using PSG data.

Wearable device data

Participants downloaded the FB mobile app to their personal phone and were assigned a FB account linked to their study identifier. Fitbit sensitivity was set to "normal", the default setting in the Fitbit mobile app. Participants synchronised their FB mobile app before the overnight PSG study, and also after each sleep.

A bespoke platform (AthenaCX, AthenaCX.com, In the Wild Research Ltd., Dublin, Ireland) was used to integrate FB data for all participants. A custom python script was used to extract FB sleep stage (30 s epochs) and HR (60 s epochs) data from this platform as .json files. It is not possible to access raw FB data, and this was the maximum resolution for FB sleep stage and HR data.

As FB sleep stage data began automatically when FB detected sleep onset, wake epochs were added to FB sleep stage data for the time after PSG "lights off" and prior to the first sleep epoch detected by FB, similar to previous studies[16]. FB HR data corresponding to each FB-detected sleep, and overlapping with PSG HR, were compared with PSG. Extracted sleep stage and HR data were then analysed using Matlab (Mathworks, Natick, MA, USA).

Sleep metrics

For each sleep, for PSG and FB, the following sleep metrics were estimated: TST (mins): total sleep time; TWT (mins): total wake time; REM (% of TST): total REM sleep time; light sleep (% of TST): total stage N1 and N2 sleep time for PSG, total light sleep time for FB; deep sleep (% of TST): total stage N3 sleep time for PSG, total deep sleep time for FB; WASO (mins): total wake after sleep onset; awakenings (/hr): average number of awakenings per hour.

Statistical analysis

Validation against PSG:

To assess the agreement between sleep metrics and HR reported by FB and PSG, Bland Altman analysis was performed and the intra-class correlation, ICC(2,1), was calculated[32]. The Wilcoxon rank sum test was performed to compare the PSG and FB estimates for each sleep metric.

Epoch-by-epoch detection of wake, REM, light and deep sleep by FB were individually assessed using the area under the receiver operating characteristic (AUC ROC), sensitivity and

specificity for each PSG night. Sensitivity, or the true positive rate, was calculated as the number of epochs correctly classified (true positives) as a particular class (wake, REM, light or deep) as a percentage of the total number of epochs classified as that class (true positives plus false negatives). Specificity, or one minus the false positive rate, was calculated as the number of true negatives as a percentage of the total number of epochs not classified as that class (true negatives plus false positives). The four-stage agreement between PSG and FB sleep staging data was assessed using the overall accuracy and Cohen's Kappa.

Minute-by-minute HR data estimated by FB were compared with PSG by calculating the bias and mean absolute error (MAE) for each night, for wake epochs only and for each sleep stage (REM, light and deep sleep). To evaluate the ability of FB to capture HR trends during each sleep stage, FB and PSG HR signals were compared using Pearson's correlation coefficient (r).

Relationship between sleep and motor impairment in HD:

Pearson partial correlation coefficients (r) were calculated to examine the strength of the linear relationships between the UHDRS-total motor score and each sleep metric, estimated using FB and PSG in the sleep lab, and using FB during the subsequent home data. Partial correlation analyses were performed to control for the effect of age.

Effect of setting and device on sleep metrics:

The effects of setting (home or sleep lab) and device (FB or PSG) on each sleep metric were examined using a linear mixed model.

Correction for multiple comparisons

The threshold for significance was initially defined as $p = 0.05$. To avoid family wise Type I errors due to multiple comparisons, a Bonferroni correction was applied by dividing the threshold for significance by the number of comparisons. FB and PSG data were compared for seven sleep metrics (TST, TWT, REM, light, deep, WASO, awakenings). Five sleep metrics (TST, TWT, REM, light, deep) estimated by FB were found to have good agreement between PSG, with the effect of UHDRS-TMS and the effect

of setting (home or sleep lab) examined for these metrics. Additionally the effect of UHDRS-TMS on participant age was examined. In total, 18 comparisons were conducted, resulting in a corrected significance threshold of $p = 0.0028$.

Results

A summary of demographics and clinical data for the study participants are presented in Table I. One participant was deemed to have REM behavioural disorder, five participants had PLMS >15 /hour, and 2 participants had mild obstructive sleep apnea (AHIs 16.6 and 16.7 15 events/hr). One participant did not take part in the 7-day home study, see Table I.

Validation against PSG

PSG and FB hypnograms and HR traces for two representative participants with HD are presented in Figure 1. Figure 1a and 1b illustrate overnight data for a participant for whom a poor fit was observed between FB and PSG sleep staging, with a 4-stage Cohen's Kappa of 0.4, with low correlation between FB and PSG heart rate traces also observed for that participant, $r = 0.49$. In contrast, Figure 1c and 1d illustrate data for a participant for whom a good fit was observed between FB and PSG sleep staging data, with a Cohen's Kappa of 0.71, and highly correlated FB and PSG heart rate traces, $r = 0.9$.

The agreement between a range of sleep and HR metrics estimated by FB and PSG is presented in Figure 2, with intra class correlation (ICC(2,1)), biases and MAE presented for each sleep metric. FB-estimated TST and TWT were in excellent agreement with PSG (ICC > 0.9), with narrow limits of agreement. Good agreement (ICC 0.79-0.85) was observed for light, REM and deep sleep times estimated by FB and PSG, however wide limits of agreement were observed for these metrics. Poor agreement was observed for WASO and awakenings (/hr). When compared using the Wilcoxon rank sum test, no significant differences were observed between PSG and FB for any of the sleep metrics.

The agreement between FB epoch-by-epoch sleep stage data and PSG is presented in Figure 3, including sensitivity, specificity and accuracy for detection of each sleep stage. The overall four-stage Cohen's Kappa was 0.59 ± 0.10 across all participants. The area under the ROC was 0.81 ± 0.08

for detection of wake, 0.86 ± 0.16 for REM sleep, 0.76 ± 0.06 for light sleep and 0.76 ± 0.14 for deep sleep.

Compared to PSG, FB minute-by-minute HR was underestimated by 0.11 ± 0.40 bpm during overall sleep (including wake), by 0.03 ± 0.66 bpm during wake, by 0.13 ± 0.20 bpm during REM sleep, underestimated by 0.13 ± 0.23 bpm during light sleep, and by 0.47 ± 0.61 bpm during deep sleep. Across all participants, the MAE in FB HR data compared to PSG was 1.16 ± 0.42 bpm for each overall sleep (including wake), 1.37 ± 0.53 bpm during wake, 1.29 ± 0.81 bpm during REM sleep, 1.07 ± 0.46 bpm during light sleep and 1.15 ± 0.54 bpm during deep sleep.

Pearson's linear correlation (r) between FB and PSG HR traces was 0.78 ± 0.16 for all data from "lights out" to "lights on", 0.77 ± 0.13 during wake, 0.68 ± 0.29 during REM, 0.79 ± 0.19 during light, and 0.72 ± 0.25 during deep sleep. All correlations were significant ($p < 0.001$).

Relationship between sleep and motor impairment in HD:

Older age was significantly associated with increased UHDRS-TMS ($r = 0.86$, $p < 0.001$), or more severe motor impairment. Based on both home and sleep lab recordings from FB and PSG, light sleep time (%) significantly increased and deep sleep time (%) significantly decreased with increasing motor impairment (UHDRS-TMS), see Figure 4. WASO and awakenings were not included in this analysis due to poor agreement between FB and PSG.

Effect of setting on sleep metrics:

TST at home, estimated using FB, was significantly longer than TST estimated using PSG and FB in the sleep lab (PSG: 289.90 ± 88.58 min, FB during PSG: 283.05 ± 117.76 min, FB at home: 316.88 ± 163.10 min, $p<0.001$). Additionally, TWT estimated using FB at home was significantly shorter than TWT estimated using FB or PSG in the sleep lab (PSG: 133.85 ± 89.05 min, FB during PSG: 140.75 ± 120.12 min, FB at home: 66.60 ± 63.68 min, $p<0.001$).

When sleep staging was examined, no significant differences were observed between home data (FB) and sleep lab data (FB and PSG) for percentage REM sleep (PSG: 14.35 ± 7.29 %, FB during PSG:

16.69 ± 7.05 %, FB at home: 20.83 ± 6.67 %, p = 0.06), percentage light sleep (PSG: 67.35 ± 12.09 %, FB during PSG: 66.61 ± 12.39 %, FB at home: 64.15 ± 11.73 %, p = 0.29) or for percentage deep sleep (PSG: 18.30 ± 10.20 %, FB during PSG: 19.09 ± 5.18 %, FB at home: 16.13 ± 5.98 %, p = 0.43).

Discussion

Dual sensor wearable devices, which monitor movement using inertial sensors and heart rate using PPG, may offer a low-cost method to monitor sleep and heart rate activity at home in HD, and to track changes over time. In this study, sleep stages and HR data obtained from one such device, the Fitbit Charge 4, were compared against PSG in participants with manifest HD. To examine sleep under more natural conditions, a home-based study was also conducted.

Fitbit Charge 4 accurately estimated total sleep time and time in each sleep stage, compared with PSG in ten participants with manifest HD, with a wide range of motor impairment (UHDRS-TMS 5-51). One previous study has validated a Fitbit device (Fitbit One) against PSG in participants with HD, in a cohort of 7 participants (3 with manifest HD). That earlier device, which did not feature a PPG sensor, and did not report sleep staging or HR, overestimated TST by 88 minutes and underestimated WASO by 39 minutes[21]. In healthy participants, previous studies have also reported superior performance for sleep metrics provided by dual-sensor Fitbit devices[23,24], compared with older devices which do not monitor HR[26]. A previous study [23] in 44 healthy adults, 9 with period limb movement of sleep, reported that Fitbit Charge 2 overestimated TST by 9 min, light sleep by 34 min, and deep sleep by 24 min, with no difference observed between PSG and FB for WASO and REM. In the current study, no significant differences between Fitbit Charge 4 and PSG were observed for any sleep metric (Figure 2). That previous study also performed an epoch-by-epoch analysis and reported that Fitbit Charge 2 had a sensitivity of 96% for sleep detection, 61% for wake, 74% for REM sleep, 81% for light sleep and 49% for deep sleep compared with PSG[23]. These results in healthy individuals are consistent with the sleep stage sensitivities observed in this study for Fitbit Charge 4 compared to PSG in participants with HD (Figure 3). The findings of the current study are encouraging for FB estimation of TST, TWT,

and overall time in each sleep stage, however estimations of awakening frequency and WASO were less accurate, related to error in sleep onset detection by FB (Figure 1).

The current study also evaluated heart rate estimated by the Fitbit Charge 4 in individuals with HD, and observed a low bias of -0.11 ± 0.40 bpm and strong correlation between FB and PSG heart rate trends ($r = 0.78 \pm 0.16$), consistent with previous findings in healthy participants. The accuracy of HR estimated by Fitbit Charge 2 was previously examined during exercise, with higher biases reported compared with ECG in 15 healthy adults during cycling[33]. A study comparing HR from Fitbit Charge 2 and an ambulatory ECG over 24 hours in one healthy participant, reported 91% agreement assessed using concordance class correlation[34]. Another study validated HR from Fitbit Charge HR against portable PSG in 25 healthy adults[30], reporting generally accurate minute-by-minute HR estimates, with a bias of -0.66 bpm, and Pearson's correlation coefficient of 0.84 ± 0.11.

A significant negative linear correlation was observed between UHDRS-TMS and percentage REM sleep in the present study (Figure 4), consistent with a previous PSG study in 30 adults with HD[2] which also found that increased motor impairment was associated with a lower percentage of REM sleep. In that study [2], no other sleep metrics were significantly correlated with UHDRS-TMS. However, a significant positive linear correlation between motor impairment (UHDRS-TMS) and percentage light sleep was additionally observed in the current study (Figure 4). As expected, age was significantly correlated with UHDRS-TMS, and all correlations were estimated using a partial correlation analysis to control for the effect of age on each relationship.

At home, participants experienced longer total sleep times, and shorter total wake times, compared to the sleep lab. These differences may be due to sleep lab protocols, where "lights off" and "lights on" times are guided by hospital routines rather than natural sleep patterns, discomfort due to PSG sensors, or a first-night effect[35] causing reduced sleep in the sleep lab. Previously, home and sleep lab data for an alternative wearable device were compared in healthy adults, with no significant differences in TST or WASO between home and lab observed[17]. Reduced sleep quality in Huntington's disease, or the behavioural and cognitive changes which occur in HD, may explain this contrasting finding in healthy adults. In this study, percentage times in each sleep stage were consistent between home and sleep lab recordings. These results indicate that home monitoring using dual-sensor wearable

devices may provide clinically meaningful sleep pattern data which could support symptom management. While not providing PSG-quality clinical sleep data, and not a replacement for PSG for many application, these devices perform well in sleep stage classification and bring the possibility of monitoring natural sleep patterns at home over multiple nights.

Conclusion

The results suggest that a dual-sensor wrist-worn device, such as Fitbit Charge 4, could be used to accurately monitor HR and overall sleep patterns at home in individuals with manifest HD (UHDRS-TMS 5-51). In this cohort, Fitbit Charge 4 provided accurate estimates of total sleep times and time in each sleep stage, with sensitive and specific detection of REM sleep compared to the gold standard PSG. Detection of deep sleep was less sensitive, but highly specific, compared to PSG. However, FB did not perform well in the detection of awakenings, or WASO estimation. When sleep at home was compared to sleep during hospital-based PSG, longer sleep times were observed, while percentage time in each sleep stage was consistent across setting. To further investigate the effect of motor symptoms on device accuracy, future research should include larger cohort sizes, and a wider range of motor impairment.

Acknowledgements

This study was conducted as part of the DOMINO-HD consortium study, which involves 7 European partners coordinated by Cardiff University, through the EU Joint Programme – Neurodegenerative Disease Research.

Research supported by EU Joint Programme for Neurodegenerative Disease Research with funding from Alzheimer's Society, Secretary of State for Health and Social Care, Health and Care Research Wales, Public Health Agency Northern Ireland, Jacques and Gloria Gossweiler Foundation, Bundesministerium für Bildung und Forschung, Narodowe Centrum Badań i Rozwoju, Swiss National Science Foundation (32ND30_185548) and Health Research Board (JPND-HSC-2018-003).

Abbreviations

AASM: American Academy of Sleep Medicine.

AHI: Apnea Hypopnea Index.

AUC ROC: Area Under the Curve of the Receiver Operating Characteristic

BMI: Body Mass Index.

CAG: Cytosine-Adenine-Guanine trinucleotide.

DBS: Disease Burden Score [36] = Age*(CAG-35.5);

ECG: Electrocardiography.

ESS: Epsworth Sleepiness Scale (scored out of 24 points).

FB: Fitbit Charge 4.

HD: Huntington's Disease.

HR: Heart Rate.

ICC: Intra-Class Correlation.

MAE: Mean Absolute Error.

PLMS: Periodic Limb Movement of Sleep.

PPG: Photoplethysmography.

PSG: Polysomnography.

REM: Rapid Eye Movement sleep

SD: Standard Deviation.

Sens = Sensitivity.

Spec = Specificity.

TST = Total Sleep Time.

TWT = Total Wake Time.

UHDRS-TMS = Unified Huntington's Disease Rater Scale Total Motor Score.

WASO = Wake After Sleep Onset.

References

1. Bliwise DL. Sleep disorders in Alzheimer's disease and other dementias. *Clin Cornerstone*. 2004;6(1 SUPPL. 1). doi:10.1016/S1098-3597(04)90014-2

2. Piano C, Losurdo A, Marca G Della, et al. Polysomnographic findings and clinical correlates in Huntington disease: A cross-sectional cohort study. *Sleep*. 2015;38(9). doi:10.5665/sleep.4996

3. Neutel D, Tchikviladze M, Charles P, et al. Nocturnal agitation in Huntington disease is caused by arousal-related abnormal movements rather than by rapid eye movement sleep behavior disorder. *Sleep Med*. 2015;16(6):754-759. doi:10.1016/j.sleep.2014.12.021

4. Yadav R, Jagadish A, Sheshagiri D, et al. Sleep and Polysomnographic abnormalities in Indian patients with Huntington's disease. *Parkinsonism Relat Disord*. 2020;79. doi:10.1016/j.parkreldis.2020.06.390

5. Emser W, Brenner M, Stober T, Schimrigk K. Changes in nocturnal sleep in Huntington's and Parkinson's disease. *J Neurol*. 1988. doi:10.1007/BF00314313

6. Wiegand M, Möller AA, Lauer CJ, et al. Nocturnal sleep in huntington's disease. *J Neurol*. 1991. doi:10.1007/BF00314781

7. Hansotia P, Wall R, Berendes J. Sleep disturbances and severity of Huntington's disease. *Neurology*. 1985;35(11). doi:10.1212/wnl.35.11.1672

8. Morton AJ. Circadian and sleep disorder in Huntington's disease. *Exp Neurol*. 2013. doi:10.1016/j.expneurol.2012.10.014

9. Soneson C, Fontes M, Zhou Y, et al. Early changes in the hypothalamic region in prodromal Huntington disease revealed by MRI analysis. *Neurobiol Dis*. 2010;40(3). doi:10.1016/j.nbd.2010.07.013

10. Lazar AS, Panin F, Goodman AOG, et al. Sleep deficits but no metabolic deficits in premanifest Huntington's disease. *Ann Neurol*. 2015. doi:10.1002/ana.24495

11. Maffi S, Scaricamazza E, Migliore S, Casella M, Ceccarelli C, Squitieri F. Sleep Quality and Related Clinical Manifestations in Huntington Disease. *J Pers Med*. 2022;12(6). doi:10.3390/jpm12060864

12. Ogilvie AC, Nopoulos PC, Schultz JL. Sleep disturbances by disease type and stage in Huntington's disease. *Parkinsonism Relat Disord*. 2021;91. doi:10.1016/j.parkreldis.2021.08.011

13. Pillai JA, Leverenz JB. Sleep and Neurodegeneration: A Critical Appraisal. *Chest*. 2017;151(6). doi:10.1016/j.chest.2017.01.002

14. Videnovic A, Golombek D. Circadian and sleep disorders in Parkinson's disease. *Exp Neurol*. 2013;243. doi:10.1016/j.expneurol.2012.08.018

15. Goodman AOG, Barker RA. How vital is sleep in Huntington's disease? *J Neurol*. 2010. doi:10.1007/s00415-010-5517-4

16. Lee XK, Chee NIYN, Ong JL, et al. Validation of a consumer sleep wearable device with actigraphy and polysomnography in adolescents across sleep opportunity manipulations. *Journal of Clinical Sleep Medicine*. 2019;15(9). doi:10.5664/jcsm.7932

17. Grandner MA, Bromberg Z, Hadley A, et al. Performance of a multisensor smart ring to evaluate sleep: in-lab and home-based evaluation of generalized and personalized algorithms. *Sleep*. 2023;46(1). doi:10.1093/sleep/zsac152

18. Chinoy ED, Cuellar JA, Huwa KE, et al. Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep*. 2021;44(5). doi:10.1093/sleep/zsaa291

19. Dunn J, Kidzinski L, Runge R, et al. Wearable sensors enable personalized predictions of clinical laboratory measurements. *Nat Med*. 2021;27(6). doi:10.1038/s41591-021-01339-0

20. Ranjan S, Kohler S, Harrison MB, Quigg M. Nocturnal Post-arousal Chorea and Repetitive Ballistic Movement in Huntington's Disease. *Mov Disord Clin Pract*. 2016;3(2). doi:10.1002/mdc3.12258

21. Maskevich S, Jumabhoy R, Dao PDM, Stout JC, Drummond SPA. Pilot Validation of Ambulatory Activity Monitors for Sleep Measurement in Huntington's Disease Gene Carriers. *J Huntingtons Dis*. 2017. doi:10.3233/JHD-170251

22. McLaren B, Drummond SPA, Glikmann-Johnston Y, et al. Greater Time in Bed and Less Physical Activity Associate With Poorer Cognitive Functioning Performance in Huntington's Disease. *Neuropsychology*. 2021;35(6). doi:10.1037/neu0000757

23. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. *Chronobiol Int*. 2018. doi:10.1080/07420528.2017.1413578

24. Menghini L, Yuksel D, Goldstone A, Baker FC, de Zambotti M. Performance of Fitbit Charge 3 against polysomnography in measuring sleep in adolescent boys and girls. *Chronobiol Int*. 2021. doi:10.1080/07420528.2021.1903481

25. Moreno-Pino F, Porras-Segovia A, López-Esteban P, Artés A, Baca-García E. Validation of fitbit charge 2 and fitbit alta hr against polysomnography for assessing sleep in adults with obstructive sleep apnea. *Journal of Clinical Sleep Medicine*. 2019;15(11). doi:10.5664/jcsm.8032

26. Hakim M, Miller R, Hakim M, et al. Comparison of the Fitbit® Charge and polysomnography for measuring sleep quality in children with sleep disordered breathing. *Minerva Pediatrics*. 2018. doi:10.23736/s0026-4946.18.05333-1

27. Baroni A, Bruzzese JM, Di Bartolo CA, Shatkin JP. Fitbit Flex: an unreliable device for longitudinal sleep measures in a non-clinical population. *Sleep and Breathing*. 2016. doi:10.1007/s11325-015-1271-2

28. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *J Affect Disord*. 2017. doi:10.1016/j.jad.2017.04.030

29. Nissen M, Slim S, Jäger K, et al. Heart Rate Measurement Accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: Device Evaluation Study. *JMIR Form Res*. 2022;6(3). doi:10.2196/33635

30. Benedetti D, Olcese U, Frumento P, et al. Heart rate detection by Fitbit ChargeHR™: A validation study versus portable polysomnography. *J Sleep Res*. 2021;30(6). doi:10.1111/jsr.13346

31. Gorny AW, Liew SJ, Tan CS, Müller-Riemenschneider F. Fitbit charge HR wireless heart rate monitor: Validation study conducted under free-living conditions. *JMIR Mhealth Uhealth*. 2017;5(10). doi:10.2196/mhealth.8233

32. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull*. 1979;86(2). doi:10.1037/0033-2909.86.2.420

33. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the fitbit charge 2 for monitoring heart rate. *PLoS One*. 2018. doi:10.1371/journal.pone.0192691

34. Nelson BW, Allen NB. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: Intraindividual validation study. *JMIR Mhealth Uhealth*. 2019;7(3). doi:10.2196/10828

35. Newell J, Mairesse O, Verbanck P, Neu D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. *Psychiatry Res*. 2012;200(2-3). doi:10.1016/j.psychres.2012.07.045

36. Orth M, Schwenke C. Age-at-onset in Huntington disease. *PLoS Curr*. 2011. doi:10.1371/currents.RRN1258

Figure Titles and Captions

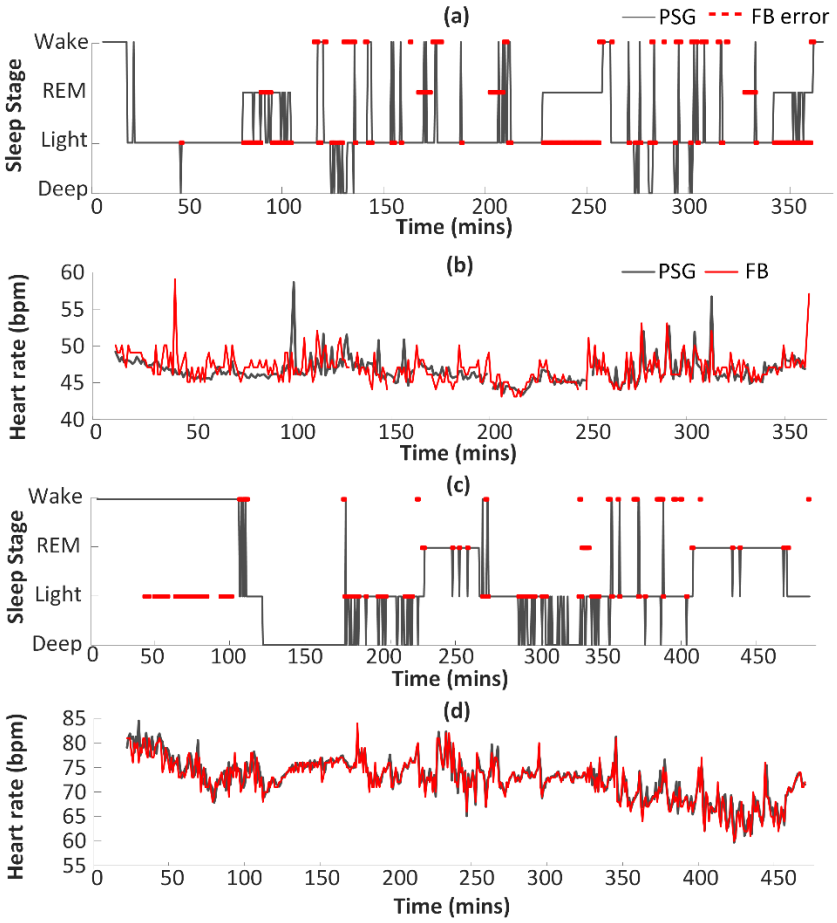

Figure 1.

Title: Hypnograms and corresponding HR traces for two sample participants with HD.

Caption: Hypnograms and corresponding HR traces for two sample participants with HD. PSG data are shown in dark grey, while FB data (for epochs where sleep staging disagreed with PSG) are shown in red. (a,b) Participant 1; Age =56 years, UHDRS-TMS = 51, DBS = 532, Kappa (4 stage classification) = 0.40, r (HR) = 0.49. (c,d); Participant 2: Age = 38, UHDRS-TMS = 5, DBS = 247, Kappa (4 stage classification) = 0.71, r (HR) = 0.90.

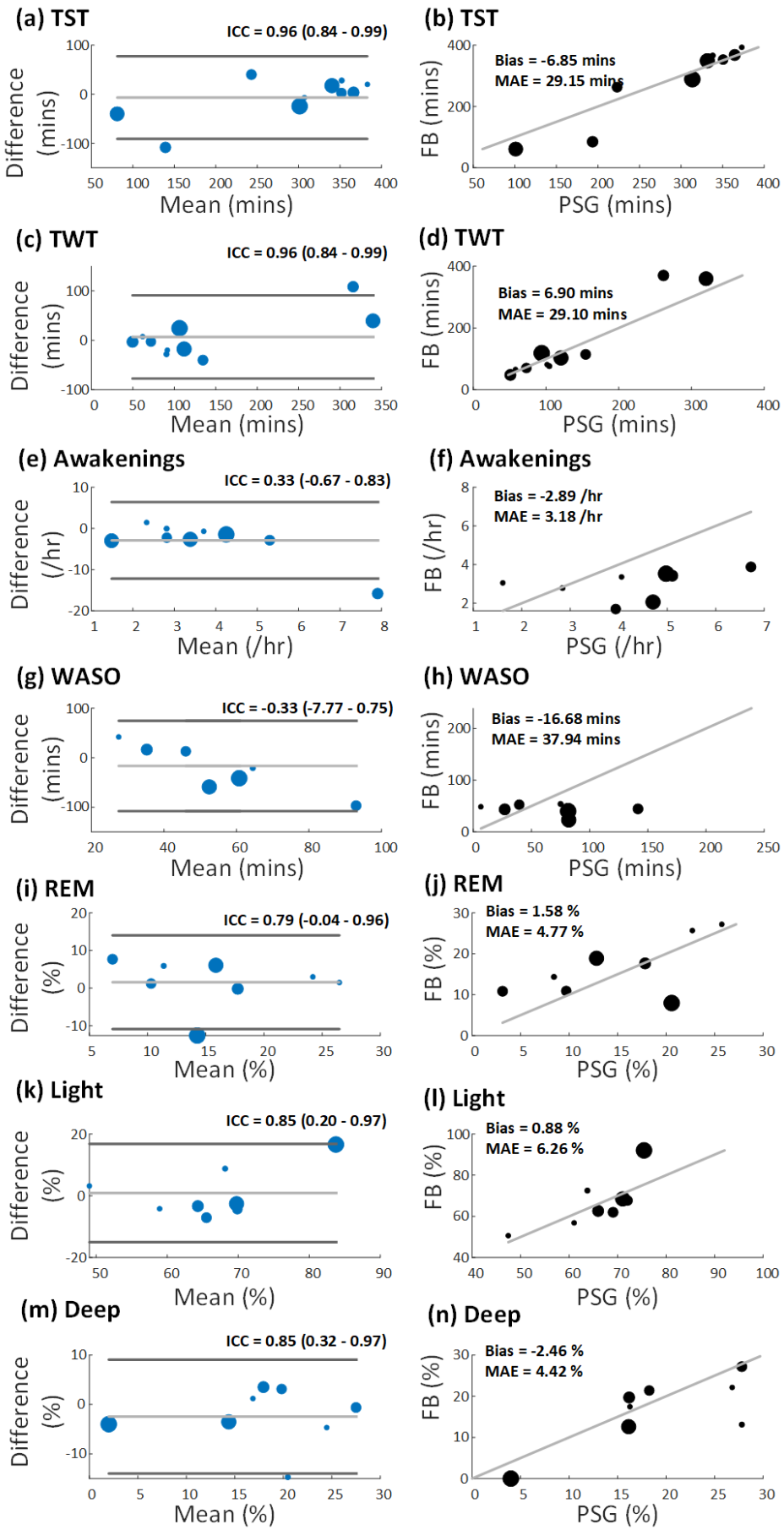
(a) TST
ICC = 0.96 (0.84 - 0.99)
Difference (mins)
Mean (mins)
(b) TST
Bias = -6.85 mins
MAE = 29.15 mins
FB (mins)
PSG (mins)
(c) TWT
ICC = 0.96 (0.84 - 0.99)
Difference (mins)
Mean (mins)
(d) TWT
Bias = 6.90 mins
MAE = 29.10 mins
FB (mins)
PSG (mins)
(e) Awakenings
ICC = 0.33 (-0.67 - 0.83)
Difference (/hr)
Mean (/hr)
(f) Awakenings
Bias = -2.89 /hr
MAE = 3.18 /hr
FB (/hr)
PSG (/hr)
(g) WASO
ICC = -0.33 (-7.77 - 0.75)
Difference (mins)
Mean (mins)
(h) WASO
Bias = -16.68 mins
MAE = 37.94 mins
FB (mins)
PSG (mins)
(i) REM
ICC = 0.79 (-0.04 - 0.96)
Difference (%)
Mean (%)
(j) REM
Bias = 1.58 %
MAE = 4.77 %
FB (%)
PSG (%)
(k) Light
ICC = 0.85 (0.20 - 0.97)
Difference (%)
Mean (%)
(l) Light
Bias = 0.88 %
MAE = 6.26 %
FB (%)
PSG (%)
(m) Deep
ICC = 0.85 (0.32 - 0.97)
Difference (%)
Mean (%)
(n) Deep
Bias = -2.46 %
MAE = 4.42 %
FB (%)
PSG (%)

Figure 2.

Title: Agreement between PSG and Fitbit for each sleep metric.

Caption: Bland Altman plots, and scatter plots, illustrating the agreement between PSG and Fitbit for each sleep metric (TST = total sleep time; TWT = total wake time; Awakenings = average number of awakenings per hour; WASO = wake after sleep onset; REM = rapid eye movement sleep stage; Light = light sleep, or stage N1 and N2 sleep; Deep = deep sleep or slow wave sleep stage) . For Bland Altman plots, intraclass correlation coefficients, ICC(2,1), are presented with the lower and upper limits of the 95% confidence interval. The grey horizontal line indicates the bias, with a negative bias indicating that FB underestimated a particular sleep metric relative to PSG, and a positive bias indicating an overestimation by FB for that metric. On each scatter plot, the bias and MAE are presented in text and the grey diagonal line represents the ideal fit if FB estimated each point with no error relative to PSG. The size of each data point is scaled by the UHDRS-TMS for each participant (Table I), with larger markers indicating higher UHDRS-TMS (increased motor impairment).

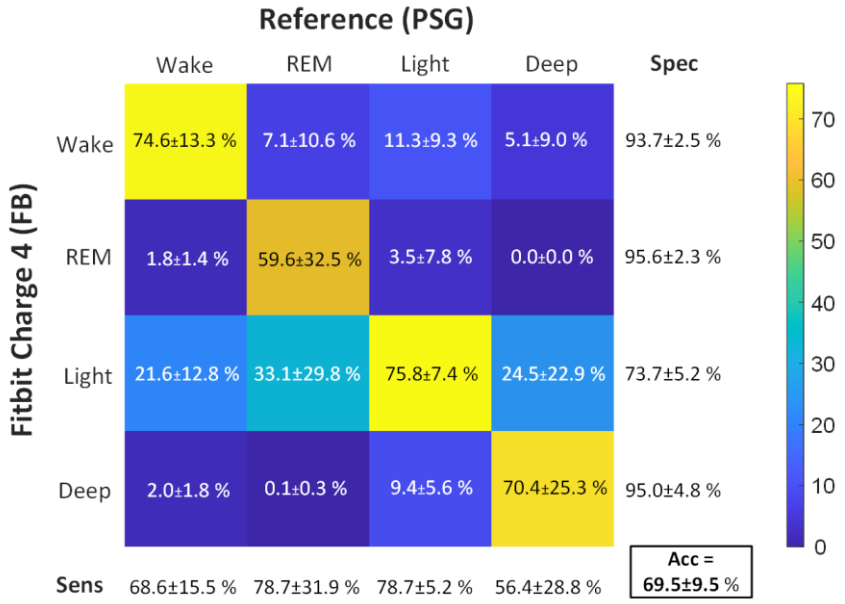

Figure 3.

Title: Confusion matrix for sleep stage detection.

Caption: Confusion matrix for sleep stages for all epochs for all participants during PSG recordings. Rows correspond to FB sleep stages and columns correspond to the PSG sleep stages. Each cell contains the number of epochs classified as each sleep stage, as a percentage of the reference (PSG) number of epochs of that sleep stage. The mean ± standard deviation across all participants is presented. The diagonal cells correspond to epochs that are correctly classified, with the values on the diagonal equal to the accuracy for each metric. The values in the off-diagonal cells correspond to incorrectly classified epochs. The specificity (Spec) is shown on the far right of the plot for each sleep stage, and the row at the bottom shows the sensitivity (Sens), or true positive rate, for each sleep stage. The overall classification accuracy is included at the bottom right.

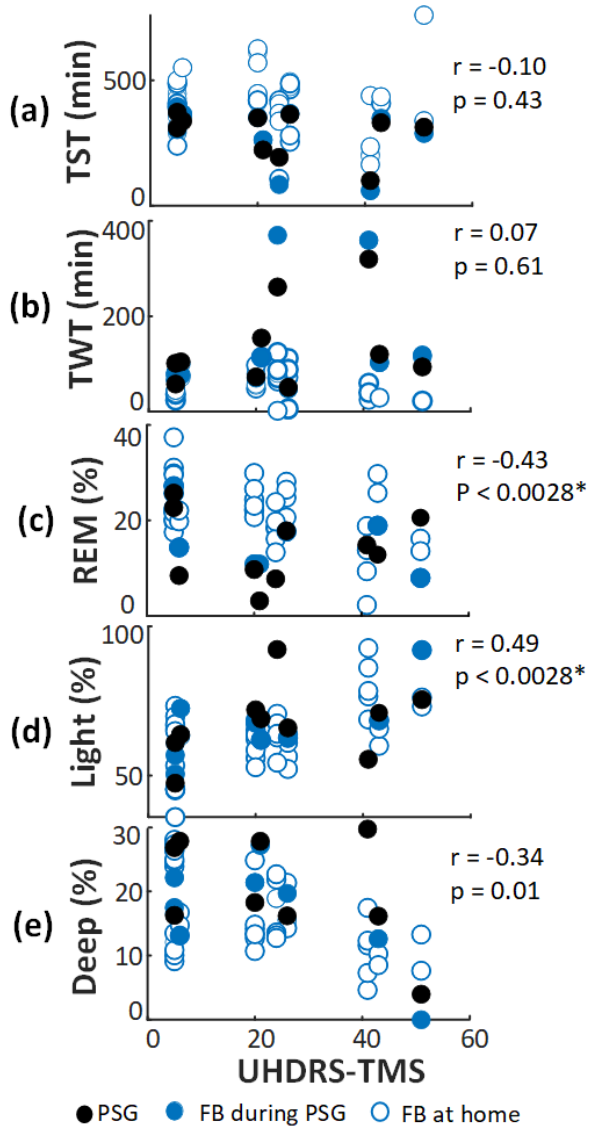

Figure 4.

Title: The relationship between UHDRS total motor score and time in each sleep stage.

Caption: The relationship between UHDRS total motor score (UHDRS-TMS) and (a) total sleep time (TST), (b) total wake time (TWT), (c) REM, (d) light and (e) deep sleep times for PSG and FB sleep lab data, and home FB data.

Tables

Table I. Demographic and clinical information for participants in the PSG study, and the subset who participated in the home study.

	Sleep lab (PSG + FB)		**Home (FB)**	
N	10		9	
Male:Female	5:5		5:4	
	Mean ± SD	**Range**	**Mean ± SD**	**Range**
Age (years)	46.3 ± 13.6	25-64	48.7 ± 12.0	25-64
BMI (Kg/m^2)	25.9 ± 5.4	20.9-37.5	26.2 ± 5.6	20.9-37.5
CAG Repeats	44.3 ± 3.7	41-54	43.2 ± 1.4	41-45
UHDRS-TMS (score out of 124)	24.2 ± 16.5	5-51	24.6 ± 17.5	5-51
Disease Burden Score[36]	382.3 ± 102.6	202.5-532	373.3 ± 104.6	202.5-532
ESS (score out of 24)	8.5 ± 5.6	2-18	8.2 ± 5.8	2-18
PLMS (events/hr)	25.7 ± 44.3	0-144.2	27.8 ± 46.5	0-144.2
AHI (events/hr)	5.7 ± 7.2	0.2-16.7	6.3 ± 7.4	0.2-16.7

SD = standard deviation; BMI = body mass index; CAG = Cytosine-Adenine-Guanine trinucleotide; UHDRS-TMS = Unified Huntington's Disease Rater Scale Total Motor Score (score out of 124 point, higher value indicated increased motor symptoms); Disease Burden Score [36] = Age*(CAG-35.5); ESS = Epsworth Sleepiness Scale (scored out of 24 points); PLMS = Periodic Limb Movement of Sleep; AHI = Apnea Hypopnea Index.