# Heart rate monitoring using wrist photoplethysmography in Parkinson's disease: feasibility and relation with autonomic dysfunction

Kars I. Veldkamp[1,2], Luc J.W. Evers[1], Twan van Laarhoven[3] , Yordan P. Raykov[4], Max A. Little[5], King Chung Ho[6], Sooyoon Shin[6], Bastiaan R. Bloem[1], Marc A. Brouwer[2], Jos Thannhauser[1,2].

[1]Radboud university medical center, Donders Institute for Brain, Cognition and Behaviour, Department of Neurology, Center of Expertise for Parkinson & Movement Disorders; Nijmegen, the Netherlands.

[2]Radboud university medical center, Department of Cardiology, Nijmegen, the Netherlands.

[3]Radboud University Nijmegen, Institute for Computing and Information Science, Nijmegen, the Netherlands.

[4]University of Nottingham, Nottingham, United Kingdom.

[5]University of Birmingham, Birmingham, United Kingdom.

[6]Verily Life Sciences, South San Fransisco, CA, United States.

**Corresponding author:** Kars Veldkamp, Radboud university medical center, Department of Neurology, Center of Expertise for Parkinson & Movement Disorders, PO Box 9101, 6500 HB Nijmegen, Netherlands. E-mail: kars.veldkamp@radboudumc.nl. Tel.: 00 31 24 3615202.

**Word count abstract: 149**

**Running title:** Heart rate monitoring using wrist photoplethysmography in Parkinson disease

**Key words:** Parkinson disease, photoplethysmography, pulse rate monitoring, autonomic dysfunction, remote monitoring, digital markers

## Abstract

Wrist photoplethysmography (PPG) allows for continuous pulse rate monitoring in people with Parkinson disease (PD), with potential to measure autonomic dysfunction. However, motion artifacts remain a challenge, as they can affect signal quality and bias pulse rate estimates. In this study of 444 people with PD, we propose a novel approach to account for motion artifacts, and explore the effect of autonomic dysfunction on pulse rate characteristics. The proportion of high-quality PPG data was unaffected by rest tremor and dyskinesia. At the individual level, filtering periodic motion artifacts significantly reduced overestimation of tremor-related maximum pulse rate. Greater autonomic dysfunction led to a lower maximum pulse rate, but did not affect resting pulse rate. In conclusion, PPG-based heart rate estimation in PD improves when periodic motion artifacts are accommodated. Future studies are warranted to confirm the association between daily-life pulse rate measurements and autonomic dysfunction, as observed in this study.


### Key contributions/strengths:

- Careful signal quality assessment, taking into account periodic artifacts by combining PPG & accelerometer
- Careful HR estimation using smoothed-pseudo Wigner-Ville, tested on an independent dataset.
- Evaluation on a unique large PPG dataset from people with PD
- Correlation to the clinical gold standard for autonomic dysfunction
- Pipeline open source available

# Introduction

Non-motor symptoms are an important part of the spectrum of symptoms in Parkinson disease (PD), and play a crucial role in shaping quality of life among people with PD[1]. One prevalent yet often underdiagnosed non-motor manifestation of PD is dysfunction of the autonomic nervous system. This can manifest in various forms, including orthostatic hypotension, constipation and urinary dysfunction[2–4]. Assessing these symptoms is challenging: current methods rely solely on questionnaires and periodic hospital visits, which are influenced by the subjective interpretations of both patients and clinicians[5–7]. These traditional assessments could lead to misdiagnosis, resulting in a high number of unnecessary follow-ups and treatment. Therefore, there is need for more objective and continuous methods to assess autonomic symptoms in PD[8–10].

A potential alternative way to assess autonomic dysfunction in PD involves analyzing heart rate variation. Autonomic regulation of heart rate is affected in PD, even in prodromal stages, and signs of non-cardiac autonomic dysfunction (e.g. bladder dysfunction, constipation) are more frequent among individuals with cardiac autonomic dysfunction[11–13]. People with PD have lower maximum heart rates, elevated resting heart rates and impaired heart rate recovery after exercise[14,15]. Insights into heart rate abnormalities currently stem mainly from episodic heart rate measurements. It would be relevant to move towards continuous measurements over a longer period, thus creating a continuous "digital marker" based on heart rate patterns. Such a digital heart rate marker could offer insights into daily fluctuations and progression of autonomic dysfunction in PD. It could also help investigate the relationship between heart rate profiles and the severity of autonomic symptoms, which currently remains poorly understood. Ultimately, this could support clinical decision making, guiding treatment adjustments or optimizing medication timing and dosing.

Wrist photoplethysmography (PPG) can be used to unobtrusively monitor pulse rate, which under typical conditions corresponds to heart rate[16], and thereby indicate impaired heart rate regulation in daily life. Despite the growing use of PPG in wrist-worn devices, their application in ambulatory

settings faces challenges due to motion artifacts[17], which may be particularly relevant for people with PD, many of whom experience excessive movements caused by tremor or dyskinesia.

Here, we aim to evaluate the feasibility of using wrist-worn PPG signals to assess pulse rate profiles and their relation with autonomic dysfunction in people with PD during daily life. Given the exploratory nature of this study and the variability in motor symptoms among individuals with PD, we do not predefine feasibility thresholds. Instead, we assess PPG signal quality throughout day and night, focusing on the artifact-inducing effects of motor symptoms on signal quality and pulse rate estimation. Additionally, to provide insight into the clinical relevance of daily-life PPG monitoring in this population, we examine the influence of autonomic dysfunction on pulse rate parameters.

# Results

## Study population

The baseline characteristics of the study population are presented in Table 1. From the 520 participants in the PPP, 34 participants had insufficient data recordings in week 0 (n=10), week 1 (n=17) or both (n=7). Moreover, 22 participants had proven rhythm disorders on the screening ECG. 20 participants selected using stratified sampling were used for training the signal quality algorithm. The remaining 444 participants were included in the main analyses.

Participants had a mean age of 61.4 years (standard deviation, SD: 9.0), of whom 187 (42.1%) were women. Mean time since PD diagnosis was 31.4 (SD: 17.5) months. MDS-UPDRS Part III scores (in the off state) reflected mild to moderate disease severity (33.2, SD: 12.9). The synchronized sensor data (accelerometer + PPG) available for analysis comprised 22.0 hours per day per participant (SD: 1.1) in week 0 and 22.1 hours per day (SD: 1.2) in week 1.

**Table 1: Baseline characteristics of the study population. Data are means (SD) or numbers (%).**

| | **N = 444** |
|---|---|
| **Age (years)** | 61.4 (9.0) |
| **Sex (number of women)** | 187 (42.1%) |
| **Time since diagnosis (months)** | 31.4 (17.5) |
| **SCOPA-AUT score, range 0-69** | 14.5 (7.2) |
| **MDS-UPDRS Part III OFF score, range 0-132** | 33.2 (12.9) |
| **Use of beta blockers** | 40 (9.0%) |
| **PASE score, range 0-400** | 173.2 (81.2) |
| **Resting pulse rate – day ($min^{-1}$)** | 66.5 (9.2) |
| **Resting pulse rate – night ($min^{-1}$)** | 61.4 (8.0) |
| **Maximum pulse rate ($min^{-1}$)** | 87.1 (13.0) |

| | |
|---|---|
| **Daily sensor data availability week 0** | 22.0 (1.1) |
| **Daily sensor data availability week 1** | 22.1 (1.2) |

93

## Signal quality - descriptives

Signal quality was defined using a two-step algorithm combining PPG signal morphology and detection of periodic motion artifacts. This approach is visually summarized in Figure 1, providing a clear overview of the criteria used to identify high-quality segments.

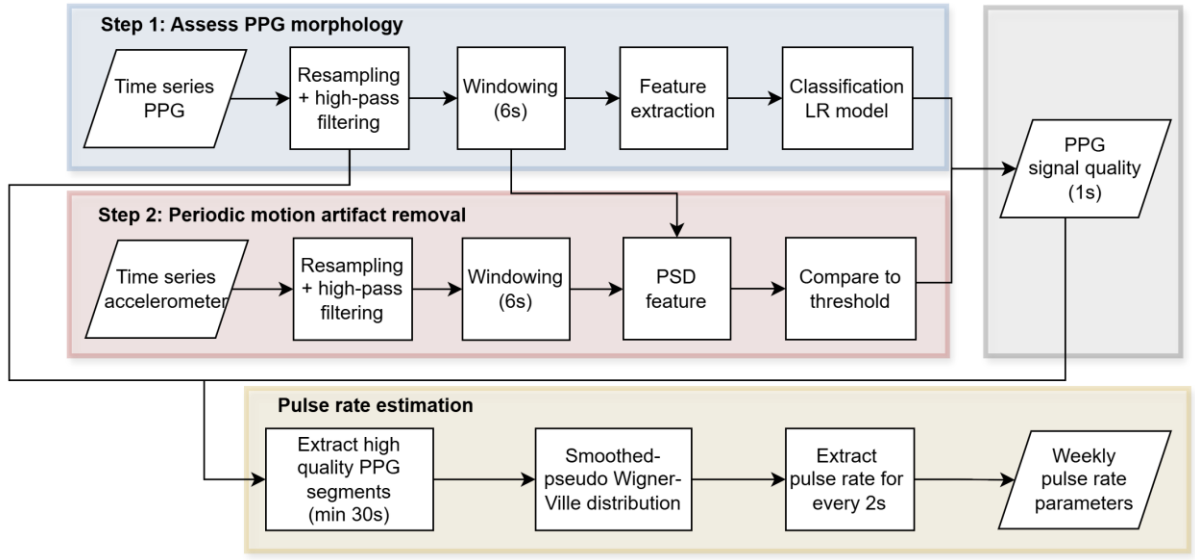



**Figure 1: An overview of the signal processing pipeline for PPG signal quality and pulse rate estimation. The pipeline consists of a PPG signal quality filter and a pulse rate estimation method: 1) signal quality filter, combining PPG morphology classification (step 1) and periodic motion artifact detection (step 2) using accelerometer data to obtain per-second quality labels; and 2) pulse rate estimation from high-quality PPG segments using the smoothed pseudo Wigner-Ville distribution. Pulse rate is estimated for every 2s and summarized weekly. LR = logistic regression, PPG = photoplethysmography; PSD = power spectral density.**

In Figure 2, the median (IQR) proportion of high quality PPG signals is shown across every hour of the circadian cycle. During daytime hours, the median proportion was 29.2% [24.0%, 35.9%]; while during nighttime hours, it was 86.1% [79.3%, 90.6%]. We observed a similar pattern in the second study week (Supplementary Figure 5, Supplementary Table 3). Removing periodic artifacts resulted in an overall reduction of 0.10% of the median proportion of high-quality PPG signals during the day (Supplementary Figure 6, Supplementary Table 3).

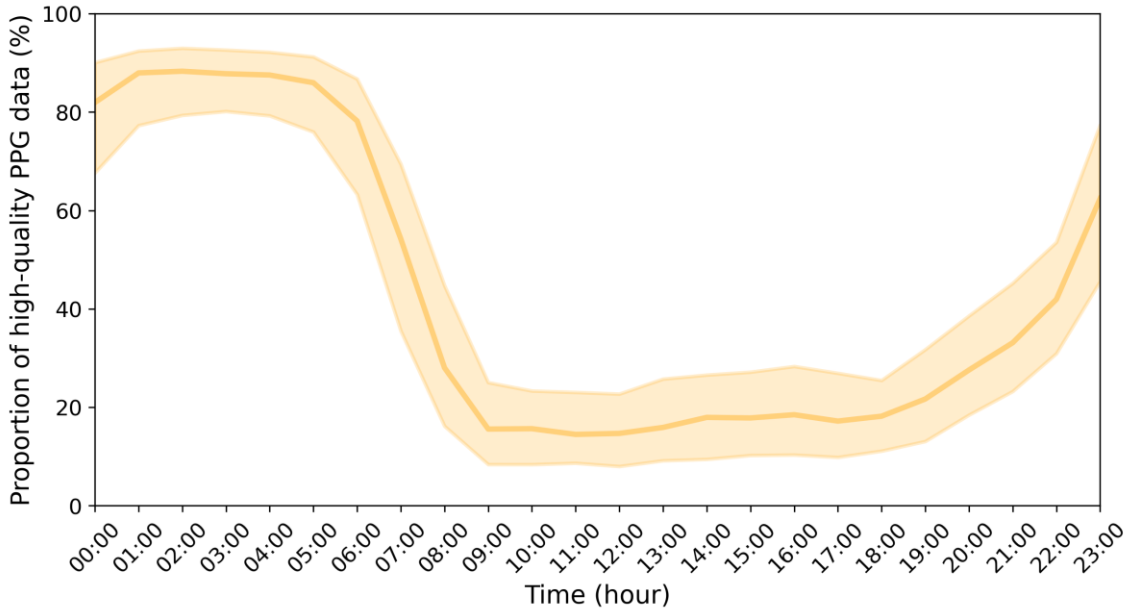



**Figure 2: Proportion of high-quality PPG data across the circadian cycle in the first study week. Data are represented as medians + IQR.**

## Impact of tremor and dyskinesia on signal quality

No significant differences in day-time data quality were found among the no tremor, mild tremor, and severe tremor groups [$F(2,441)=0.19$, $p=0.83$; Figure 3]. Similarly, no significant differences in day-time data quality were found among the no dyskinesia, mild dyskinesia and severe dyskinesia groups [$F(2,441)=1.07$, $p=0.34$; Figure 4]. All comparisons between study groups based on tremor and dyskinesia can be found in Supplementary Tables 3 and 4.

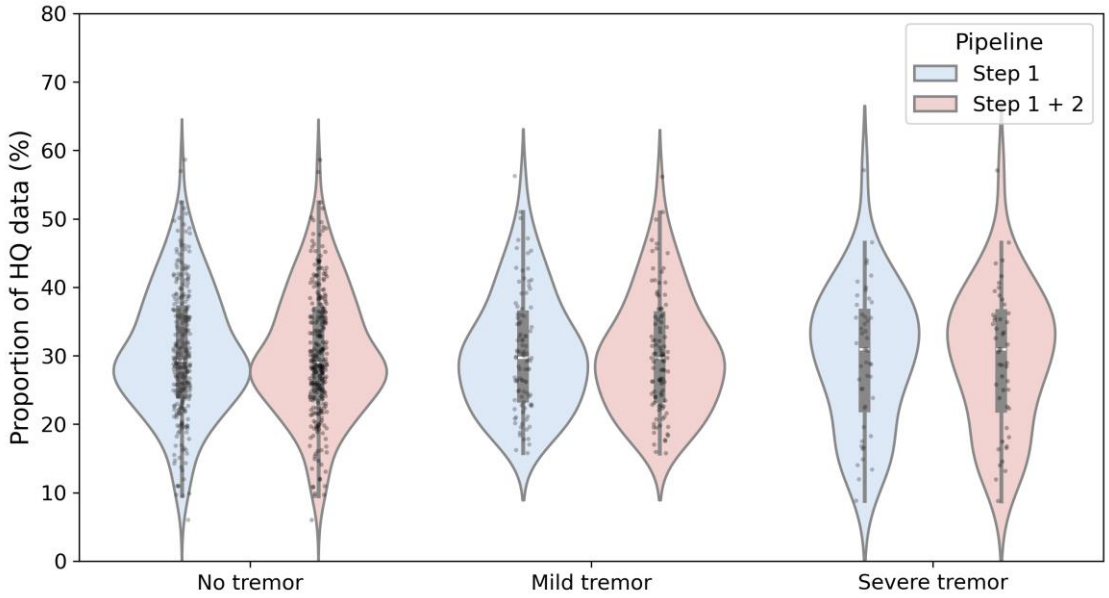



**Figure 3*:* The proportion of high-quality PPG data in relation to tremor severity (MDS-UPDRS 3.17) in the first study week. The different colors represent the data quality when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red).**

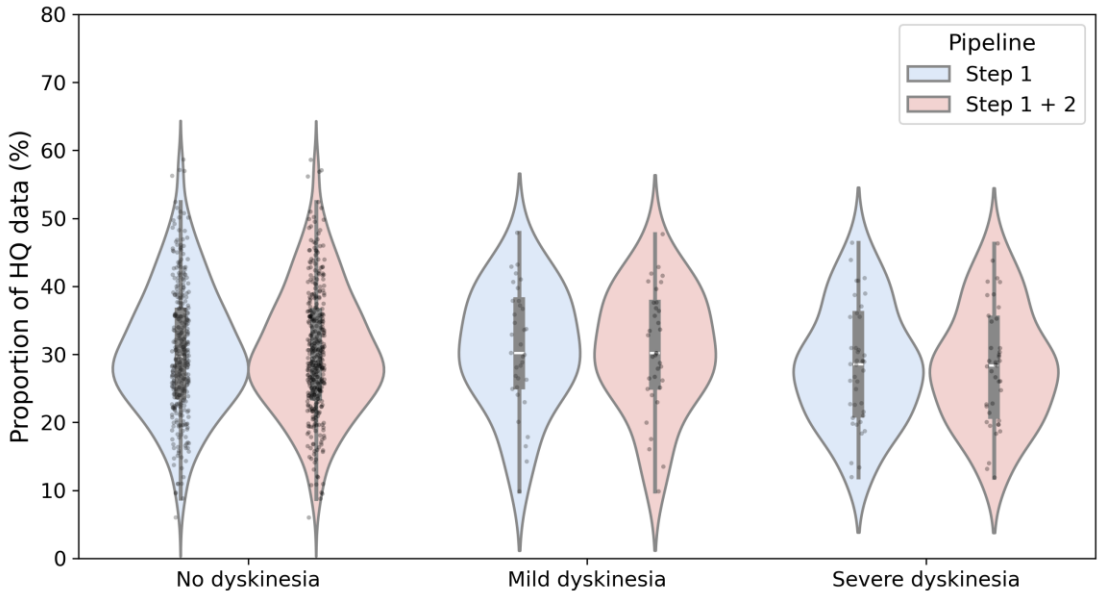



**Figure 4: The proportion of high-quality PPG data in relation to dyskinesia severity (MDS-UPDRS 4.1 + 4.2) in the first study week. The different colors represent the data quality when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red).**

## Impact of periodic artifact removal on pulse rate estimates

Next, we assessed the effect of periodic artifact removal on two-second pulse rate estimates in the physiological range (40-180 beats per minute). Pulse rate was estimated every two seconds using smoothed pseudo Wigner-Ville distribution (SPWVD) time-frequency analysis, applied to 30-second segments of high-quality PPG signals.

Figure 5 illustrates the mean absolute removal of pulse rate estimates per subject across the different tremor groups in the first study week. Individuals with higher tremor scores demonstrated numerically more detection and removal of high-frequency periodic motion artifacts, particularly in the >160 bpm range. No difference is seen between the different dyskinesia groups (Supplementary Figure 7). Furthermore, in all study groups, low-frequency (40-120 bpm) periodic motion artifacts were removed, with the highest proportion in the 40-80 bpm range.

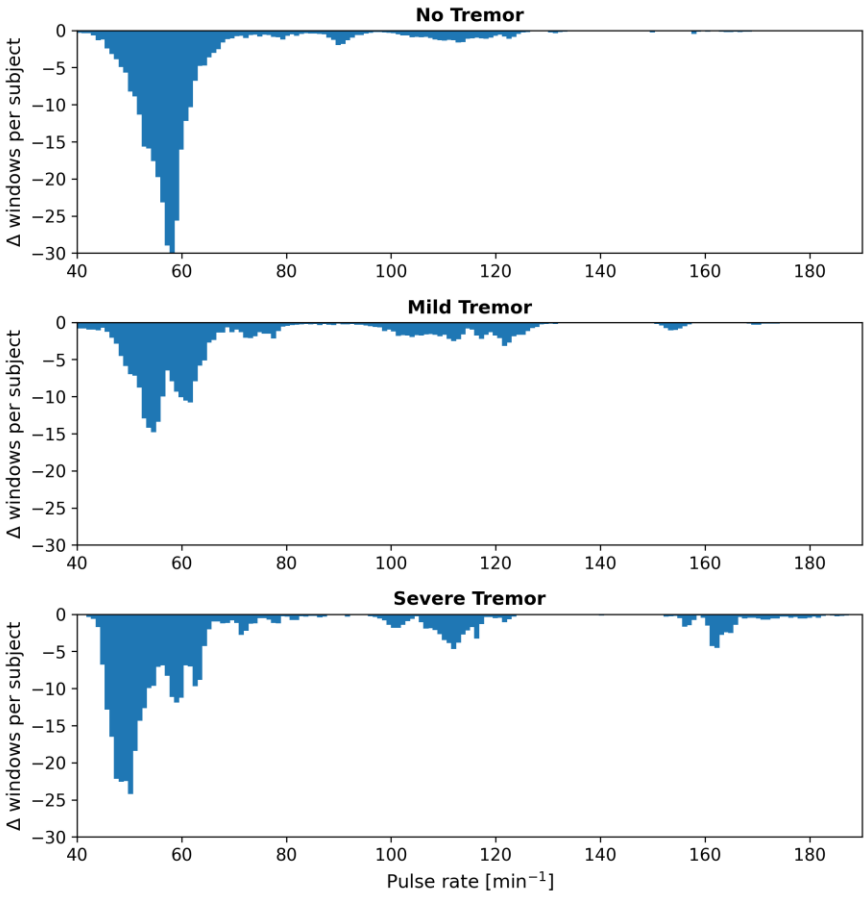


**Figure 5: Removed pulse rate estimates by applying step 2 of the signal quality algorithm, for the different tremor groups. The data are represented as the mean absolute removal per subject per pulse rate value. People with a higher tremor score show more exclusion of higher pulse rate estimates.**

Figure 6 shows a representative example of a data fragment from a severe tremor subject where periodic movement artifacts influenced the PPG signal. By applying the periodic motion artifact filter in step 2 of the signal quality algorithm, this segment is excluded from further pulse rate analysis.

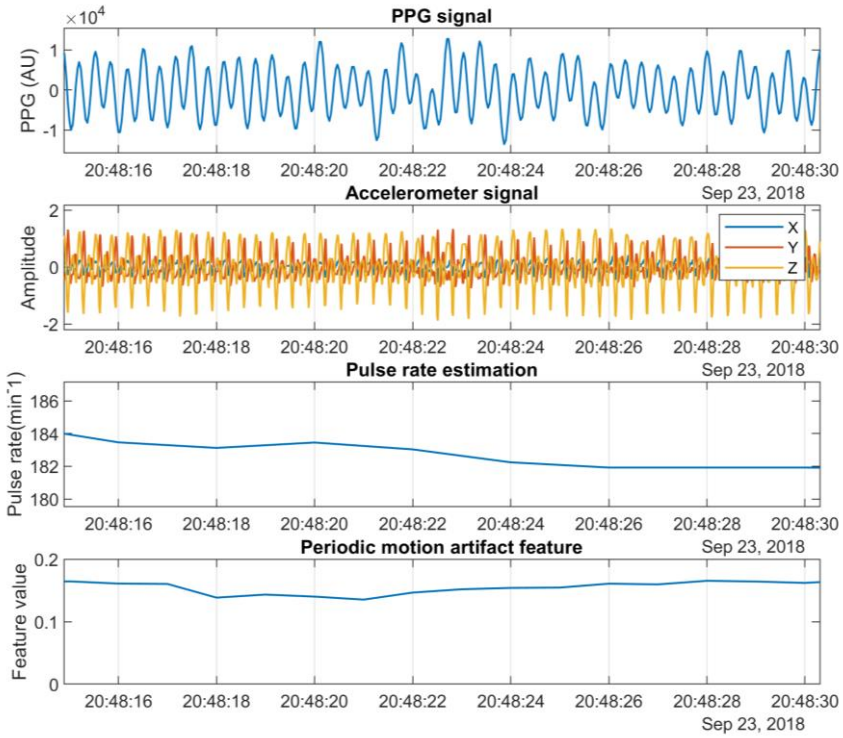



**Figure 6: Data fragment of periodic artifact leaking into PPG in a severe tremor subject. In the upper two panels, the PPG signal shows substantial alignment with acceleration, indicating that a potential motion artifact is present in the PPG signal. Without artifact removal, the extracted pulse rate of the PPG signal is approximately 180 beats per minute (third panel). However, the relative power in the accelerometer at the dominant PPG frequency (fourth panel) exceeds the motion artifact threshold of 0.10, leading to the exclusion of this estimate from further analysis.**

Next, we assessed the impact of periodic artifact removal on the weekly aggregated pulse rate parameters. After filtering, the mean resting pulse rate was 61.4 $min^{-1}$ (SD: 8.0) at night and 66.5 $min^{-1}$ (SD: 9.2) during the day. The mean maximum pulse rate was 87.1 $min^{-1}$ (SD: 13.0). Full distributions of all pulse rate parameters are shown in Supplementary Figure 8. The effect of periodic motion artifact removal (Step 2) on the aggregated maximum pulse rate across the three tremor groups is illustrated in Figure 7. There were significant differences in maximum pulse rate between the two configurations across all tremor groups. In participants without tremor, the mean maximum pulse rate decreased from 87.2 bpm to 85.2 bpm after periodic motion artifact removal, a decrease of 2.0 bpm (SD 8.2; $p < 0.001$). In those with mild tremor, the mean maximum pulse rates dropped from 91.3 bpm to 87.5 bpm after,  a difference of 3.8 bpm (SD 11.3; $p = 0.001$). Participants with severe tremor showed a decrease from 97.6 bpm before and 92.7 bpm, representing a difference of 4.9 bpm (SD 14.7; $p = 0.02$). On an individual level, several maximum pulse rates in all groups were affected by more than 30 BPM. We saw a similar pattern for the three different dyskinesia groups (Supplementary Figure 10).

In contrast to the maximum pulse rate, individual resting pulse rate parameters were not affected by removal of periodic motion artifacts across tremor and dyskinesia groups (Supplementary Figure 11-14).

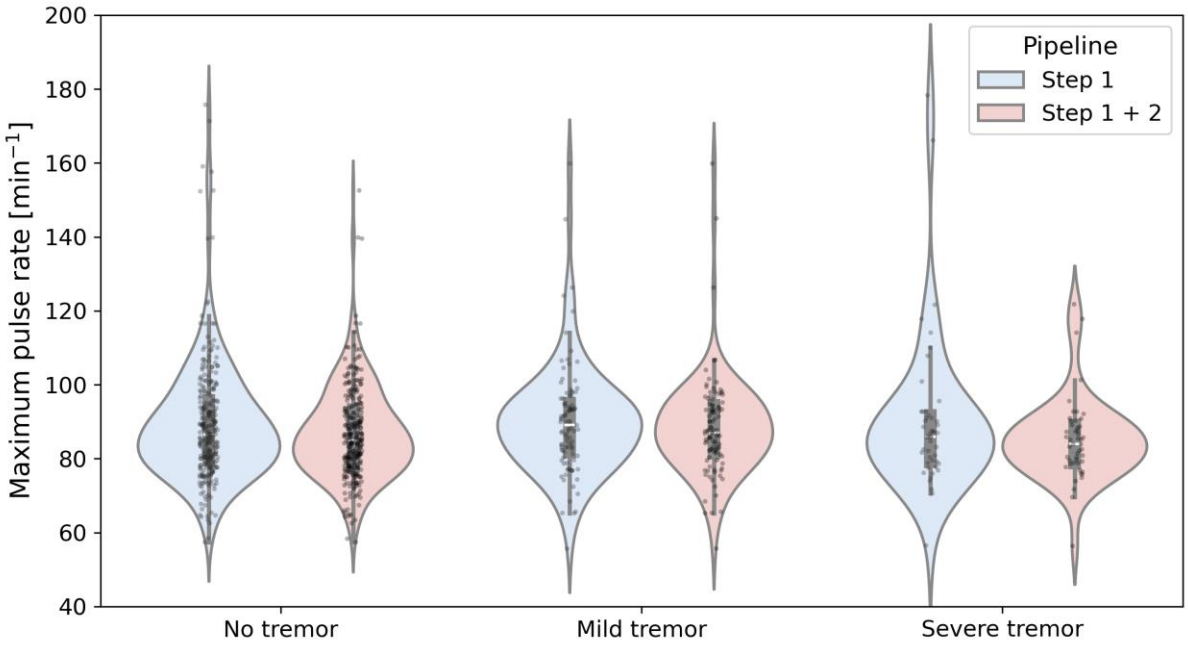



**Figure 7*:* The maximum pulse rate in relation to tremor severity (MDS-UPDRS 3.17) in the first study week. The different colors represent the maximum pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). While filtering for periodic motion artifacts significantly affects the aggregated maximum pulse rate at the group level, its effect was particularly substantial at a individual level – especially in the severe tremor group.**

## Effect of autonomic dysfunction on pulse rate parameters

More severe autonomic dysfunction led to a lower maximum pulse rate during the day in the first study week (β:-0.17, 95% CI: [-0.33, -0.00]; Figure 8), and in the second study week (β:-0.26, 95% CI: [-0.41, -0.04]; Supplementary Figure 15). In contrast, we found no effect of autonomic dysfunction on resting pulse rate, during either the night (β:0.02, 95% CI: [-0.08, 0.13]) or the day (β:-0.02, 95% CI: [-0.13, 0.10]) in the first study week (Figure 8). Similar results were obtained in the second study week (Supplementary Figure 15).

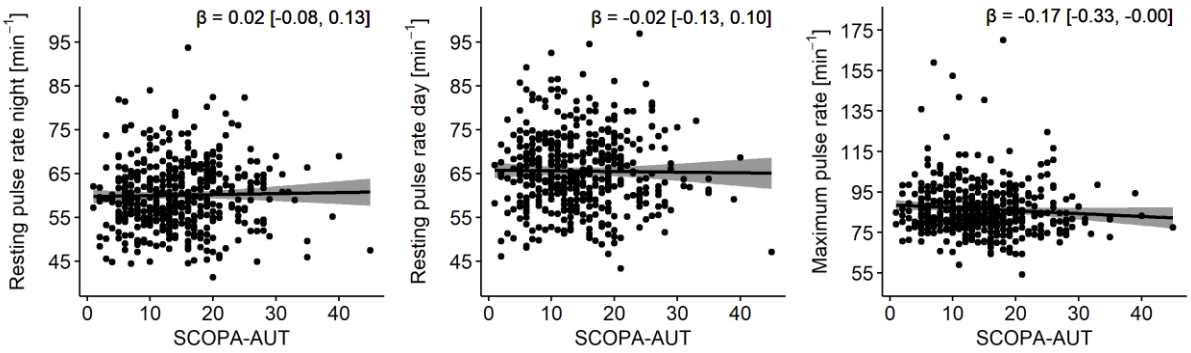



**Figure 8. Multivariate regressions between the pulse rate parameters in the first study week and SCOPA-AUT. Every data point is corrected using the following covariates: age, sex, use of betablockers and baseline physical activity.** SCOPA-AUT = SCales for Outcomes in PArkinson's disease - Autonomic dysfunction.

## Discussion

This study explored the feasibility of wrist-worn PPG for measuring pulse rate parameters in people with PD during daily life and to assess the effect of autonomic dysfunction on these parameters. First, we addressed the problem of PPG motion artifacts in the PD population. Our findings indicate that the severity of tremor and dyskinesia does not affect the overall availability of high-quality PPG data. However, periodic motion artifacts, such as those caused by tremors, can impact pulse rate estimates. Segments with such periodic artifacts can be filtered out by combining PPG and accelerometer signals, reducing overestimation of the maximum pulse rate at the individual level with minimal data removal. Second, we studied the effect of autonomic dysfunction on pulse rate parameters. We found a weak effect of severity of autonomic dysfunction on weekly maximum pulse rate, but no effect on weekly resting pulse rate. This indicates that features related to autonomic dysfunction can be captured using PPG in people with PD, but further research is needed to identify more specific pulse rate (variability) markers to monitor autonomic dysfunction.

Compared with existing literature, our study provides novel insights into the impact of periodic motion artifacts on the quality of PPG signals in people with PD. The impact of these artifacts on

daily life assessments had not been studied before. We demonstrate that filtering out such artifacts (e.g., tremor) minimally reduces the proportion of high-quality PPG data, but it can improve pulse rate estimation in segments affected by periodic motion artifacts. The extraction of pulse rates from PPG signals can be approached using traditional signal analysis methods[18–20] or more advanced techniques, such as deep learning combined with sensor fusion (accelerometer and PPG)[21]. Traditional methods, as employed in this study, typically focus on identifying and extracting high-quality PPG signals prior to pulse rate estimation[22]. However, most studies focus solely on PPG signal characteristics, whereas our study also integrates accelerometer data to explicitly account for periodic motion artifacts. This approach is particularly important in the context of PD, as our results suggest that tremor-related motion can mimic physiological patterns in the PPG signal. This insight may extend to broader contexts, such as lower-frequency periodic movements during e.g. gait (around 1 Hz)[23]. This is supported by our data, showing that most removed periodic artifacts, regardless of tremor group, were concentrated in the 40-80 bpm range (0.7-1.3 Hz). Although sophisticated approaches like statistical models and machine learning, including deep learning incorporating frequency characteristics of both PPG and accelerometer, have been explored primarily in sports research, they have not yet been specifically developed for diseased populations[21,24,25]. We posit that such approaches are suited to also address PD-specific challenges. Future research should explore these methods, as they may enhance pulse rate estimation reliability during lower-quality PPG segments encountered in daily activities.

Previous lab-based studies demonstrated that people with PD have a lower maximum heart rate during exercise compared to controls due to autonomic dysfunction. We extended this by examining the effect of *severity* of autonomic dysfunction on maximum pulse rate in daily life. We found that more severe autonomic dysfunction leads to a lower maximum pulse rate. This supports the notion that, in addition to parasympathetic dysfunction, impaired sympathetic dysfunction is a pathophysiological feature of autonomic involvement in PD[3,26]. Based on pathophysiological reasoning, we expected to observe an effect of autonomic dysfunction on resting pulse rate[27].

However, our findings did not support this hypothesis, aligning instead with mixed empirical evidence, with some studies reporting no difference in daytime resting heart rate between PD and controls[28]. Similarly, findings on how PD affects nighttime resting heart rate are also inconsistent[29,30], even though this is considered to be a more stable metric for autonomic function (because of lower measurement variability due to fewer external influences[31,32]).

We are the first to study these relationships using real-life PPG data, rather than during controlled settings. We observed only a weak effect of autonomic dysfunction, as assessed through a patient-reported outcome (SCOPA-AUT) , on maximum pulse rate. Multiple factors could explain this weak effect. First, maximum pulse rate is highly influenced by physical activity, which not only elevates the heart rate but also introduces motion artifacts that can lower PPG signal quality[33,34]. Although we adjusted for physical activity using the Physical Activity Scale for the Elderly (PASE), maximum pulse rate remains highly dependent on the amount and level of physical activity and the ability to capture high-quality PPG data during these activities. This is reflected in our findings, where the observed maximum pulse rates were relatively low compared to participants' age-predicted maximum heart rates. Second, it is possible that the patient-reported outcome to assess autonomic dysfunction does not reliably capture all the physiological aspects of autonomic function. A more robust and objective alternative would be to use quantitative measures such as cardiac innervation imaging (123I-mIBG scintigraphy [35]). Given these limitations, future research should consider not only global pulse rate parameters but also zoom in on heart rate responses in more specific behavioral contexts, such as sleep or daily activities (e.g. during walking or other intense physical activities). Nocturnal analyses offer a promising avenue for research because reduced levels of movement during sleep will likely improve the PPG data quality. This would also allow for pulse rate variability measurements, which could yield more accurate assessments of autonomic function[36–38]. The merits of this approach remain to be formally demonstrated, as persons with PD can manifest a wide range of nocturnal movements, including dream enactment behavior and periodic leg movements during sleep[39].

A key strength of this study was the use of a large dataset collected from a representative population of people with early-stage PD through continuous real-world monitoring. Participants wore the smartwatch on average 22 hours per day, indicating strong adherence and consistent device use in daily life. This extensive dataset provides a comprehensive insight into pulse rate patterns in daily life of people with PD, including both daytime and nighttime recordings. Moreover, this study establishes a first link between free-living pulse rate data and autonomic dysfunction, offering new insights into autonomic regulation in PD. Another notable strength was the approach to data annotation and validation of pulse rate estimations. The careful annotations ensured the accuracy and reliability of the PPG signal quality, while the validation of the pulse rate estimation (using the smoothed pseudo Wigner-Ville distribution with alternative methods using an external dataset) reinforces the robustness of the methodology. Importantly, signal quality assessment explicitly accounted for periodic motion artifacts, such as tremors, by combining PPG and accelerometer signals. This multimodal approach allowed us to filter out segments with periodic disturbances, thereby improving the reliability of the pulse rate parameters. This supports the potential of PPG as a tool for monitoring pulse rate and autonomic function in people with PD. The Personalized Parkinson Project offers a unique opportunity to further study this concept, as we have digital data available for all participants for minimally up to 2 years (and up to 3 years for a sizeable subgroup)[8,40].

This study was not without limitations. One key drawback was the reliance on a self-administered questionnaire (SCOPA-AUT) to assess autonomic dysfunction. While this questionnaire is considered a gold-standard assessment for one's autonomic dysfunction, it remains inherently subjective and prone to high within-subject variability[41]. This could explain why we found no effect of autonomic dysfunction on resting pulse rate and only a weak effect on maximum pulse rate. Future studies should consider incorporating objective, quantitative measures of autonomic function, such as orthostatic blood pressure testing, sweat tests, and cardiac innervation imaging using 123I-mIBG scintigraphy. In addition, the estimation of the effect of autonomic dysfunction on pulse rate

parameters from observational data should be interpreted with caution, given its underlying assumptions. Potential residual confounding and imperfect measurement of key variables may have influenced the results. To increase transparency about these assumptions, we included a DAG, which may be further refined in future research. Another limitation is the lack of validation for the impact of filtering using ECG data. Also, we did not explicitly assess whether segments with a high correlation between the spectral content of the accelerometer and PPG indeed resulted in erroneous pulse rate estimates when compared to ECG. Such an analysis would have helped to confirm that the high pulse rate estimates removed by filtering indeed reflected periodic motion artifacts, rather than physiological changes. However, our feature for periodic motion artifact detection was designed to identify segments with high accelerometer-PPG correlation. Therefore, we expect that filtering mainly removes implausible pulse rate estimates, while discarding only a minimal amount of real pulse rate data. Future research should confirm this by directly comparing filtered PPG data with ECG recordings. Finally, the potential impact of tremor on pulse rate estimation was only assessed at the group level using clinically based tremor assessments. Future work could benefit from integrating tremor and gait detection algorithms, particularly given the presence of low-frequency artifacts suggestive of possible gait-related motion. Such approaches would allow for per-segment analysis of periodic motion artifacts, providing a more precise evaluation of their effects on pulse rate measurements.

Identifying a reliable digital marker for autonomic dysfunction in PD remains a complex challenge, but this study represents an important first step. The current pipeline provides a novel method for estimating pulse rate in the PD population, supporting the development of PPG-based digital markers of autonomic function. However, further refinements – such as improving performance during motion – could increase the accuracy and robustness of these pulse rate markers in real-world, longitudinal applications. In parallel, future research should prioritize identifying more specific digital markers for autonomic dysfunction, such as pulse rate variability. This will be crucial for developing a more comprehensive set of markers for tracking autonomic dysfunction.

Longitudinal evaluations in clinical trials could provide valuable insights into the progression of autonomic dysfunction, with trends over time likely offering more actionable information than cross-sectional comparisons between subjects, particularly for pulse rate variability parameters during nighttime. As we continue to explore and validate these approaches, the development of more specific parameters will be crucial for tracking disease progression.

# Methods

## Study design

The primary objectives are to 1) study signal quality of wearable PPG signals in daily life, especially in the presence of tremor and dyskinesia, and 2) assess the impact of periodic artifact removal on pulse rate estimates. As a secondary objective, we investigate the effect of autonomic dysfunction, as assessed by the SCOPA-AUT questionnaire, on pulse rate parameters (resting pulse rate and maximum pulse rate).

As this is an exploratory study, no predefined thresholds for feasibility were set. However, feasibility is evaluated based on two aspects: (1) technical reliability, defined by the proportion of high-quality data despite the presence of motor symptoms, and (2) clinical meaningfulness, reflected by plausible effects of autonomic dysfunction on pulse rate parameters. Together these analyses aim to assess whether continuous wrist-worn PPG monitoring in PD is both technically viable and potentially informative in a clinical context.

## Study cohort

Data were obtained from the Personalized Parkinson Project (PPP), a single-center cohort study (NCT03364894) including 520 people with early PD (i.e. time since diagnosis ≤5 years)[8]. In brief, participants were monitored in daily life using a wrist-worn sensor device for a minimum of two years (and up to three years in a large subgroup). Participants were also assessed during yearly in-clinic study visits, which included a detailed clinical assessment, MRI, and collection of biosamples.

In the current study, we included participants who completed the baseline clinical assessment and the first two weeks of wearable sensor data collection. The exclusion criteria were insufficient sensor data collection (weekly average wearable wear time < 12 hours per day) and presence of atrial rhythm disorders (as confirmed by screening for cardiac anomalies using electrocardiogram (ECG) recordings (12-lead Holter) or based on medical records). The latter is particularly relevant to ensure

that heart rate patterns reflect autonomic function rather than underlying heart rhythm abnormalities.

## Demographics and clinical data

Data on patient demographics, medical history, medication use and symptom severity were obtained from the baseline study visit. To quantify the severity of autonomic dysfunction, we used the total score on the SCales for Outcomes in PArkinson's disease - Autonomic Dysfunction (SCOPA-AUT) questionnaire[42]. The severity of tremor and dyskinesia was obtained from the Movement Disorders Society-Unified Parkinson Disease Rating Scale (MDS-UPDRS) part III and IV assessments, which was conducted by trained assessors[43].

## Wearable sensor data

Continuous monitoring of the participants was facilitated using a wearable sensor device (Verily Study Watch, Verily Life Sciences, CA, USA). This watch enables data collection by several sensors, including PPG and accelerometer. During the baseline study visit, participants were instructed on the correct watch placement on the arm, charging, and maintenance. PPG data were obtained at a sampling rate of 30 Hz and accelerometer data at a sampling rate of 100 Hz. In this study, we used the PPG and accelerometer data obtained during the first two weeks of follow-up.

## Wearable sensor data preprocessing

The raw sensor data were stored in the *TSDF* data format to allow for efficient processing[44]. Minor variations in the sampling frequency were resolved by resampling the PPG data through cubic spline interpolation to exactly 30 Hz (PPG) and 100 Hz (accelerometer). Two fourth-order high-pass Butterworth filters with the cut-off frequencies of 0.4 Hz (PPG) and 0.2 Hz (accelerometer) were applied to eliminate offset and trend components in the signal.

## Signal quality assessment

We employ an approach to filter out PPG segments with significant motion artifacts which could impede reliable pulse rate estimation using the combined outputs of the following steps:

1. Assessment of PPG morphology
2. Periodic motion artifact removal

In the first step, segments are filtered out based on the typical PPG morphology. In the second step, remaining segments with periodic motion artifacts (e.g. introduced by tremor) which mimic pulse waves were filtered out by combining the PPG and accelerometer data. The schematic representation of this approach is depicted in Figure 1.

### Step 1: Assessment of PPG morphology

We assessed PPG morphology based on the typical sinusoidal pulse waves with discernible peaks from which pulse rate can be determined[18,19,22]. To detect typical PPG morphology, we employed a machine learning modelling approach – a logistic regression (LR) classifier. The LR classifier was trained using an annotated dataset consisting of PPG data from 20 PPP subjects (training set, not used in the other evaluations), selected through stratified sampling based on age, sex, and rest tremor scores. The PPG signals were annotated by visual inspection by two experienced investigators (KV, JT) following a standardized annotation protocol, which included a detailed definition for the PPG morphology of low- versus high-quality segments (Supplementary Methods 1.1.1.2). First, the two investigators independently annotated the same segments and discussed any disagreements to refine the annotation protocol. After sufficient agreement was reached (kappa of 0.8), the remaining segments were annotated by a single investigator. A detailed description of the annotation process can be found in Supplementary Methods 1.1.1.

The LR model to classify PPG morphology was trained using window-based time and frequency features. L1 regularization was applied for feature selection to enhance classifier efficiency. The performance was evaluated using leave-one-subject-out cross-validation. Compared to the visual annotations, the classifier's accuracy was 98.4% (SD: 0.95%), with a sensitivity of 98.2% (SD: 1.32%) and a specificity of 98.5% (SD: 1.32%).

A detailed description of the training process, including feature selection and the specific features used in the LR classifier, can be found in Supplementary Methods 1.1.2-1.1.4.

### Step 2: Periodic motion artifact removal

Due to the sensitivity of PPG to motion artifacts, we hypothesized that periodic movements such as tremor can cause periodic artifacts very similar to the typical pulse wave in the PPG signal. During the annotation process, we indeed found several examples of oscillations in the PPG signal that were accompanied by oscillations of the same frequency in the accelerometer signal (see Figure 9 for one example). These examples occurred in the low rest tremor range (around 2.5 to 3 Hz), which is still within the physiological pulse rate range.

To filter out PPG segments with such periodic artifacts, we combined the PPG and accelerometer signals. We considered a segment to be a periodic artifact if the dominant frequency of the PPG matched the dominant frequency of the accelerometer. To identify these artifacts, we calculated the window-based relative power in the accelerometer at the dominant PPG frequency (±0.05 Hz) and its first harmonic (±0.05 Hz). If the relative power exceeded a threshold of 0.10 the window was classified as an artifact. The threshold was selected based on the training dataset. Supplementary Methods 1.2 provide more details on the calculation of the relative power and the threshold selection.

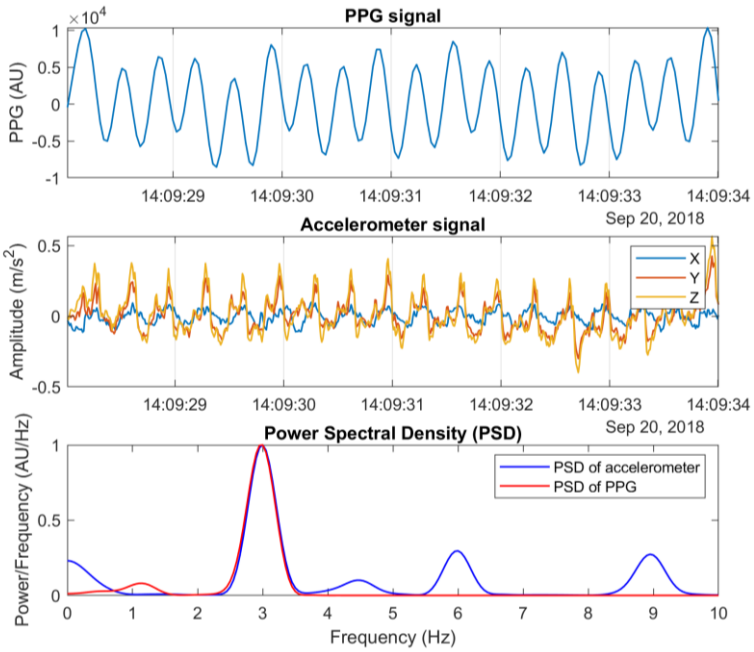



**Figure 9: Possible leakage of periodic movements into the PPG signal. Upper panel: PPG signal, middle panel: three axes of the accelerometer, lower panel: power spectral density of PPG and the sum of the three accelerometer axes. The dominant frequency in both signals is +/- 3 Hz, which falls within both the physiological pulse rate range and**

## Composite signal quality classification

Both steps were applied to 6s windows with a 5s overlap, after which the final signal quality classification was made for every 1s. A window was classified as high-quality, only if both steps classified it as such. To produce the final quality prediction for every 1s, majority voting was applied to the overlapping windows. More details can be found in Supplementary Methods 1.3.

## Impact of tremor and dyskinesia on signal quality

To investigate the impact of tremor and dyskinesia on PPG signal quality, participants were stratified into sub-groups based on their rest tremor scores on the device-sided arm (item 3.17) and combined dyskinesia scores (items 4.1 and 4.2) from the MDS-UPDRS at the baseline visit. Rest tremor scores were categorized into three groups: no tremor (score of 0), mild tremor (score of 1), and severe tremor (score of 2 or higher). Similarly, dyskinesia scores, derived from both the duration and

functional impact of dyskinesias, were divided into three groups: no dyskinesia (combined score of 0), mild dyskinesia (combined score of 1), and severe dyskinesia (combined score of 2 or higher). Signal quality estimates were analyzed for either daytime hours (08:00 – 21:59) or nighttime hours (00:00 – 05:59).

## Pulse rate estimation

Following evaluation of the PPG signal quality, we proceeded with pulse rate estimation, as depicted in Figure 1. For this, 30 consecutive high-quality 1s windows were used, as determined by the signal quality algorithm. To identify the most suitable method for pulse rate estimation from PPG, we compared various approaches using the PPG DaLiA dataset[45], with ECG as reference. The highest accuracy was obtained using smoothed-pseudo Wigner-Ville Distribution (SPWVD) time-frequency analysis, hence we used the SPWVD for all subsequent pulse rate estimates. See Supplementary Methods 1.4.1-1.4.6 for the results of the comparison.

Using the SPWVD method, we estimated the pulse rate for every 2s. An example of this approach is provided in Supplementary Methods 1.4.7. Pulse rate estimates were divided into either daytime (08:00 – 21:59) or nighttime (00:00 – 05:59) and aggregated per week into three parameters: resting pulse rate (day and night) and maximum pulse rate during the day. The resting pulse rate was defined as the most frequently occurring pulse rate value across high-quality windows within the respective time windows. To ensure a meaningful mode calculation, pulse rate estimates were assigned to frequency bins of 1 beat per minute, based on the approximate frequency resolution from the 30-second SPWVD window. Maximum pulse rate during the day was defined as the 99% percentile of all pulse rate estimates.

## Impact of periodic artifact removal on pulse rate estimates

To evaluate the impact of periodic motion artifacts on pulse rate estimations, we compared the results of two configurations of the processing pipeline: (1) using only PPG morphology for high-quality segment selection (step 1), and (2) using the combined PPG and accelerometer approach to

also account for periodic artifacts (steps 1 and 2). The two seconds pulse rate estimates and weekly pulse rate aggregates derived from both configurations were analyzed across the previously defined study groups for rest tremor and dyskinesia severity.

## Effect of autonomic dysfunction on pulse rate parameters

We used a linear regression model to investigate the effect of autonomic symptom severity, as assessed by the continuous sum score of the SCOPA-AUT questionnaire, on the weekly pulse rate parameters. The pulse rate parameters served as dependent variables in the regression analyses. To reduce bias and strengthen the interpretation of this effect estimate in an observational setting[46], we constructed a direct acyclic graphs (DAG) to identify potential confounding factors known to influence both pulse rate parameters and autonomic dysfunction (see Supplementary Methods 1.5). Based on this DAG, we adjusted the regression models for age, sex, baseline physical activity (assessed by the Physical Activity Scale for the Elderly; PASE), and medication (specifically betablockers). Statistical significance for comparisons between study groups (e.g. tremor and dyskinesia severity) was assessed using one-way ANOVA or Kruskal-Wallis tests, whichever was appropriate. For within-subject comparisons (e.g., pulse rate parameters before and after periodic artifact removal), paired t-tests or Wilcoxon signed-rank tests were applied, whichever was appropriate. A significance threshold of $p < 0.05$ was used. The linear regression models were implemented using the LM function in R (RStudio, version 4.3.2).

## Data availability

Data from the Personalized Parkinson Project used in the present study were retrieved from the PEP database (https://pep.cs.ru.nl/index.html). The PPP data is available upon request via: ppp-data@radboudumc.nl. More details on the procedure can be found on the website www.personalizedparkinsonproject.com/home.

## Code availability

Code for the running the individual PPG signal processing pipeline (preprocessing, signal quality assessment and pulse rate estimation) is available in the ParaDigMa toolbox: https://doi.org/10.5281/zenodo.15223364. The code to generate the aggregated results in this study is publicly available in the Git repository: *https://github.com/biomarkersParkinson/PPP_PPG_feasibility.git*.

## Acknowledgements

This work was financially supported by the Michael J Fox Foundation (grant #020425), the Dutch Research Council (grant #ASDI.2020.060 & grant #2023.010), the Dutch Research Council Long-Term Program (project #KICH3.LTP.20.006, financed by the Dutch Research Council, Verily, and the Dutch Ministry of Economic Affairs and Climate Policy), and by the Dutch Ministry of Economic Affairs (Topconsortium voor Kennis en Innovatie, Life Sciences & Health, grant #LSHM20090-H048). The Radboudumc Center of Expertise for Parkinson's and Movement Disorders was supported by a center of excellence grant from the Parkinson's Foundation.

## Author contributions

Research project (1): A. Conception, B. Organization, C. Execution;

Statistical Analysis (2): A. Design, B. Execution, C. Review and Critique;

Manuscript Preparation (3): A. Writing of the first draft, B. Review and Critique;

K.I.V.: 1A, 1B, 1C, 2A, 2B, 3A.

L.J.W.E.: 1A, 1B, 2A, 2C, 3B.

T.L.: 1A, 2A, 2C, 3B.

J.P.Y.: 1A, 2C, 3B.

M.A.L.: 1A, 2C, 3B.

K.C.H.: 2C, 3B.

S.S.: 2C, 3B.

B.R.B.: 1A, 1B, 2C, 3B.

M.A.B.: 1A, 2C, 3B.

J.T.: 1A, 1C, 2A, 2C, 3B.

## Competing interests

Authors KCH and SS are currently employed by, and currently hold shares in, Verily Life Sciences, but declare no nonfinancial competing interests. Author BRB serves as the co-Editor in Chief for the Journal of Parkinson's disease, serves on the editorial board of Practical Neurology and Digital Biomarkers, has received fees from serving on the scientific advisory board for the Critical Path Institute, Gyenno Science, MedRhythms, UCB, Kyowa Kirin and Zambon (paid to the Institute), has received fees for speaking at conferences from AbbVie, Bial, Biogen, GE Healthcare, Oruen, Roche, UCB and Zambon (paid to the Institute), and has received research support from Biogen, Cure Parkinson's, Davis Phinney Foundation, Edmond J. Safra Foundation, Fred Foundation, Gatsby Foundation, Hersenstichting Nederland, Horizon 2020, IRLAB Therapeutics, Maag Lever Darm Stichting, Michael J Fox Foundation, Ministry of Agriculture, Ministry of Economic Affairs & Climate Policy, Ministry of Health, Welfare and Sport, Netherlands Organization for Scientific Research (ZonMw), Not Impossible, Parkinson Vereniging, Parkinson's Foundation, Parkinson's UK, Stichting Alkemade-Keuls, Stichting Parkinson NL, Stichting Woelse Waard, Health Holland / Topsector Life Sciences and Health, UCB, Verily Life Sciences, Roche and Zambon. Author BRB does not hold any stocks or stock options with any companies that are connected to Parkinson's disease or to any of his clinical or research activities. All other authors declare no competing interests.

# References


1. Prakash, K. M., Nadkarni, N. V., Lye, W. K., Yong, M. H. & Tan, E. K. The impact of non-motor symptoms on the quality of life of Parkinson's disease patients: a longitudinal study. *Eur J Neurol* **23**, 854–860 (2016).

2. Micieli, G., Tosi, P., Marcheselli, S. & Cavallini, A. Autonomic dysfunction in Parkinson's disease. *Neurological Sciences* **24**, s32–s34 (2003).

3. Sharabi, Y., Vatine, G. D. & Ashkenazi, A. Parkinson's disease outside the brain: targeting the autonomic nervous system. *Lancet Neurol* **20**, 868–876 (2021).

4. Marras, C. & Chaudhuri, K. R. Nonmotor features of Parkinson's disease subtypes. *Movement Disorders* **31**, (2016).

5. Evers, L. J. W., Krijthe, J. H., Meinders, M. J., Bloem, B. R. & Heskes, T. M. Measuring Parkinson's disease over time: The real-world within-subject reliability of the MDS-UPDRS. *Mov Disord* **34**, 1480–1487 (2019).

6. van Wamelen, D. J. *et al.* Digital health technology for non-motor symptoms in people with Parkinson's disease: Futile or future? *Parkinsonism Relat Disord* **89**, 186–194 (2021).

7. Lima, A. L. S. de *et al.* Large-Scale Wearable Sensor Deployment in Parkinson's Patients: The Parkinson@Home Study Protocol. *JMIR Res Protoc 2016;5(3):e172 https://www.researchprotocols.org/2016/3/e172* **5**, e5990 (2016).

8. Bloem, B. R. *et al.* The Personalized Parkinson Project: Examining disease progression through broad biomarkers in early Parkinson's disease. *BMC Neurol* **19**, 160 (2019).

9. Janssen Daalen, J. M. *et al.* Digital biomarkers for non-motor symptoms in Parkinson's disease: the state of the art. *npj Digital Medicine 2024 7:1* **7**, 1–22 (2024).

10. Schalkamp, A. K., Harrison, N. A., Peall, K. J. & Sandor, C. Digital outcome measures from smartwatch data relate to non-motor features of Parkinson's disease. *NPJ Parkinsons Dis* **10**, 1–7 (2024).

11. Yang, J. H. *et al.* Association of heart rate variability with REM sleep without atonia in idiopathic REM sleep behavior disorder. *J Clin Sleep Med* **17**, 461–469 (2021).

12. Alonso, A., Huang, X., Mosley, T. H., Heiss, G. & Chen, H. Heart rate variability and the risk of Parkinson disease: The Atherosclerosis Risk in Communities study. *Ann Neurol* **77**, 877–883 (2015).

13. Duijvenboden, S. van *et al.* Cardiac autonomic function during exercise and incident Parkinson's disease. *medRxiv* 2024.11.08.24316979 (2024) doi:10.1101/2024.11.08.24316979.

14. Roberson, K. B. *et al.* Hemodynamic responses to an exercise stress test in Parkinson's disease patients without orthostatic hypotension. *Appl Physiol Nutr Metab* **44**, 751–758 (2019).

15. Maetzler, W. *et al.* Time- and frequency-domain parameters of heart rate variability and sympathetic skin response in Parkinson's disease. *J Neural Transm* **122**, 419–425 (2015).

16. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas* **28**, R1 (2007).

17. Ismail, S., Akram, U. & Siddiqi, I. Heart rate tracking in photoplethysmography signals affected by motion artifacts: a review. *EURASIP J Adv Signal Process* **2021**, 1–27 (2021).

18. Charlton, P. H. *et al.* Wearable Photoplethysmography for Cardiovascular Monitoring. *Proc IEEE Inst Electr Electron Eng* **110**, 355–381 (2022).

19. Elgendi, M. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* **3**, (2016).

20. Vest, A. N. *et al.* An open source benchmarked toolbox for cardiovascular waveform and interval analysis. *Physiol Meas* **39**, 105004 (2018).

21. Reiss, A., Indlekofer, I., Schmidt, P. & Van Laerhoven, K. Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks. *Sensors 2019, Vol. 19, Page 3079* **19**, 3079 (2019).

22. Pradhan, N., Rajan, S. & Adler, A. Evaluation of the signal quality of wrist-based photoplethysmography. *Physiol Meas* **40**, 065008 (2019).

23. Ghislieri, M., Agostini, V., Rizzi, L., Knaflitz, M. & Lanotte, M. Atypical Gait Cycles in Parkinson's Disease. *Sensors 2021, Vol. 21, Page 5079* **21**, 5079 (2021).

24. Guo, Z., Ding, C., Hu, X. & Rudin, C. A supervised machine learning semantic segmentation approach for detecting artifacts in plethysmography signals from wearables. *Physiol Meas* **42**, 125003 (2021).

25. Motin, M. A., Karmakar, C. K. & Palaniswami, M. PPG Derived Heart Rate Estimation during Intensive Physical Exercise. *IEEE Access* **7**, 56062–56069 (2019).

26. Andersen, K. B. *et al.* Sympathetic and parasympathetic subtypes of body-first Lewy body disease observed in postmortem tissue from prediagnostic individuals. *Nat Neurosci* **28**, 925–936 (2025).

27. Gourine, A. V. & Ackland, G. L. Cardiac Vagus and Exercise. *Physiology* **34**, 71 (2019).

28. Haapaniemi, T. H. *et al.* Ambulatory ECG and analysis of heart rate variability in Parkinson's disease. *J Neurol Neurosurg Psychiatry* **70**, 305–310 (2001).

29. Brisinda, D. *et al.* Cardiovascular autonomic nervous system evaluation in Parkinson disease and multiple system atrophy. *J Neurol Sci* **336**, 197–202 (2014).

30. Sauvageot, N., Vaillant, M. & Diederich, N. J. Reduced sympathetically driven heart rate variability during sleep in Parkinson's disease: A case-control polysomnography-based study. *Movement Disorders* **26**, 234–240 (2011).

31. Sammito, S., Thielmann, B. & Böckelmann, I. Update: factors influencing heart rate variability–a narrative review. *Front Physiol* **15**, 1430458 (2024).

32. Damoun, N., Amekran, Y., Taiek, N., Jalil, A. & Hangouche, E. Heart rate variability measurement and influencing factors: Towards the standardization of methodology. *Glob Cardiol Sci Pract* **2024**, e202435 (2024).

33. Väliaho, E. S. *et al.* Continuous 24-h Photoplethysmogram Monitoring Enables Detection of Atrial Fibrillation. *Front Physiol* **12**, (2022).

34. Charlton, P. H., Kyriacou, P., Mant, J. & Alastruey, J. Acquiring Wearable Photoplethysmography Data in Daily Life: The PPG Diary Pilot Study. *Engineering Proceedings 2020, Vol. 2, Page 80* **2**, 80 (2020).

35. Pitton Rissardo, J. & Letícia Fornari Caprara, A. Cardiac 123I-Metaiodobenzylguanidine (MIBG) Scintigraphy in Parkinson's Disease: A Comprehensive Review. *Brain Sci* **13**, 1471 (2023).

36. Merrick, J., Grippo, A. J., Bartlett, G., Shaffer, F. & Ginsberg, J. P. An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health* **5**, 258 (2017).

37. Memon, A. A. *et al.* Heart rate variability during sleep in synucleinopathies: a review. *Front Neurol* **14**, (2024).

38. Heimrich, K. G., Lehmann, T., Schlattmann, P. & Prell, T. Heart Rate Variability Analyses in Parkinson's Disease: A Systematic Review and Meta-Analysis. *Brain Sciences 2021, Vol. 11, Page 959* **11**, 959 (2021).

39. Poewe, W. Non-motor symptoms in Parkinson's disease. *Eur J Neurol* **15**, 14–20 (2008).

40. Burq, M. *et al.* Virtual exam for Parkinson's disease enables frequent and reliable remote measurements of motor function. *npj Digital Medicine 2022 5:1* **5**, 1–9 (2022).

41. Stewart, C. B. *et al.* The longitudinal progression of autonomic dysfunction in Parkinson's disease: A 7-year study. *Front Neurol* **14**, 642 (2023).

42. Visser, M., Marinus, J., Stiggelbout, A. M. & van Hilten, J. J. Assessment of autonomic dysfunction in Parkinson's disease: the SCOPA-AUT. *Mov Disord* **19**, 1306–1312 (2004).

43. Goetz, C. G. *et al.* Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders* **23**, 2129–2170 (2008).

44. Claes, K. *et al.* TSDF: A simple yet comprehensive, unified data storage and exchange format standard for digital biosensor data in health applications. (2022).

45. PPG-DaLiA - UCI Machine Learning Repository. https://archive.ics.uci.edu/dataset/495/ppg+dalia.

46. Hernán, M. A. Methods of Public Health Research — Strengthening Causal Inference from Observational Data. *New England Journal of Medicine* **385**, 1345–1348 (2021).

# 1 Supplementary methods

## 1.1 Development of PPG morphology classifier

The goal of the classifier was to distinguish high quality PPG data from poor quality, based on short (6s) PPG segments derived from manually annotated 1-minute epochs. The classifier was trained on handcrafted features extracted from these segments and evaluated using a nested cross-validation procedure to ensure generalizability.

### 1.1.1 Annotation pipeline

The annotation pipeline relied on visual inspection by experts, supported by consistent signal visualization. First, PPG recordings were divided into 1-minute epochs. Then, filtered signals were visualized alongside their spectrograms using a custom graphical user interface (GUI). Annotations were performed independently by two trained investigators.

#### *1.1.1.1 Annotation data set*

A representative sample of 20 PPP subjects was selected from the data set to create the annotation set. Stratification was performed based on the following characteristics: sex, age and resting tremor scores (UPDRS III: Motor examination in OFF state). From each subject, 1-minute PPG epochs were selected from two timepoints: week 1 and week 52. For each week, annotation was performed on the first full 24h of the PPG recording (1440 epochs). This approach ensured that the total annotation set (57600 epochs) included variability in both temporal aspects and individual signal characteristics.

#### *1.1.1.2 Data quality labeling*

All PPG recordings were divided into non-overlapping 1-minute epochs, which allowed for sufficient context for reliable visual interpretation of PPG morphology while remaining also practical for large-scale annotations. We determined high quality PPG data as the typical waveform as described in literature[1–3]. Each epoch was assigned to one of the initial four quality categories, defined by the proportion of the signal displaying the typical morphology:

- Label 1 – Very high quality: ≥95% of the epoch
- Label 2 – Moderate quality: 50-95%
- Label 3 – Low quality: 10-50%
- Label 4 – Very poor quality: <10%

#### *1.1.1.3 Annotation protocol*

Annotations were performed using a custom GUI in MATLAB. Each 1-minute epoch was shown with:

- Time-domain PPG signal (filtered, fixed amplitude scale)
- Spectrogram (computed using pspectrum, 3s time resolution, 50% overlap)
- Timestamp (absolute time in the recording)

Spectrograms were min-max normalized to a range between 0 and 1 for better comparability of the dominant frequencies. Investigators underwent a training phase using three independent 24h datasets and proceeded to annotate after reaching inter-rater agreement (Cohen's kappa > 0.8).

### 1.1.2 Feature extraction

For the development of the classifier in this research, we only used the epochs annotated as label 1 (very high quality) or label 4 (very poor quality). These epochs were subdivided into 6s non-overlapping segments. From each 6s PPG segment, a set of 10 time- and frequency-domain features was constructed. The following 10 features were calculated:

- Standard deviation
- Mean amplitude of the absolute PPG signal
- Median amplitude of the absolute PPG signal
- Skewness
- Kurtosis
- Dominant frequency: calculated using Welch's method with a 3s Hann window and 50% overlap.

- Relative power: power within ± 0.2 Hz around the dominant frequency, normalized to the total power
- Spectral entropy: computed as the Shannon entropy of the normalized power spectral density
- Signal-to-noise ratio: ratio of the variance of the absolute signal (signal variance) to the variance of the signal (noise variance), as described by Elgendi et al.[2]
- Autocorrelation peak: highest non-zero-lag peak of the autocorrelation function

### 1.1.3 Training of the classifier

#### *1.1.3.1 Classifier*

A Logistic Regression (LR) classifier was selected due to its interpretability and effectiveness for binary classification.

#### *1.1.3.2 Hyperparameter optimization*

For hyperparameter optimization, a grid search approach was used to identify the best configuration of hyperparameters for the LR classifier. The following hyperparameters were tuned:

- Regularization strength (λ): varied over a logarithmic range, including 0 and log-spaced values from $10^{-6}$ and $10^{0}$. The regularization method was fixed to lasso (L1 regularization)
- Solver: the optimization process explored two solvers:
    - Stochastic Gradient Descent (SGD)
    - Sparse Reconstruction by Separable Approximation (SpaRSA)

#### *1.1.3.3 Leave-One-Out Cross-Validation (LOOCV)*

In the LOOCV, the dataset split such that one subject is held out for testing, and the remaining subjects are used for training. In the inner loop, hyperparameters are optimized through grid search based on the highest average performance across the training set. In the outer loop, the model is evaluated on the held-out subject using the optimal hyperparameters. The performance is averaged

across all iterations. Finally, hyperparameters are selected based on the average performance of the held-out validation data, and the model is retrained on the full dataset using the optimal configuration.

### 1.1.4 Results

The LR classifier achieved high performance in distinguish high-quality from poor quality segments. Supplementary Table 2 shows the classification metrics averaged over all LOOCV iterations.

**Supplementary Table 1: Performance metrics of the logistic regression classifier evaluated using nested leave-one-out cross-validation (LOOCV) on 6s PPG segments. Metrics are averaged across all LOOCV iterations.**

| Metric | Mean (± SD) |
|---|---|
| **Accuracy** | 98.4 (1.0) |
| **Sensitivity** | 98.2 (1.4) |
| **Specificity** | 98.5 (1.3) |
| **F1-score** | 98.1 (1.1) |
| **Precision** | 97.9 (1.7) |

The best performance of the final LR classifier was achieved using a LR classifier with a regularization parameters of 0.0001 using Lasso regularization and the SpaRSA solver.

**Supplementary Table 2: Final hyperparameter configuration for the logistic regression classifier after nested leave-one-out cross-validation.**

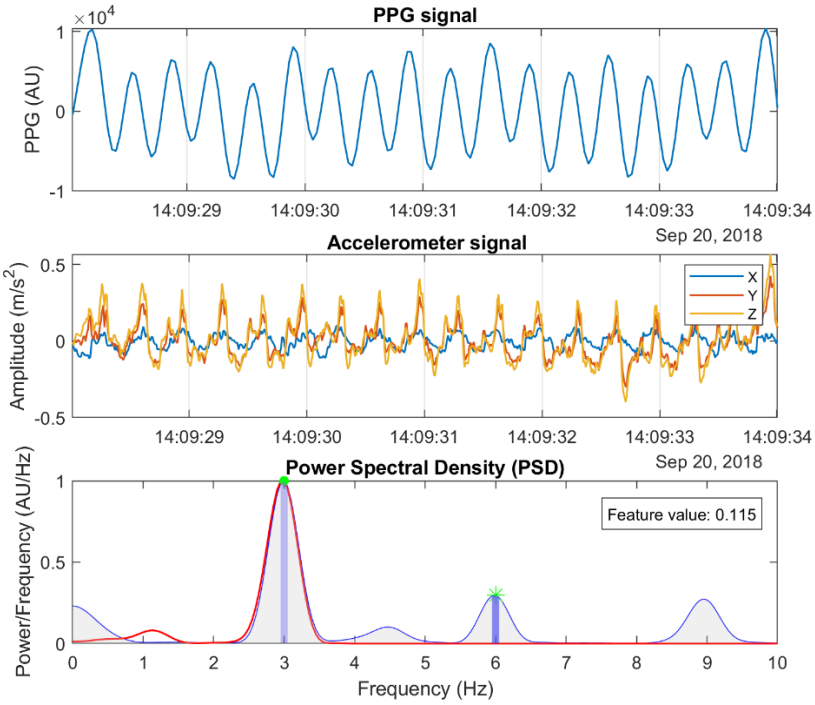
| Hyperparameter | Mean (± SD) |
|---|---|
| **Regularization strength (λ)** | 0.0001 |
| **Regularization type** | Lasso (L1) |
| **Solver** | SpaRSA |

## 1.2 Periodic motion artifact removal

### 1.2.1 Calculation of Relative Power

To identify periodic artifacts in the PPG signal caused by periodic motion artifacts, we employed a method to calculate the window-based relative power in the accelerometer signal. This process involved the following steps (also illustrated in Supplementary Figure 1):

1. **Frequency Analysis**: We performed the Welch's method on both the PPG and accelerometer signals to obtain the frequency distribution. We defined the dominant frequency as the frequency at which the signal has the highest power.
2. **Frequency Band Selection**: For each window, we focused on the dominant PPG frequency (±0.05 Hz) and its first harmonic (±0.05 Hz). The first harmonic is determined as twice the dominant frequency and is included to capture higher-order periodic components.
3. **Power Calculation**: Using the PSD estimates from the Welch's method, we calculated the power within these frequency bands for the accelerometer signal. The power within a frequency band is the integral of the PSD over that band, which we computed using the trapezoidal rule.
4. **Relative Power**: The relative power is defined as the ratio of the power within the specified frequency band to the total power in the accelerometer signal. This ratio quantifies the proportion of the signal's energy that is concentrated at the dominant PPG frequency and its harmonic.

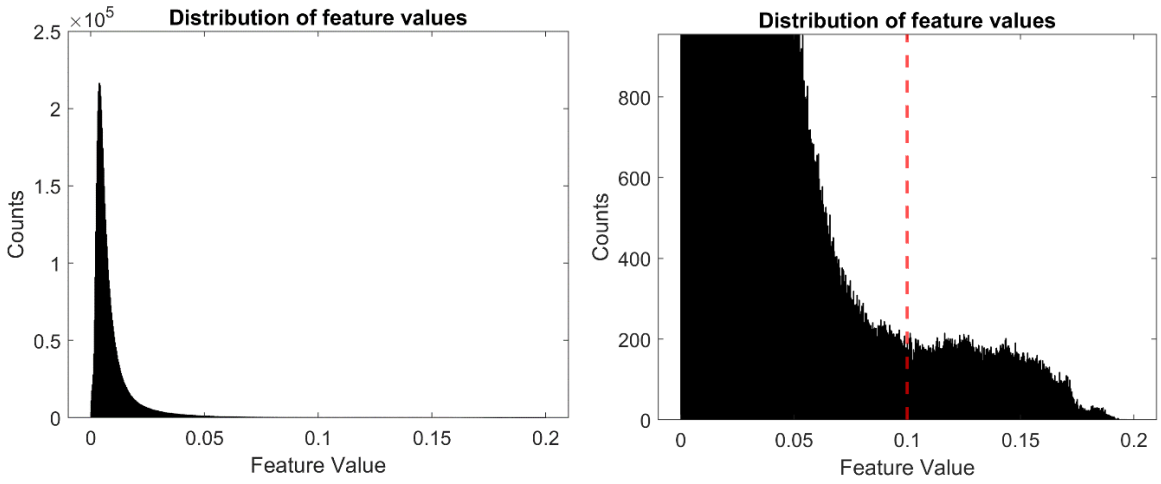


**Supplementary Figure 1: Schematic illustration of the relative power feature of the accelerometer. Upper panel: PPG signal, middle panel: three axes of the accelerometer, lower panel: visualization of the relative power feature. Frequency content of both signals is obtained using Welch's method. Thereafter, the dominant frequency (green dot) and first harmonic (green star) of the PPG signal are detected. Around both frequencies (±0.05 Hz) we calculate band powers (blue areas) using trapezoidal rule and calculate the relative power compared to the total power (grey area). In this example the feature value is 0.115.**

### 1.2.2 Threshold Selection

The threshold for classifying a window as an artifact was determined through empirical analysis of the annotation data set. The steps involved were:

1. **Training Dataset Analysis**: We analyzed the same data set as was used for developing the LR classifier to assess PPG morphology. We used every 6s segment which was classified as high-quality PPG based on this classifier for all 20 subjects in the first study weeks. This resulted in 5.2 million segments.
2. **Relative Power Distribution**: We calculated the relative power for all segments in the training dataset and examined the total distribution of relative power values (Supplementary Figure 2).
3. **Threshold Determination**: We set a threshold of 0.10 for the relative power. This threshold was chosen based on total distribution as shown in Supplementary Figure 2. While the overall distribution resembles an exponential decay, we observed a deviation from this trend around 0.10, where the feature values occur more frequently than expected. This suggests that values above this threshold may represent a qualitatively different subset of data, making 0.10 a reasonable and data-driven cut-off.




**Supplementary Figure 2: The left panel shows the overall distribution of the relative power values, following an exponential decay. The right panel provided a zoomed-in view showing a notable deviation from the exponential decay trend around the feature value of 0.10. Based on this deviation, a threshold of 0.10 was established for classifying artifacts.**

## 1.3 Signal quality classification

The LR classifier produced posterior probabilities, where higher probabilities indicated a higher likelihood of the PPG signal being of high quality, characterized by clear pulse waves with discernible peaks. For each 1-second window, the mean posterior probability was calculated, and a window was classified as high-quality if its mean probability exceeded 0.5. Accelerometer-derived binary labels were used to refine the classification by detecting periodic motion artifacts. A label of 0 indicated substantial motion artifact (above the threshold). If the majority of accelerometer labels in a 1-second window were 0, the overall signal quality label was set to 0, overriding the PPG morphology classification. For each 1-sec window, the mean posterior probability was calculated. A window was classified as high-quality if its mean posterior probability exceeded 0.5. The accelerometer-derived labels binary were used to refine the signal quality classification by detecting periodic motion artifacts. A label of 0 indicated substantial motion artifact (exceeding the threshold). If the majority-voted accelerometer label for a 1-sec window was equal to 0, the overall signal quality label was set to 0, overriding the PPG morphology classification.

## 1.4 Validation of Pulse Rate Estimation

To evaluate the most accurate approach for heart rate estimation from photoplethysmography signals, we compared six different methods, including four time-frequency analysis techniques and two beat-detection algorithms. While beat-detection methods directly identify individual pulses, time-frequency methods offer a continuous representation, which can be more beneficial In cases of noisy signals. To ensure reliable comparisons, we applied the signal quality algorithm to extract high-quality PPG segments of at least 30s from an external dataset, containing synchronized PPG and ECG recordings during free-living activities. Pulse rate was then estimated on these segments for every 2s, using each of the six methods. By comparing both approaches, we aimed to assess their strengths and determine which method provides the most reliable estimates under ambulatory conditions.

### 1.4.1 Validation dataset

The performance of each method was assessed using the PPG DaLiA dataset[4], which contains synchronized PPG and electrocardiogram (ECG) recordings from subjects engaged in daily activities. ECG-derived heart rate of every 2s (8s windows, 6s overlap) served as the reference standard.

### 1.4.2 Time frequency methods

Time-frequency distributions (TFDs) provide a representation of the PPG signal's spectral content over time, allowing for pulse rate estimation by identifying the dominant frequency component in the expected physiologically heart rate range (40-180 bpm). The following four TFD methods were considered and implemented in MATLAB:

**Short-Time Fourier Transform (STFT):** The STFT partitions the signal into windows and applies the Fourier transform to each segment[5]. This approach provides a straightforward way to analyze frequency variations. However, its resolution is limited by the fixed window length, leading to trade-offs between time and frequency precision. A shorter window improves the temporal resolution but reduces frequency precision while a longer window has the opposite effect[6]. To match the ECG labels, we used 8s windows with 6s overlap.

**Continuous Wavelet Transform (CWT):** The CWT decomposes the signal into wavelet coefficients across multiple scales, capturing both transient and oscillatory components[5]. This multi-scale

approach provides adaptability to pulse rate variations, but its resolution depends on the choice of wavelet function and may be affected by noise in non-stationary signals. For this specific analysis, the Generalized Morse Wavelet was implemented using the in-build *cwt* function.

**Wigner-Ville Distribution (WVD):** The WVD provides a high-resolution time-frequency representation by computing the time-dependent spectral energy of the signal. This method offers precise frequency localization but is highly susceptible to cross-term interference when multiple frequency components are present.[5]

**Smoothed Pseudo Wigner-Ville Distribution (SPWVD):** The SPWVD is a modification of the WVD that incorporates time and frequency smoothing to mitigate the presence of cross-terms, which can obscure the true spectral content of the PPG signal.[5] The smoothing functions improve robustness against noise, making it particularly effective in tracking pulse rate variations during noisier segments. Compared to standard WVD, SPWVD tries to maintain a high-frequency resolution without suffering from cross-terms. For the WVD and SPWVD, a fast and memory-efficient implementation of both algorithms was used.[7] The SPWVD was smoothed using 1s hamming window in the time domain and a 8s hamming window in the frequency domain, determined empirically.

### 1.4.3 Beat-detection algorithms

In addition to time-frequency distribution methods, we also evaluated two robust beat-detection algorithms[8], which aim to identify individual heartbeats from PPG signals and derive 2s pulse rate estimates based on inter-beat-intervals. To do this, we detected peaks within 8s windows with a 6s overlap. The two algorithms were:

**Multi-Scale Peak and Trough Detection (MSPTD):** This method detects peaks and onsets at multiple scales in the PPG signal to enhance robustness against noise and artifacts.

**Adapted onset detector (qPPG):** This method detects heartbeats by analyzing the steepest upward slopes in the PPG signal. It uses a slope function over a short window and applies an adaptive threshold to identify heartbeats.

### 1.4.4 Evaluation metrics

Accuracy of the different methods was assessed using the following metrics:

- Mean absolute error: The mean of the absolute differences between the 2s pulse rate estimates from PPG and the reference ECG label. This provides a general indication of accuracy.
- Median absolute error: The median of the absolute differences between the 2s pulse rate estimates from PPG and the reference ECG label. This provides an indication of accuracy without influence of outliers.
- Percentage of windows absolute error <2 BPMs: To evaluate the proportion of highly accurate estimates, reflecting reliability in estimating the correct heart rate.
- Percentage of windows absolute error > 5 BPM: To assess the frequency of large errors, indicating robustness against significant inaccuracies in heart rate estimations.

### 1.4.5 Results

**Supplementary Table 3: Evaluation metrics of the validation for six different methods to perform heart rate estimation on PPG signals. CWT = Continuous Wavelet Transform, MSPTD = Multi-Scale Peak and Trough Detection, qPPG = Adapted Onset Detector, SPWVD = Smoothed-Pseudo Wigner-Ville Distribution, STFT = Short-Time Fourier Transform, WVD = Wigner-Ville Distribution.**

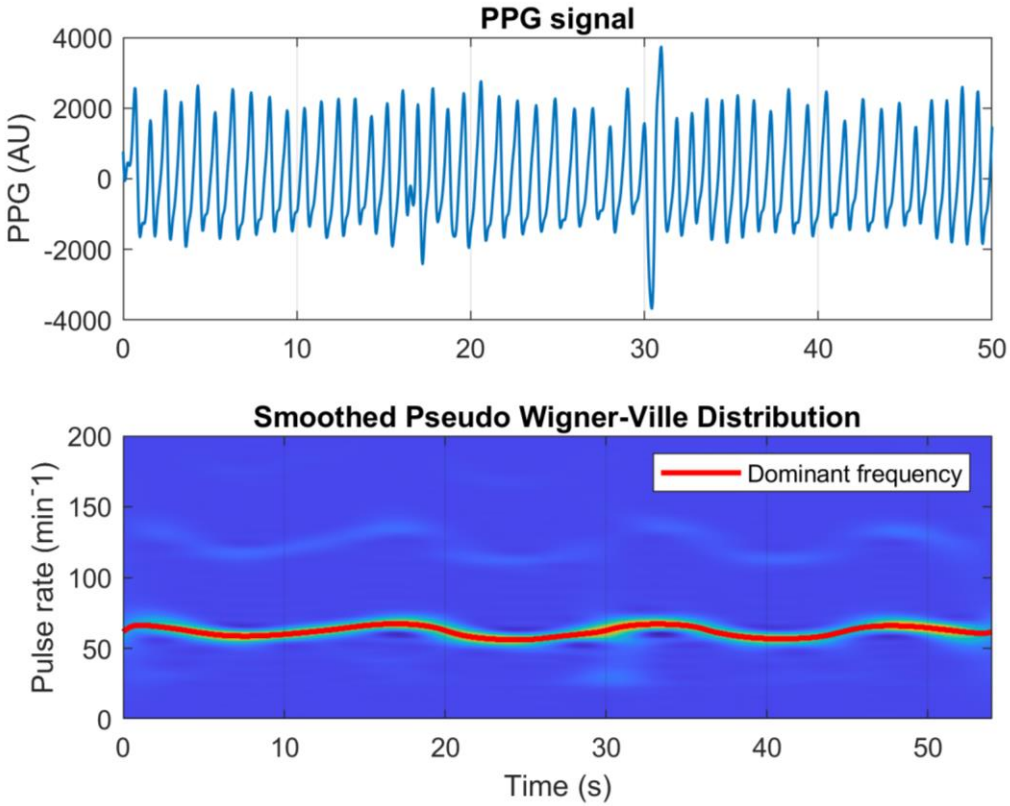
| | Time frequency analyses | | | | Beat detectors | |
|---|---|---|---|---|---|---|
| **Method** | **STFT** | **CWT** | **WVD** | **SPWVD** | **qPPG** | **MSPTD** |
| **Mean absolute error (SD)** | 1.34 (2.30) | 1.12 (2.03) | 1.34 (2.48) | 0.87 (2.11) | 1.98 (4.31) | 1.38 (2.22) |
| **Median absolute error (IQR)** | 0.80 (0.37 – 1.60) | 0.75 (0.35 – 1.33) | 0.67 (0.30 – 1.36) | 0.44 (0.20 – 0.89) | 0.93 (0.41 – 1.86) | 0.90 (0.40 – 1.71) |
| **% absolute error < 2** | 82.0 | 88.4 | 84.3 | 92.5 | 77.1 | 79.9 |
| **% absolute error > 5** | 3.0 | 1.1 | 4.8 | 1.1 | 5.5 | 2.4 |

### 1.4.6 Conclusion

The results suggest that time-frequency approaches are more reliable than the evaluated beat detection algorithms for heart rate estimation from PPG. Among these methods, the SPWVD

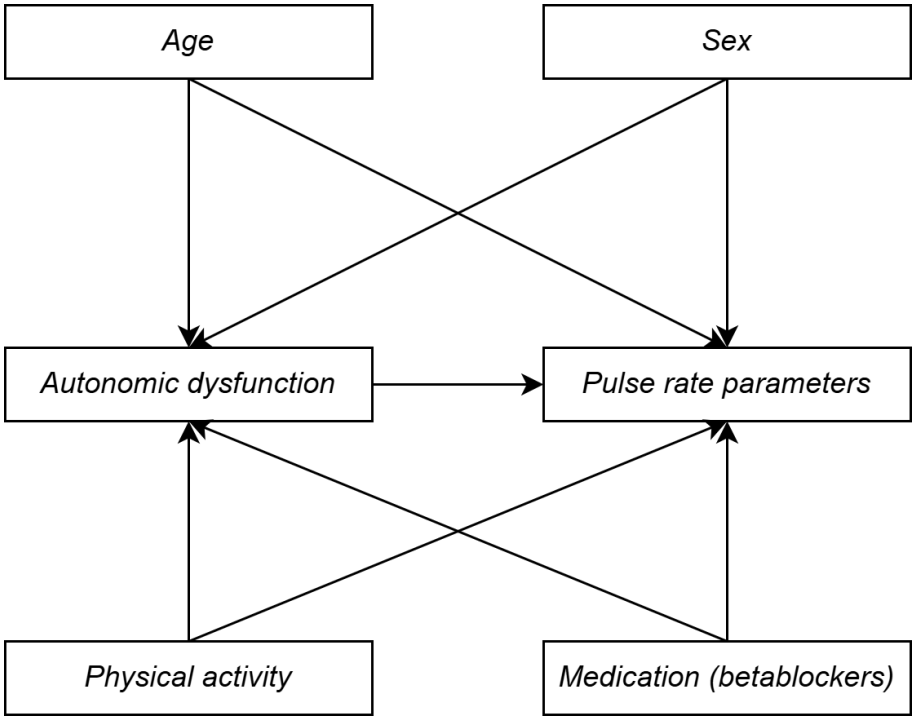
provided the most accurate pulse rate estimates, with the lowest overall error metrics and the highest proportion of accurate estimates. This highlights SPWVD as the most robust choice for analysis.

### 1.4.7 Example of SPWVD




**Supplementary Figure 3: Example of the smoothed pseudo Wigner-Ville distribution (SWPVD) applied to a the PPG signal segment. The upper panel shows the preprocessed PPG segment, while the lower panel displays its SPWVD. The red line indicates the dominant frequency at each time point. Pulse rate was estimated for every two-second window by averaging the dominant frequencies within that window.**

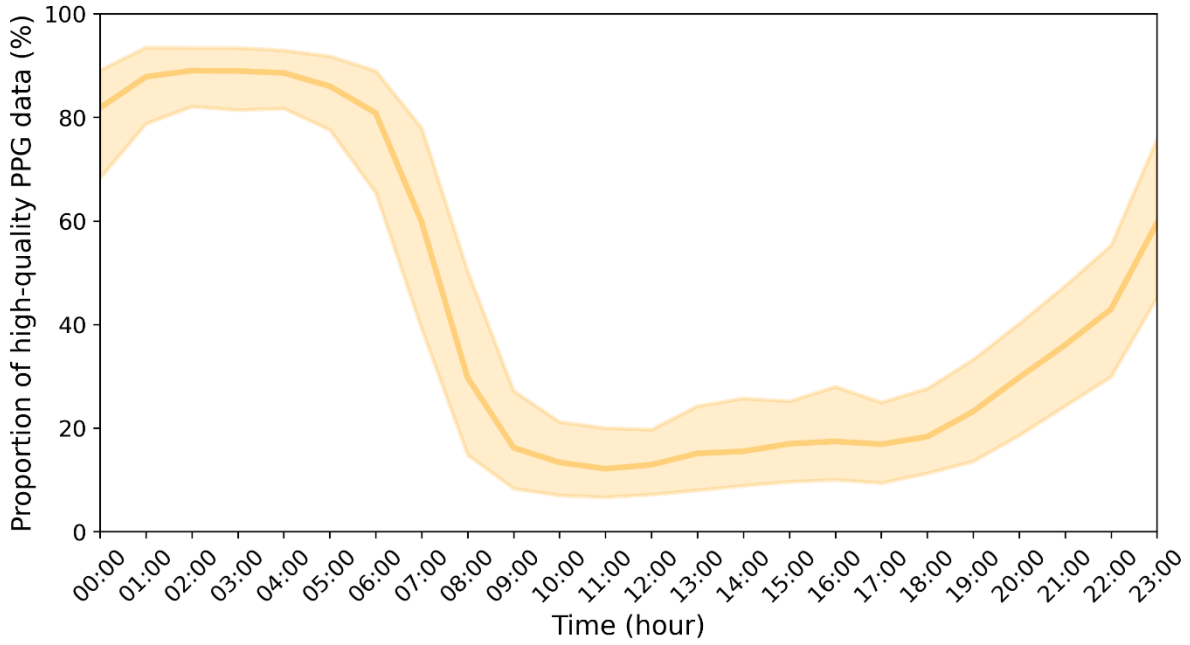
## 1.5 Directed acyclic graph pulse rate analyses




**Supplementary Figure 4: Directed Acyclic Graph illustrating the known relationships between covariates and pulse rate parameters and autonomic dysfunction. Arrows indicate the direction of influence. Age, sex, physical activity and medication (specifically beta-blockers) are included as covariates in this study model to adjust for their potential confounding effects.**

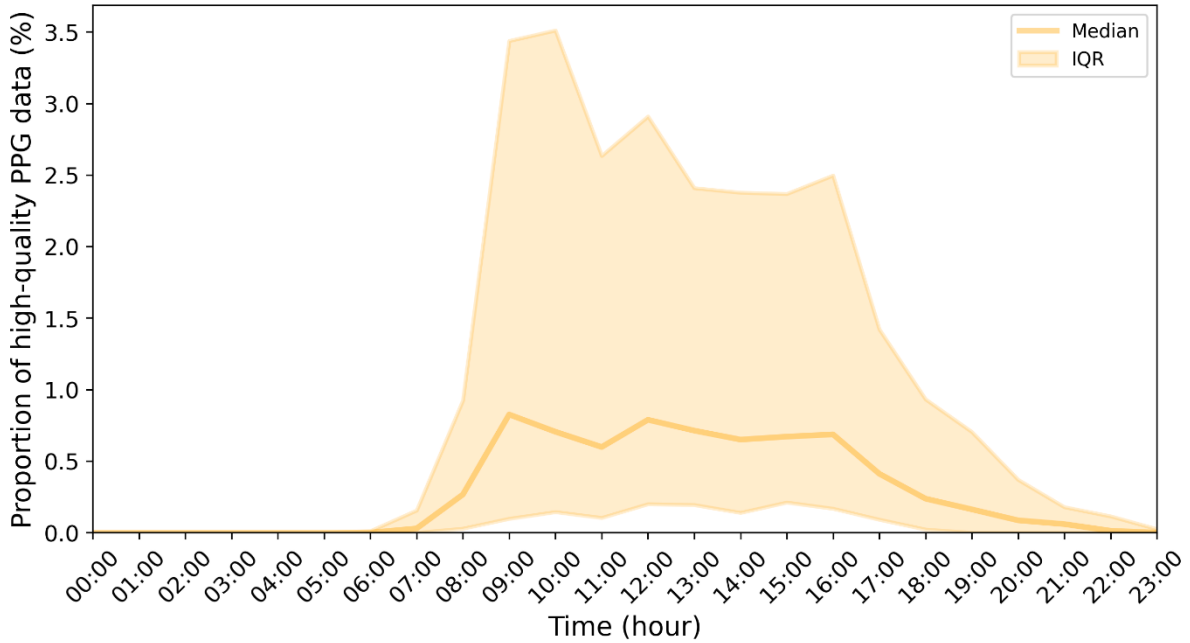
# 2 Supplementary Results

## 2.1 Data quality week 1




**Supplementary Figure 5: Proportion of high-quality PPG data across the circadian cycle in the second study week. Data are represented as medians + IQR.**

## 2.2 Influence of periodic artifact removal (step 2)




**Supplementary Figure 6: Removal of the proportion of high-quality PPG data after applying periodic artifact removal (step 2) compared to only step 1 across the circadian cycle in the first study week. Data are represented as medians + IQR.**

## 2.3 Impact of dyskinesia on signal quality

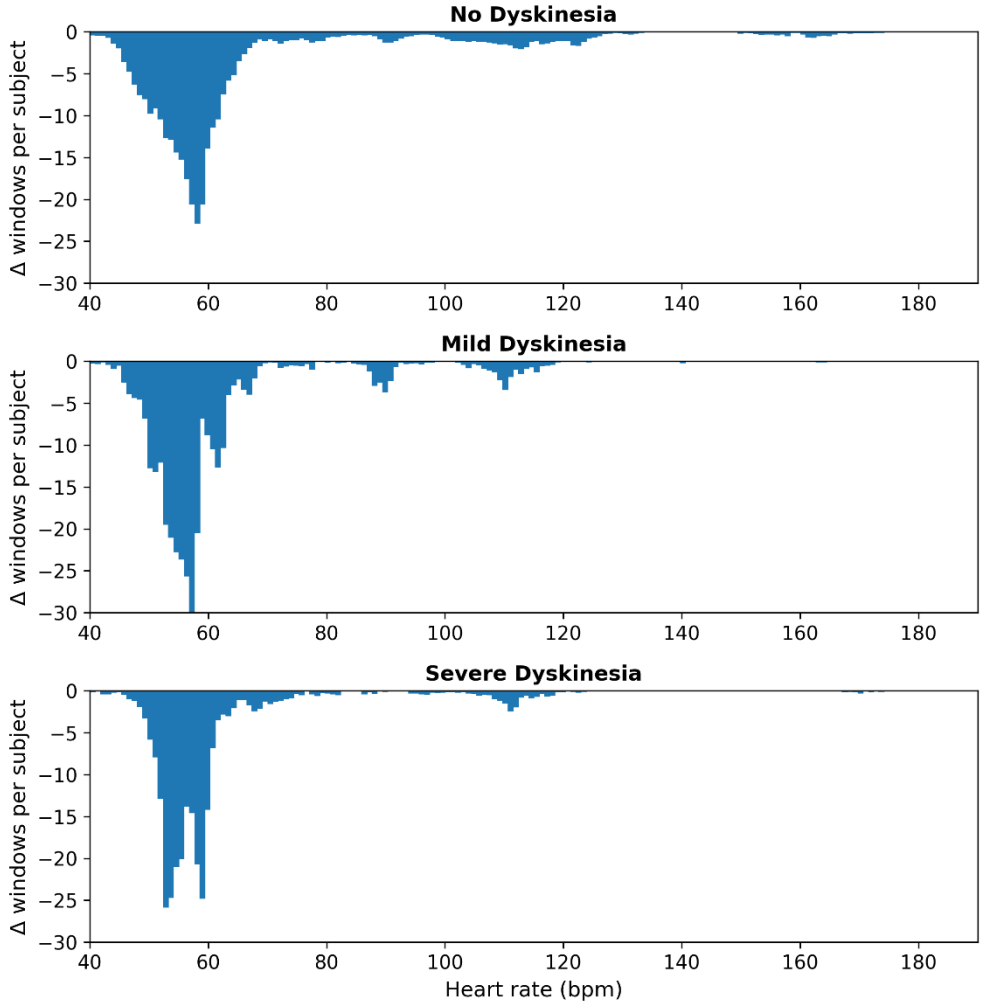
***Supplementary Table 3: Data quality across different tertiles of different rest tremor severities. Data are presented as median + IQR.***

| | No tremor (n=293) | Mild tremor (n=101) | Severe tremor (n=50) | All (n=444) |
|---|---|---|---|---|
| ***Week 0*** | | | | |
| Data quality day (%) | 28.95 [24.33, 36.29] | 29.68 [23.98, 35.67] | 30.72 [22.36, 35.86] | 29.17 [24.01, 35.90] |
| Data quality night (%) | 85.90 [79.24, 90.57] | 86.73 [78.68, 90.46] | 85.38 [79.73, 90.74] | 86.08 [79.30, 90.57] |
| Step 1 removal: Atypical morphology (PPG) (%) | 70.78 [63.48, 75.39] | 70.28 [64.23, 75.99] | 69.18 [64.07, 77.59] | 70.58 [63.97, 75.88] |
| Step 2 removal: Periodic movement (accelerometer) (%) | 0.11 [0.04, 0.24] | 0.08 [0.02, 0.23] | 0.08 [0.03, 0.20] | 0.10 [0.03, 0.22] |
| ***Week 1*** | | | | |
| Data quality day (%) | 29.48 [24.38, 36.50] | 28.97 [24.66, 34.86] | 27.34 [19.94, 37.12] | 29.01 [23.82, 36.19] |
| Data quality night (%) | 86.37 [79.39, 90.70] | 87.66 [79.19, 91.47] | 85.16 [80.65, 91.40] | 86.34 [79.37, 91.03] |
| Step 1 removal: Atypical morphology (PPG) (%) | 70.48 [63.20, 75.41] | 71.01 [65.12, 75.24] | 72.24 [62.19, 79.84] | 70.76 [63.46, 75.89] |
| Step 2 removal: Periodic movement (accelerometer) (%) | 0.12 [0.04, 0.29] | 0.08 [0.02, 0.22] | 0.09 [0.03, 0.26] | 0.11 [0.03, 0.28] |

***Supplementary Table 4: Data quality across tertiles with different dyskinesia severities. Data are presented as median + IQR.***

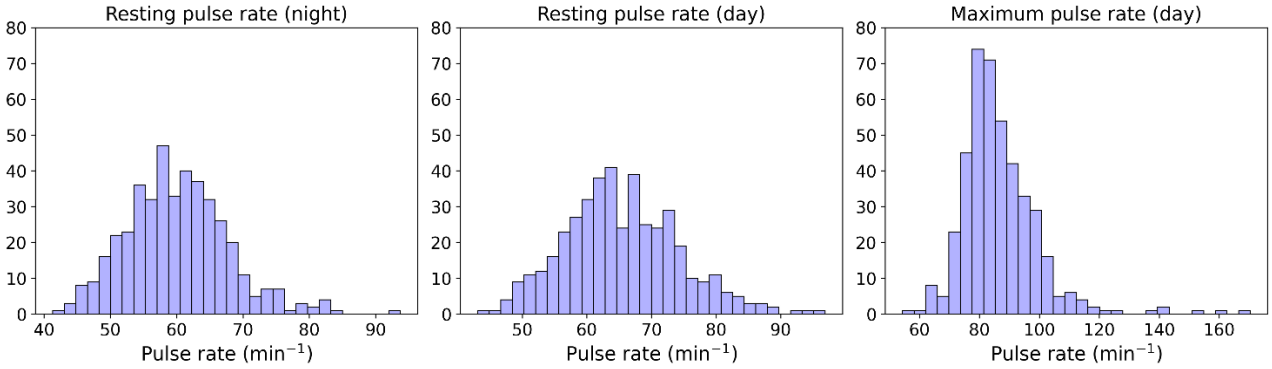
| | No dyskinesia (n = 367) | Mild dyskinesia (n = 35) | Severe dyskinesia (n = 41) | All (n=444) |
|---|---|---|---|---|
| ***Week 0*** | | | | |
| Data quality day (%) | 29.27 [24.03, 35.91] | 30.18 [25.64, 37.20] | 28.36 [21.39, 34.83] | 29.17 [24.01, 35.90] |
| Data quality night (%) | 86.16 [79.15, 90.54] | 84.56 [82.11, 89.61] | 85.92 [78.52, 91.07] | 86.08 [79.30, 90.57] |
| Step 1 removal: Atypical morphology (PPG) (%) | 70.42 [64.01, 75.81] | 69.79 [62.48, 74.29] | 71.43 [64.50, 78.45] | 70.58 [63.97, 75.88] |
| Step 2 removal: Periodic movement (accelerometer) (%) | 0.10 [0.03, 0.23] | 0.11 [0.03, 0.24] | 0.10 [0.03, 0.20] | 0.10 [0.03, 0.22] |
| ***Week 1*** | | | | |
| Data quality day (%) | 29.16 [24.38, 36.39] | 28.69 [24.51, 35.58] | 28.52 [22.32, 34.45] | 29.01 [23.82, 36.19] |
| Data quality night (%) | 86.41 [79.24, 91.17] | 86.05 [80.09, 89.78] | 87.39 [82.73, 90.98] | 86.34 [79.37, 91.03] |
| Step 1 removal: Atypical morphology (PPG) (%) | 70.51 [63.31, 75.46] | 71.29 [64.29, 75.41] | 70.98 [65.42, 77.58] | 70.76 [63.46, 75.89] |
| Step 2 removal: Periodic movement (accelerometer) (%) | 0.12 [0.03, 0.29] | 0.08 [0.04, 0.13] | 0.10 [0.03, 0.18] | 0.11 [0.03, 0.28] |

## 2.4 Influence of periodic artifact removal (step 2) on pulse rate estimates

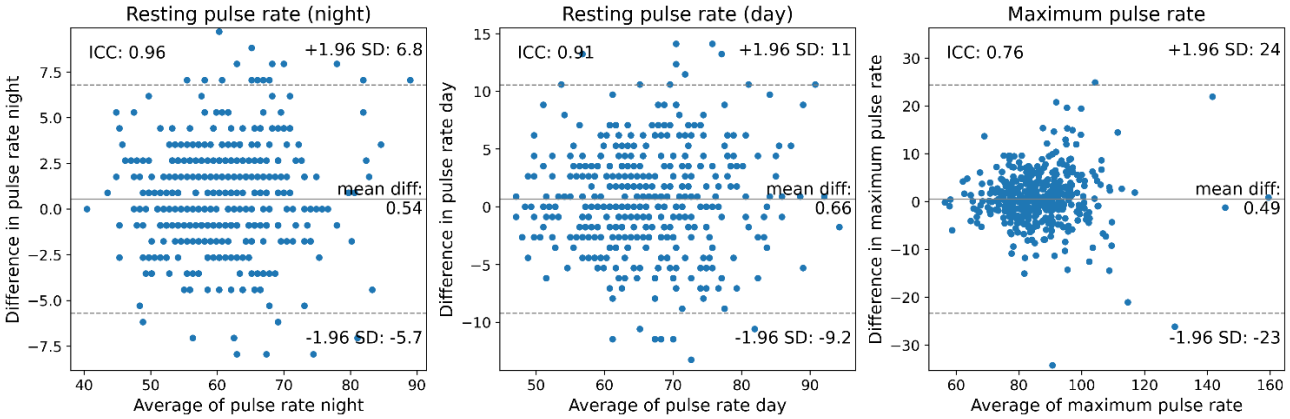



**Supplementary Figure 7: Removed pulse rate estimates by applying step 2 of the signal quality algorithm, for the different dyskinesia tertiles. The data are represented as the mean absolute removal per subject per pulse rate value. There is no association between removed pulse rate estimates and dyskinesia severity.**

## 2.5 Pulse rate parameters - descriptives

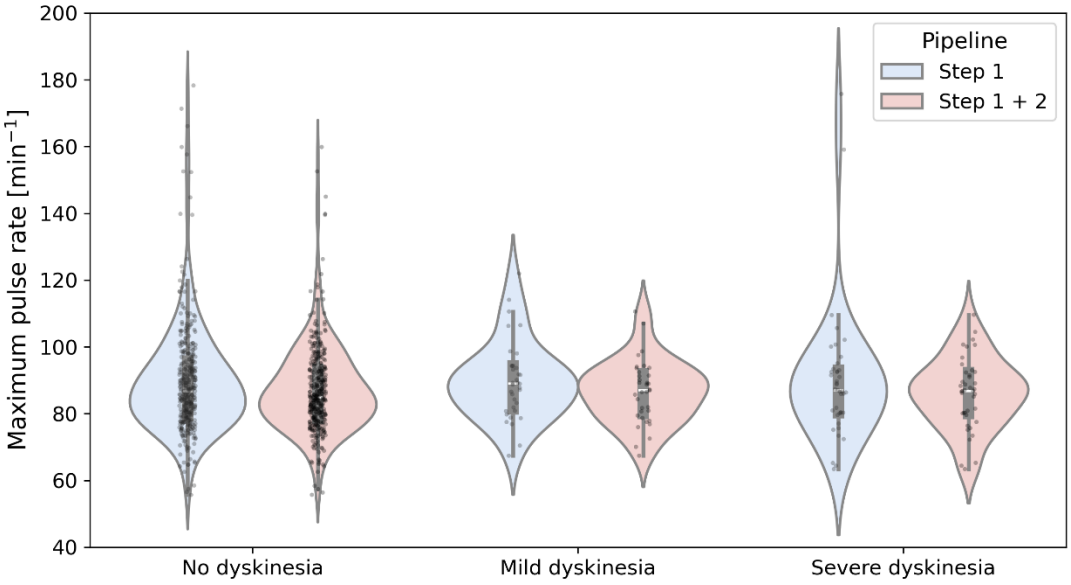



**Supplementary Figure 8: Distributions of the three different pulse parameters in the first study week across the study population.**

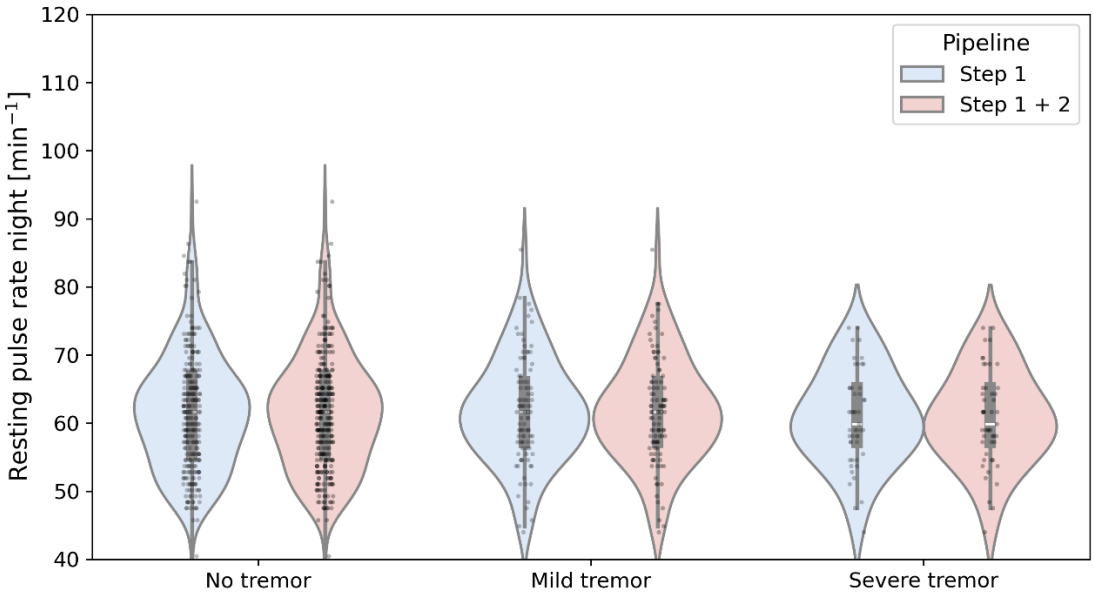



**Supplementary Figure 9: Bland-Altman plots of the pulse rate parameters in the first two study weeks with the intraclass correlation coefficient.**

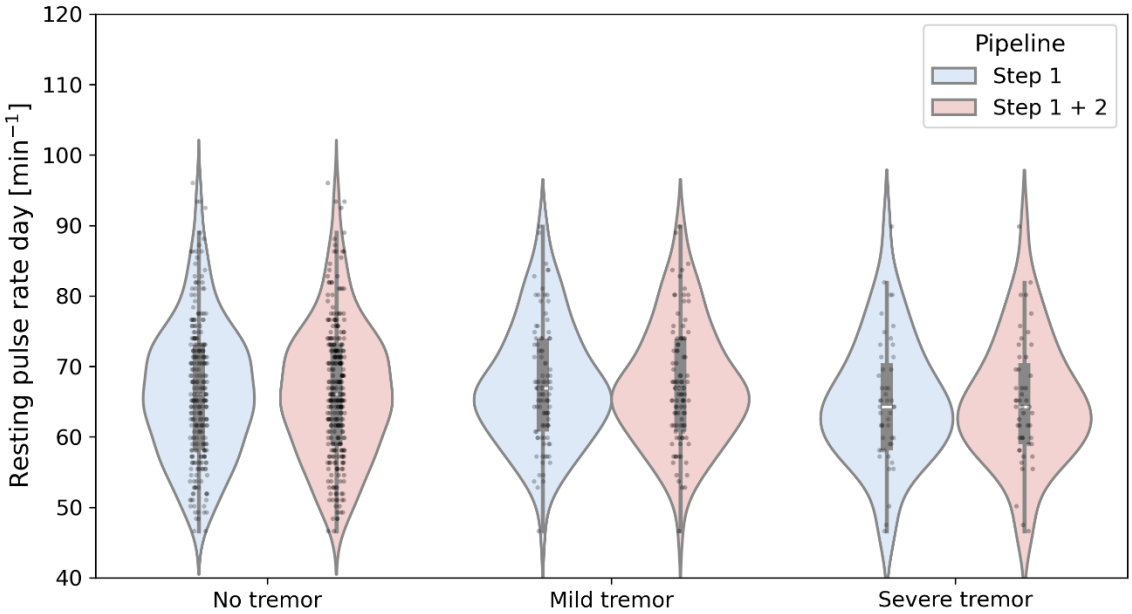
## 2.6 Influence of periodic artifact removal (step 2) on pulse rate parameters



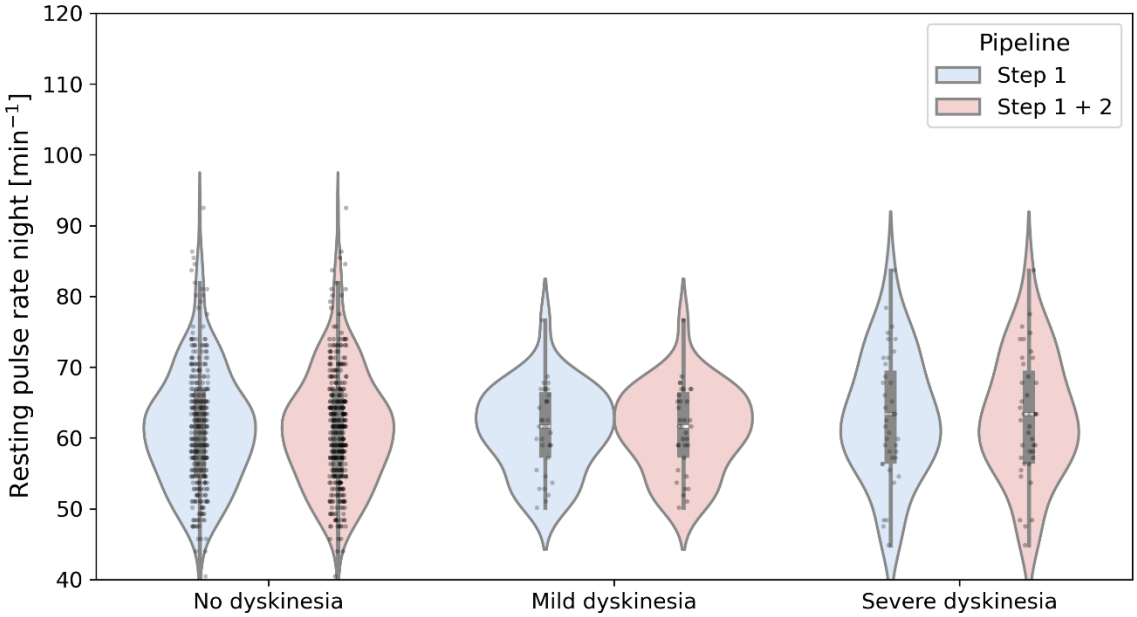

**Supplementary Figure 10: Violins of the maximum pulse rate in relation to dyskinesia severity (MDS-UPDRS 4.1 + 4.2) in the first study week. The different colors represent the maximum pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). Filtering for periodic motion artifacts does not significantly affect the aggregated maximum pulse rate at the group level, but does influence estimates at the individual level in any dyskinesia group.**



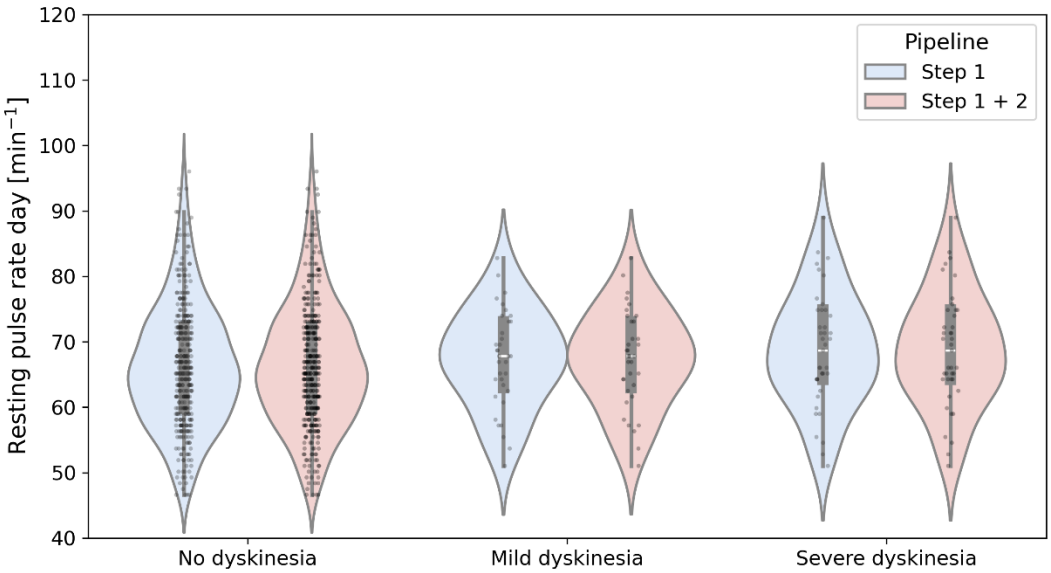

**Supplementary Figure 11: Violins of the resting pulse rate during the night in relation to tremor severity (MDS-UPDRS 3.17) in the first study week. The different colors represent the resting pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). Filtering for periodic motion artifacts does not affect the aggregated resting pulse rate in any tremor group.**




**Supplementary Figure 12: Violins of the resting pulse rate during the day in relation to tremor severity (MDS-UPDRS 3.17) in the first study week. The different colors represent the resting pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). Filtering for periodic motion artifacts does not affect the aggregated resting pulse rate in any tremor group.**

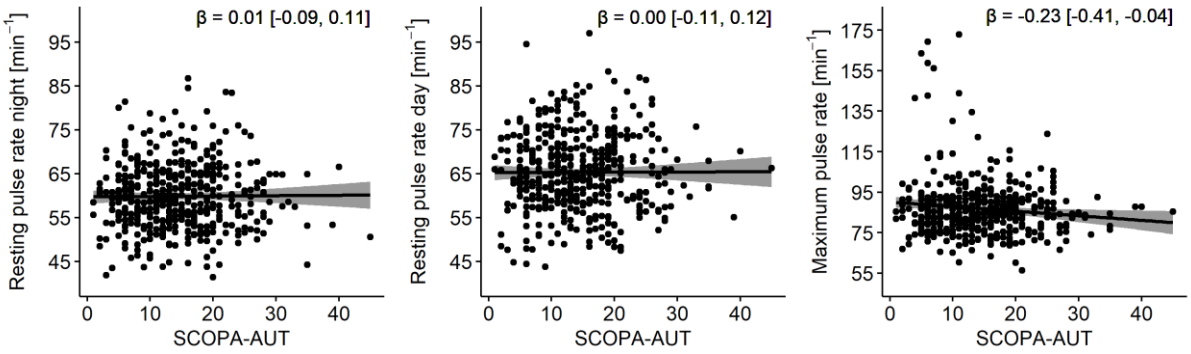



**Supplementary Figure 13: Violins of the resting pulse rate during the night in relation to dyskinesia severity (MDS-UPDRS 4.1 + 4.2) in the first study week. The different colors represent the resting pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). Filtering for periodic motion artifacts does not significantly affect the resting pulse rate in any dyskinesia group.**




**Supplementary Figure 14: Violins of the resting pulse rate during the day in relation to dyskinesia severity (MDS-UPDRS 4.1 + 4.2) in the first study week. The different colors represent the resting pulse rate when assessing solely PPG morphology (blue) or also incorporating periodic artifact removal using the accelerometer (red). Filtering for periodic motion artifacts does not significantly affect the resting pulse rate in any dyskinesia group.**

## 2.7 Regression results




**Supplementary Figure 15. Multivariate regressions between the pulse rate parameters in the second study week and SCOPA-AUT. Every data point is corrected using the following covariates: age, sex, use of betablockers and baseline physical activity.** SCOPA-AUT = SCales for Outcomes in PArkinson's disease - Autonomic dysfunction.

# 3 References


1. Charlton, P. H., Pilt, K. & Kyriacou, P. A. Establishing best practices in photoplethysmography signal acquisition and processing. *Physiol Meas* **43**, (2022).
2. Elgendi, M. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* **3**, (2016).
3. Nikhilesh Pradhan, by. Evaluation of the Signal Quality of Wrist-Based Photoplethysmography. (2017).
4. PPG-DaLiA - UCI Machine Learning Repository. https://archive.ics.uci.edu/dataset/495/ppg+dalia.
5. Scholl, S. Fourier, Gabor, Morlet or Wigner: Comparison of Time-Frequency Transforms.
6. Wright, P. S. Short-time fourier transforms and Wigner-Ville distributions applied to the calibration of power frequency harmonic analyzers. *IEEE Trans Instrum Meas* **48**, 475–478 (1999).
7. OToole, J. M. & Boashash, B. Fast and memory-efficient algorithms for computing quadratic time–frequency distributions. *Appl Comput Harmon Anal* **35**, 350–358 (2013).
8. Charlton, P. H. *et al.* Detecting beats in the photoplethysmogram: benchmarking open-source algorithms. *Physiol Meas* **43**, 085007 (2022).