Quantifyingperipheral sympathetic activations during sleep by means of an automatic method for pulse wave amplitude drop detection

M. Betta[1], *G. Handjaras*[1], *E. Ricciardi* [1], *P. Pietrini* [1],
J. Haba-Rubio [2], *F. Siclari*[2], **R. Heinzer* [2,3], **G. Bernardi*[1,2]

[1]MoMiLab Unit, IMT School for Advanced Studies, Lucca, Italy
[2] Center for Investigation and Research on Sleep (CIRS), Lausanne University Hospital(CHUV), Lausanne, Switzerland
[3] Pulmonary Department, Lausanne University Hospital(CHUV), Lausanne, Switzerland

*Equal contribution

Short Title: Detection of sleep-related pulse wave amplitude drops

Abstract word count:250
Introduction word count:849
Discussion word count: 1334
Total word count:7926

Correspondence
Monica Betta: monica.betta@imtlucca.it/monicabetta87@gmail.com
IMT School for Advanced Studies Lucca, Italy
Piazza San Francesco, 19
I-55100, Lucca, Italy

Conflict of Interest
The authors have no conflict of interest to declare.

Author contributions (CRediT taxonomy)
Conceptualization, M.B., J.H.R., R.H.,G.B., F.S..; Methodology, M.B.,G.H.; Formal Analysis, M.B.; Visualization, M.B.; Resources, E.R., R.H.; Supervision, E.R., F.S., R.H., G.B.; Writing – Original Draft, M.B., G.B.; Writing – Review & Editing, M.B., G.H., P.P., E.R., J.H.R., F.S., R.H., G.B.

ABSTRACT

Sudden drops in pulse wave amplitude (PWA) measured by finger photoplethysmography (PPG) are known to reflect peripheral vasoconstriction resulting from sympathetic activation. Previous work demonstrated that sympathetic activations during sleep typically accompany the occurrence of pathological respiratory and motor events, and their alteration may be associated with the arising of metabolic and cardiovascular diseases. Importantly, PWA-dropsoften occurin the absence of visually identifiable cortical micro-arousalsand may thus represent a more accurate marker of sleep disruption/fragmentation. In this light,an objective and reproduciblequantification and characterization of sleep-related PWA-drops may offer avaluable, non-invasive approach for the diagnostic and prognostic evaluation of patients with sleep disorders. However, the manual identificationof PWA-drops represents a time-consuming practice potentially associated with high intra/inter-scorer variability. Sincevalidated algorithms are not readily available for research and clinical purposes, here we present a novel automated approach to detect and characterize significant drops in the PWA-signal. The algorithm was tested against expert human scorers who visually inspected correspondingPPG-recordings.Results demonstrated that the algorithm reliably detects PWA-drops and is able to characterize them in terms of parameterswith a potential physiological and clinical relevance, includingtiming, amplitude, duration and slope. Themethod is completely user-independent,processesall-nightPSG-data, automatically dealing with potential artefacts, sensor loss/displacements, and stage-dependent variability in PWA-time-series. Such characteristicsmake thismethod a valuablecandidate for the comparative investigation of largeclinical datasets, to gain a betterinsightinto the reciprocal links betweensympathetic activity, sleep-related alterations, and metabolic and cardiovascular diseases.

Keywords: *autonomic nervous system, sympathetic activation, photoplethysmography, Pulse Wave Amplitude, sleep, sleep disorders.*

1.INTRODUCTION

The autonomic nervous system (ANS) plays an integral role in the fine-tuning of a great number of physiological processesthrough the complementary and synergic activity of its main divisions, namely the sympathetic and parasympathetic nervous systems.Across a night of sleep, ANS activityundergoessignificant fluctuationsin relation tosleep-stage transitions(Somers et al. 1993; Trinder et al. 2001, 2012; Whitehurst et al. 2016) as well as phasic physiological and pathological sleep-related events, including cortical arousal as well as respiratory (e.g., apnea/hypopnea) and motor (e.g., limb-movement) events(Bosi et al. 2018; de Zambotti et al. 2018). Of note, abrupt and transient autonomic activations, which lead to increased sympathetic activity and result in peripheral vasoconstriction(Bartels et al. 2016; Catcheside et al. 2002; Grote et al. 2003a; Johnson and Lubin 1967), have been shown tooccur also in the absence of visually identifiable cortical EEG arousals, and have thus been suggested to represent aneven more accurate marker of sleep fragmentation and disruption(Dresler et al. 2012; Haba-Rubio et al. 2005; Janackova and Sforza 2008; Lévy and Pépin 2003; Martin et al. 1997).Moreover, pathological conditions characterized by recurrent autonomic arousals, such as obstructive sleep apnea syndrome (OSAS) and periodic leg-movement syndrome (PLMS), are typically associated with an increased risk of cerebrovascular, cardiovascular and metabolic disorders (e.g., hypertension, myocardial infarction, stroke) (Shimizu et al. 1992)(Vargas-Pérez, Bagai, and Walters 2017). Indeed, evidence indicates that nocturnal arousal-related autonomic and hemodynamic alterations may be associated with sustained daytime sympathetic modifications(Biaggioni and Calhoun 2016; Carlson et al. 1993; Fletcher 2003; Hedner et al. 1988; Somers et al. 1993, 1995) that are in turn implicated in the etiology of the above-mentioned disorders(Brook and Julius 2000; Esler 2000; Fletcher 2003; Mark 1996; Sinski et al. 2006; Thorp and Schlaich 2015; Tsioufis et al. 2011; Vinik, Maser, and Ziegler 2011).

In light of these premises, the identification of a reliable,non-invasiveapproach for the characterization ofsleep-related autonomic activations wouldoffera new valuabletool for the diagnostic and prognostic evaluation of patients with sleep disorders. Among many different approaches for the assessment of sympathetic activity, finger photoplethysmography (PPG)emerged as a very promising candidate, since it is typically recorded during both standard polysomnographic (PSG) and polygraphic (PG) studies, and offers a portable, low-cost and non-obtrusivetechnology to continuously monitor relativevariations in peripheral blood flow in the microvascular bed of tissue(Allen 2007). The PPG-signal comprises a pulsatile physiological waveform (‘pulse wave’) attributed to cardiac synchronous changes, superimposed on a more slowly varying baseline with various lower frequency components attributed to respiration, sympathetic nervous system activity and thermoregulation(Nitzan et al. 1996, 1998, 2001). In particular, drops in pulse wave amplitude (PWA)(Korpas, Hálek, and Dolezal 2009), asmeasured at each cardiac cycle by PPG, are known to directly reflect changes in peripheral blood flow due to vasoconstriction (British Editorial Society of Bone and Joint Surgery. 1954), and may thus offer a relatively simple index of autonomicactivation (Grote et al. 2003b).In line with this, several studies already demonstrated a strong association of PWA-drops withobstructive respiratory events, like apneas and hypopneas,(Bosi et al. 2018; Grote et al. 2003a; Haba-Rubio et al. 2005; Karmakar et al. 2014),as well as with spontaneous and induced EEG arousals(Adler et al. 2013; Catcheside et al. 2002; Delessert et al. 2010).

While most PWA-drops can be easily identified in the PPG-signal, their actual detection still largely relies upon human scorers, who have to mark each event manually. Manual PWA-drops scoringis a laborious and time-consuming process that inevitably limits the potential applicability of the PWA-drop analysis to large databases of patients. Moreover, procedures based on manual scoring are known to suffer from reproducibility issues related to intra- and inter-scorer variability. This latter aspect is even more relevant in light of the fact that there is currently no consensus regarding the minimum amplitude threshold for the definition of clinically relevant PWA-drops, and that such parameterscommonlydiffer across studies and laboratories. Finally, even if some PSG-software

recently started to offer automated PWA-drop detection methods, these are often not validated and only allow the computation of simple, basic parameters, such as the overall number of drops per hour(Pulse-wave-drop index, PDI).

To overcome the above-mentioned issues, here we describea novel,automated approach for the detection and characterizationof PWA-drops from whole-night PPG data.The proposed approach allows the extraction of variousparameters of potential interest, including relative timing (used to compute PDI or the association with other scored events of interest), amplitude, descending slope, ascending slope and duration. Several data-quality checks are included in the procedure in order to automatically deal with potential artefacts caused by movement and/or sensor loss/displacement throughout the night.The performance of this PWA-drop detection approach was evaluated by comparing the detections of the algorithm with those performed manually by two expert scorers board-certified in sleep medicine.In order to further showcase the possible advantages of the automated procedure for PWA-drop detection, we also applied the algorithm to investigate relative differencesacross sleep stages in terms of PWA-drop number and morphological properties.

2. MATERIALS AND METHODS

2.1 The PWA-drop detection algorithm

The PWA-drop detection algorithm includes three main‘steps’ that are described in detail in the following paragraphs and are graphically summarized in Figure 1. First, the PWA-time-series is extracted from the raw PPG-signal and potential artifactual segments are identified and excluded from subsequent evaluations.Then,candidate PWA-drops corresponding to local peaks in the variance of the PWA time-seriesare identified.Finally, significant drops are selected among all candidates, based on a-priori defined criteria, and their main characteristics (e.g., timing, amplitude, slopes, duration,etc.) are stored for further evaluation.

2.1.1 PWA-signal extraction and preliminary artifact detection

The PPG-signal obtained from conventional pulse-oximeters, measuring pulsatile blood volume in the fingertip, is used as primary input for the algorithm. The Pulse Wave Amplitude ($PWA(i)$) time-seriesis thendefined,at each cardiac cycle (i), as the difference between the maximum (peak) and minimum (nadir) values of the correspondent blood volume pulse-wave. Prior toPWA-signal extraction, the PPG-signal is smoothed (Savitzky-Golay filter with order 2 and 200 msspan) and constant and linear trends are removed.

Prior to the actual PWA-drop detection, segments of the PWA-signal containing potentialartefacts are automatically identified and excluded from further evaluations. Specifically, an automatic procedure is used to exclude time-points for which: i) the PPG-waveform does not clearly show a consecutive maximum and minimum couple, but only two or more consecutive local maxima or minima (in other words, the blood volume pulse-wave of a particular cardiac cycle is not clearly identified); ii) the temporal distance from the previous PWA-time-point is not consistent with the physiologically plausible range of values for the heart-rate (a fixed threshold correspondingto 250 heart-beats per minuteisused); iii) the PWA-valueat a specific time-point shows differenceswith respect to both the previous and the following PWA-time-pointsthat differ significantlyfrom the total distribution of difference values computed between all consecutive PWA-points (Modified Thompson Tau Test (Thompson 1985)).Moreover, in order to take into account the possibility of sensor loss or temporary displacement (e.g., due to subject movements throughout the night), the time-course of the root-mean-square (RMS)envelope of the PPG-signal is evaluatedusing a 100

samples moving-window, and PWA-drop detections are prevented within tracts in which the RMS-value falls below a predefined arbitrary threshold that should be adjusted depending on the particular instrument used (here it was empirically set to 5 μV).

2.1.2 Definition of the baseline mask and identification of candidate PWA-drops

Theobtained PWA-signal is smoothed using a moving average filter with a 5 heart-beats span. Then, the time-varyingPWA local variance and first derivative arecomputed using a 5 heart-beats widthwindow moving with steps of one heart-beat. The time-courses of these PWA-features areusedbothfor the definition of a '*baseline mask*'(i.e., stable segments of PWA-signal used as reference for the detection of relative variations) and the identification of candidate time-points including potentialPWA-drops.

The baseline mask corresponds to the ensemble of all the 'stable' PWA-signal tracts lasting at least 2 consecutive heart beats, and is obtained by excluding all time-points corresponding to local PWA-variance-outliers, as computed with respect to the whole-night local variance distribution (Modified Thompson Tau Test (Thompson 1985)).Of note, the baseline mask includes no drops or discontinuities and is also progressively updated during the PWA-drop detection procedure by removing periods containing confirmed PWA-drops.

Candidate time-points corresponding to potential PWA-drops are insteaddefinedas local peaks in the time-course of the PWA-local-variance that simultaneously display correspondent negative values for the first-derivative estimates. As detailed below, these candidate PWA-drops are subsequently confirmed only if they fulfil specifica-priori conditions. For each candidate time-point (i_c) an observation interval is defined, ranging from the closest previous (i_{obs1}) to the closest following (i_{obs2}) maxima of the smoothed PWA-signal(PWA_f).Absolute percentsignal decreases$P_{\%}$are then computed for each time-point within theobservation intervalwith respect to a baselinevalue(b), defined as the mean of the closest previous 5 PWA-points belonging to the baseline mask ($PWA(i_{obs1})$, $PWA(i_{obs2})$, …, $PWA(i_{obs5})$.).Of note, in the case of consecutive PWA-drops,stable baseline points may be availableonly far from the considered candidatedrop, and this could result in an inaccurate estimate of the PWA-drop amplitude. For this reason, if the distance between the candidate pointand the baseline points(N)is greater than 10 heart beats, all the points belonging to this separating interval are considered, in addition to baseline points, in order to compute b.

Baseline definition:

$$if\ N < 10\ ,\ b = \frac{\sum_{i=1}^{5} PWA(i_{b1})}{5}$$

$$if\ N > 10\ ,\ b = \frac{\sum_{k=1}^{5} PWA(i_{bk}) + \sum_{k=1}^{N} PWA(i_{b5}+k)}{5+N} \qquad (1)$$

Where $i_{b1}, i_{b2} \ldots i_{b5}$ represent the closest time-points belonging to the baseline mask and smaller than i_{obs}, and $N = \ i_{obs1} - i_{b5} - 1$ (number of time-points included between the baseline points and the current drop). Thus, the percent signal decrease is then computed as:

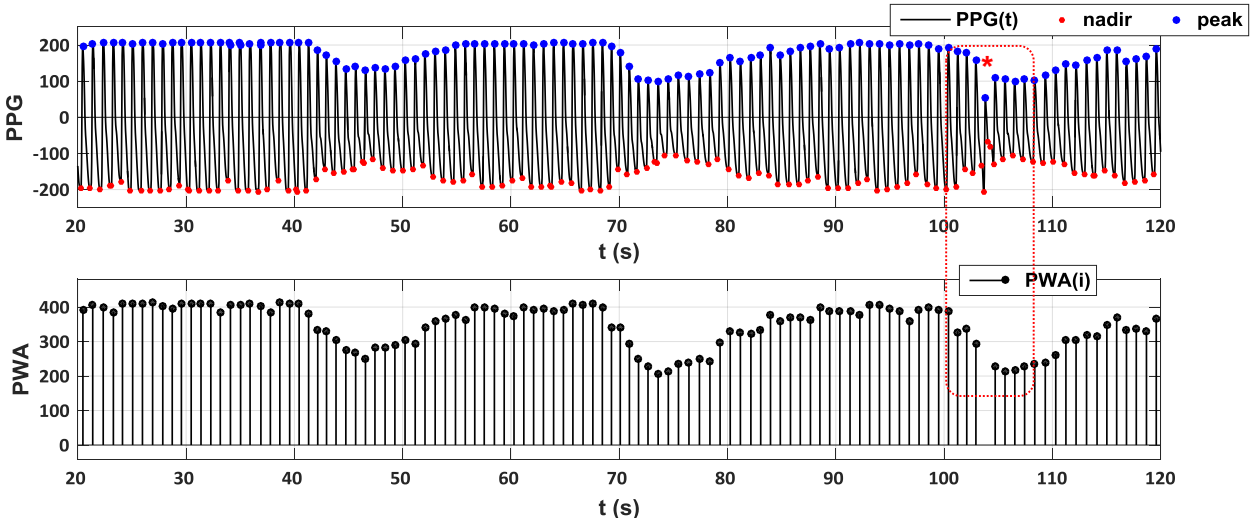
$$P_{\%}(i) = \left|\frac{PWA(i)-b}{b}\right|,\ i = \ i_{obs1}, i_{obs1}+1, \ldots\ i_{obs2};\ (2)$$

200 *2.1.3 Selection and characterization of the 'significant' PWA-drops*

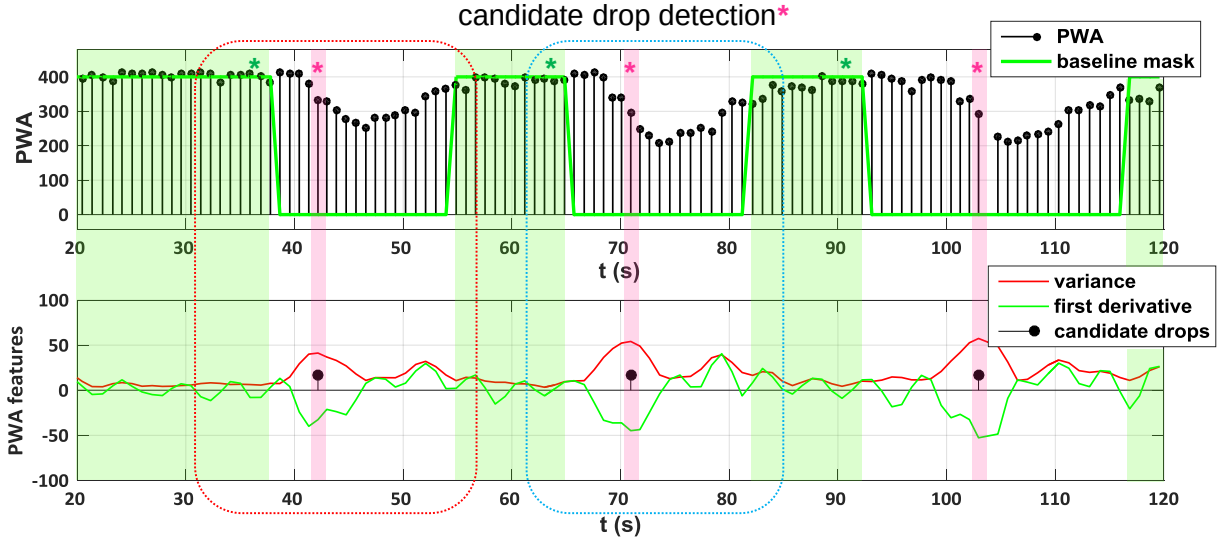
201 A 'candidate' PWA-drop is ultimately confirmedas a significant drop if the following empirical
202 criteria are simultaneously fulfilled for the corresponding observation interval: i) at least two time
203 points separated by less than 2 heart-beats show a percent decrease$P_{\%}$ greater than a
204 specificthreshold$T$, selected by the user ($P_{\%} > T$); ii) at least fourtime points separated by less than
205 2 heart-beats show a percent signaldecrease $P_{\%}$greater than $T - 10$.Indirectimplication of these
206 criteria is that,in order to be eventually detected, a PWA-drop must be at least 4-heart-beats long.

207

(a) *PWA signal extraction and preliminary artifact detection* *



(b) Time-varying PWA feature evaluation:baseline mask definition* and candidate drop detection*



(c) *Drop selection among candidates*

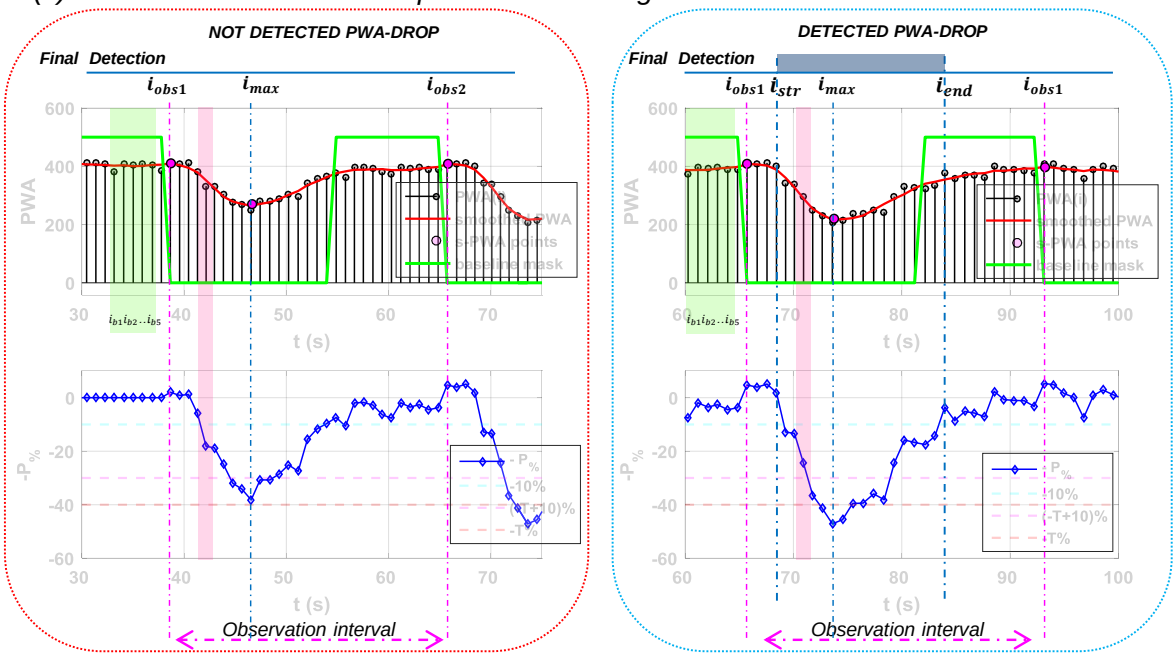
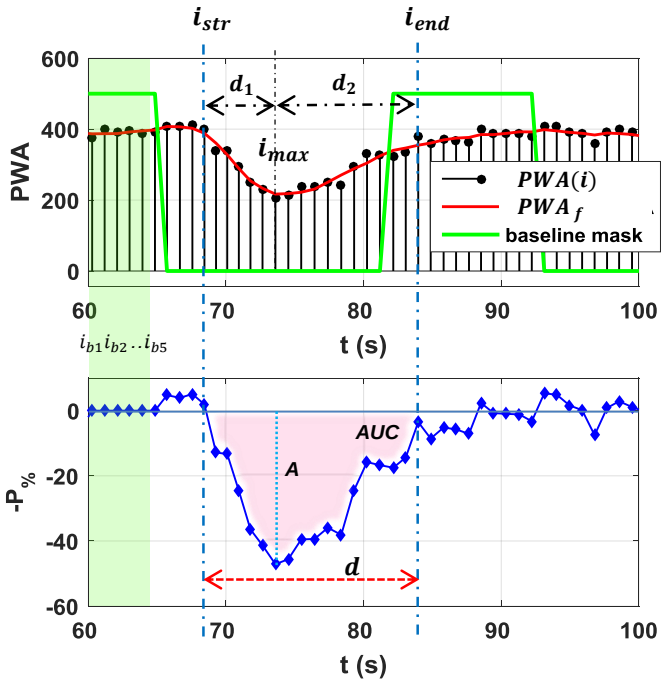


Figure 1.PWA-drop detection algorithm.Panels (a),(b) and (c) report representative examples of the main stepsof the PWA-drop detection algorithm. (a) First, the PWA-signal ($PWA(i)$) time-series (lower row)is extracted from the raw PPG-data (upper row). The red * indicates a representative exampleof an artefactual signal-change that is automatically identified and removed from the final PWA-time-series. (b) Local variance and first derivative (respectively drawn in red and green in the second row) are then evaluated for the PWA-signal in order to define a baseline mask (green line in the upper row) and detect candidate time-points for possible PWA-drops (indicated with magenta color).Finally, in panel (c), two examples of candidate drops are reported, with, below, the correspondent values of percent signal decrease (blue diamonds) in respect to the current baseline (mean value within the green shaded area). Only the candidate in the right panel fulfills the given conditions and it is eventually detected.

Once a PWA-drop is finally detected its temporal extension is more precisely re-defined within the observation interval byusing as a reference the time-point corresponding to the greatest percent decrease(i_{max}, local minima in the smoothed PWA-signal). In particular, the drop starting point is defined within the $i_{obs1}: i_{max}$ range as the first time point i_{str} after which $P_{\%}$continuouslyremains greater than 10% ($P_{\%}(i) \geq 10\% \ \forall \ i_{str} \leq i \leq i_{max}$). The ending point of each drop i_{end} is instead selected in the range $i_{max}: i_{obs2}$, in correspondence of the first time-point for whichat least one of the following conditions becomes true: i) $P_{\%}$ falls below 10%; ii) the PWA-first-derivative approaches zero; iii) the duration of the ascending tract of the PWA-drop becomes greater than 30 heart-beats. The reason for posing multiple conditions for defining the dropending-point is related to the fact that the PWA-signaloften does not return to previous baseline values, but rather 'resets' to a different stable value after the drop. Moreover, mainly in correspondence of sleep-stage transitions, the PWA-signal may undergo slow variations that could reflect a stage-dependent adaptation of the basal sympathetic activation level, rather than a sudden variation.

For eachconfirmed PWA-drop a set of properties are estimated as shown in Figure 2. These properties include: the total duration d(time-interval comprised between i_{str} and i_{end}), the amplitude A (defined as the maximum absolute signal percent decrease within the drop),and the descending and ascending slopes (defined respectively as the decrement and increment in $P_{\%}$values divided by the time expressed in seconds). The area under the curve (AUC) of the PWA-drop (computed approximating the integral of the absolute instantaneous percent decrease over the time expressed in seconds via the trapezoidal method), which depends on both the duration and the amplitude, is also computed for each event.



$$slp_1 = \frac{100 \cdot \left(\frac{(PWA_f(i_{max}) - b)}{b} - \frac{(PWA_f(i_{str}) - b)}{b}\right)}{d_1}$$

$$slp_2 = \frac{100 \cdot \left(\frac{(PWA_f(i_{end}) - b)}{b} - \frac{(PWA_f(i_{max}) - b)}{b}\right)}{d_2}$$

A	%
d	s
slp_1	$\%/s$
slp_1	$\%/s$
AUC	$\% \cdot s$

Figure 2 *.Description of the parameters extracted for each PWA-drop detected by the algorithm.The amplitude A, the duration d and the area under the curveAUC are graphically represented on the left, while the descending $slp_1$and ascending slp_2 slopes are mathematically defined on the right. Finally, the table on the right summarizes the unit of measurement for each parameter.*

2.2 Validation of the algorithm in a clinical dataset

2.2.1 PSG recordings and manual scoring

The detection algorithm was applied to the PPG-data extracted from the PSG overnight recordings of 16 patients (age 50.9 ± 6.33 yrs, 13F) randomly sampled from the *HypnoLaus*Sleep Cohort database, collected between 2009 and 2013 in Lausanne, Switzerland (Heinzer et al. 2015). The following selection constrains were imposed: absence of excessive daytime sleepiness (as measured using the Epworth Sleepiness Scale); BMI < 25 kg/m2; absence of hypertension, diabetes, metabolic syndrome and current or past cardiovascular diseases in the last 4 years; absence of self reported traffic accident in the last 4 years; absence of depression in the last 4 years. All the 16 analyzed subjects were not under psychotropic medicament affecting the central nervous system, and none of them had a prior diagnosis of a central nervous system disease. As described in previous work, all recordings took place in the patients' home environment in accordance with the 2007 AASM recommended setup specifications, using a portable PSG recorder (Titanium, Embla Flaga, Reykjavik, Iceland). Specifically, an Embla Titanium 8000J Nonin adult oximeter was used to simultaneously collect SpO2 and PPG data. SpO2 data was acquired with a sampling rate of 16Hz and a low-pass filter of 1.99Hz, while the PPG signal was collected with a sampling rate of 32 Hz and a low-pass filter of 15.9 Hz. Two trained sleep technicians manually scored the PSG recordings using the Somnologica software(version 5.1.1, Embla Flaga, Reykjavik, Iceland), according to the 2007 AASM recommendations(Silber et al. 2007)(Table 1). Moreover, two physicians board-certified in sleep medicine (hereinafter referred to as 'scorer 1' and 'scorer 2'), blind to algorithm detections, visually inspected thePPG data of each subject and manually marked individual PWA-drops (Somnologica software, 3-min windows) presenting a minimum percent signal decrease of 30%. The scorers relied on an electronic ruler provided in the software GUI next to the PPG-signal in order to assess the percent amplitude decrease of each PWA-drop. Information provided by non-PPG signals, including EEG, EMG and EOG were not taken into account during

the visual detection of PWA-drops. Obtained visual detections, were used to evaluate the accuracy of the automated PWA-drop detection algorithm, as detailed below.

PARAMETER	AVG	SD
Age	50.9	6.3
Gender	13F/3M	-
BMI	22.7	1.8
RDI	8.2	7.9
AHI	5.6	6.7
ODI	6.3	7.2
SAT	95.5	1.6
T90%	1.1	3.7
PLM	3.3	7.7

Table 1. Demographic and clinical characteristics (average and standard deviation) of included subjects. BMI = body mass index (kg/m2); RDI = respiratory disturbance index (events/h); AHI = apnea-hypopnea index (events/h); oxygen desaturation index (events/h); SAT = mean pulse oxygen saturation; T90% = percentage of total sleep time under a 90% oxygen saturation threshold; PLM = Periodic limb movement index during sleep (events/h). These parameters were calculated based on the AASM 2013 criteria. Based on AHI,10 subjects had no sleep disordered breathing (SBD), 5 subjects had mild SBD and 1 subject had moderate SBD.:

PARAMETER	AVG	SD
Total Sleeptime (min)	382.45	65.35
N1 time (min)	44.50	25.40
N2 Time (min)	152.44	36.30
N3 Time (min)	93.72	23.98
REM time (min)	91.81	29.55
N1 proportion (%)	11.66	6.37
N2 proportion (%)	39.81	5.34
N3 proportion (%)	24.92	6.82
REMproportion (%)	23.61	5.83
WASO (min)	31.56	18.92
ArousalIndex (n/h)	15.00	7.20

Table 2.*Sleep structure of the 16 subjects included in the analysis. The table includes both the total duration (min) and the relative proportion (%) of each sleep stage (N1,N2,N3,REM), as well as the number of arousals per hour and total time spent awake after the sleep onset (WASO; min). Group average (AVG) and standard deviation (SD) values for each property are reported in the first and the second column, respectively.*

2.2.2 Application of the PWA-drop detection algorithm

For each subject, the PWA-drop detectionalgorithm was repeatedly applied using different thresholds for the absolute PWA-signal decrease, with values (T)that varied between 10 and 80%, with10% steps. All other algorithm parameters were set to fixed values, as described above.For each detected PWA-drop,the following properties were computed and stored for further evaluation (Figure 2):maximum absolute amplitude A (%),descending slope$slp_1(\%/s)$,ascending slopeslp_2

298 ($\%/s$),duration in seconds (s) and area under the curve $AUC(\% \cdot s)$.The PWA-drop index (PDI),
299 corresponding to the number of detected drops per hour, was also calculated.

300

301 *2.2.3 Comparison of automated and manual scoring*

302 EachPWA-drop detected by the algorithmwas defined as either atrue positive (TP)or a false positive
303 (FP)depending on whether (TP) or not (FP) it overlapped for at least10% of its length with the
304 human scorer's detections. Cases in which a PWA-drop was detected by the human scorer but did
305 not overlapfor at least 10% of its duration with algorithm detections were marked as false negatives
306 (FN).In this context, true negative (TN) cases (generally used for computation of 'specificity')
307 could havebeen expressed as the total length of the recording thatwas free from both human and
308 algorithm detections divided by the mean duration of the PWA-drops detected by the
309 scorers.However,this kind of definition can lead to inflated specificity values, in particular when
310 target events are very rare withrespect to the total duration of the recording, as could be expected
311 for PWA-drops(Yetton et al. 2016).For this reason, we optedfor quantifying the performance of the
312 algorithmin terms ofsensitivity (or recall) and precision, separately for each human scorer, at the
313 varying of the amplitude threshold T(10-80%).Global values of sensitivity and precision across all
314 the analyzed subjects were obtained by 'concatenating' all subjects' recordings.In order to describe
315 witha singlemeasurethe overall accuracy of the algorithm with respect to the two scorers, the F-
316 score was also computed on the whole recordings. Adopted definitions for sensitivity, precision and
317 F-score are reported below.

318 $$Sensitivity = \frac{TP}{TP+FN}\ ,\ Precision = \frac{TP}{TP+FP}\ ,\ F\text{-}score = 2 * \frac{Sensitivity * Precision}{Sensitivity + Precision}$$

319

320 Analyses were performed both considering the whole period of sleep independently of the sleep-
321 stage ('ALL SLEEP') as well as separately for 'NREM' (N1+N2+N3) and 'REM' periods.

322

323 *2.2.4 Evaluation of stage-dependent differences in PWA-drop properties*

324 In order to provide an example of how the present algorithm could be applied to investigate changes
325 in the properties and distribution of PWA-drops in physiological or pathological conditions, a
326 specific analysis was conducted by evaluating stage-dependent differences in the characteristics of
327 PWA-drops. As described below (see section 3.1), this analysis was conducted using a minimum
328 absolute amplitude threshold of 40%. For each of the 16 subject included in the validation
329 procedure, the following parameters were calculated within each individual sleep stage (N1, N2,
330 N3, REM): PDI, amplitude, duration, slope-1, slope-2, AUC. Then, potential effects of the sleep
331 stage on the examined propertieswereinvestigated using a repeated measures (rm)ANOVA (the
332 sleep stage was considered as within-subjects factor; N=16).

333

334 **3. RESULTS**

335 *3.1 Performance of the algorithm with respect to human scorers*

336 The performance of the PWA-drop detection algorithm, expressed in terms of both sensitivity and
337 precision,are reported separately for each of the two human scorers,in panel (a) and panel (b) of
338 Figure 3, respectively. Panel (c) of Figure 3directly compares the meannumber of detections per
339 hour (PWA-drop index; PDI), for the two scorers and for the automated algorithm. All results are
340 reported as a function of the amplitude threshold ($T = 10\%, 20\%, \ldots, 80\%$) used for the detection

algorithm. Of note, the human scorers applied instead a rough amplitude threshold of about 30% to mark PWA-drops.

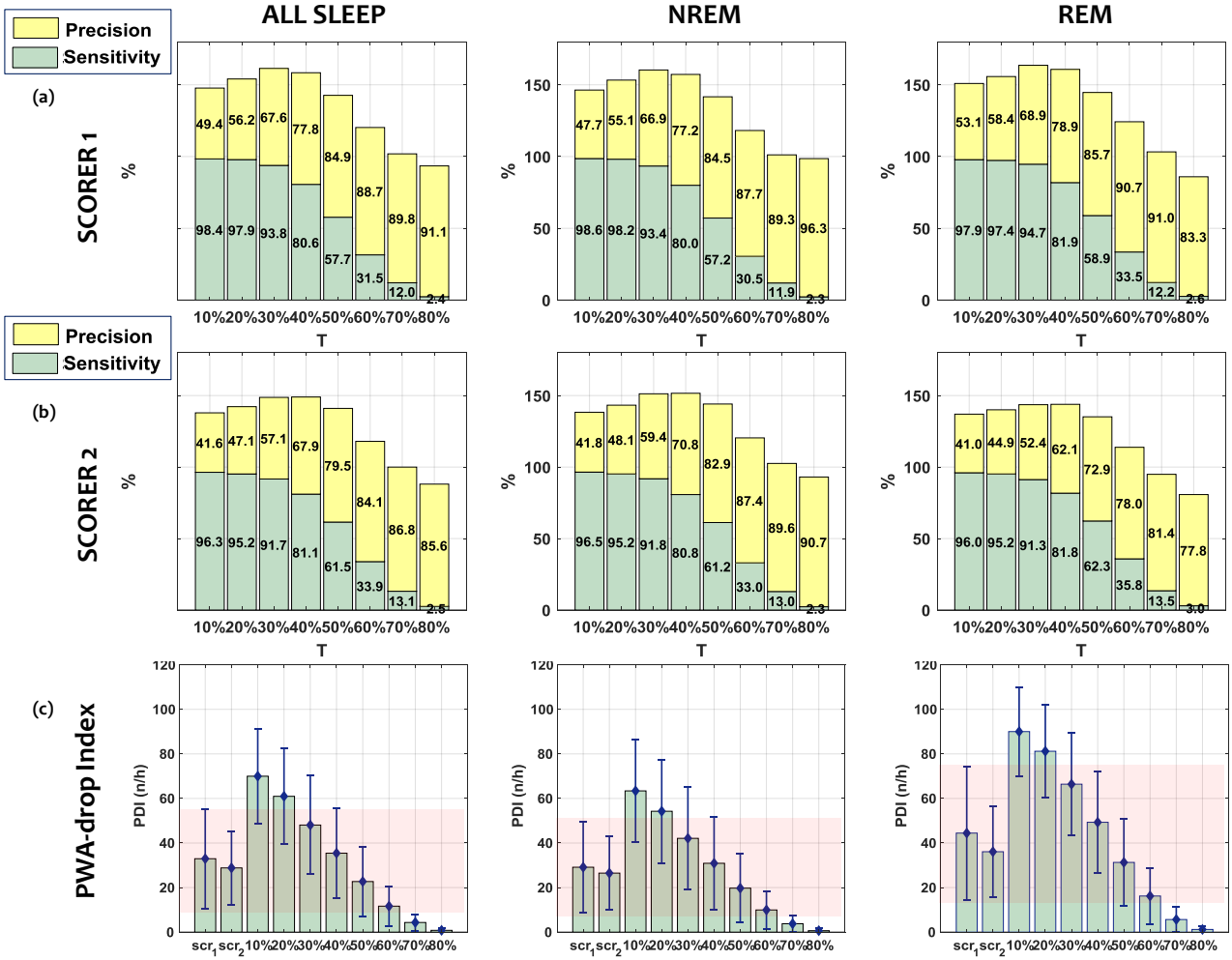


Figure 3.*Performance of the algorithm with respect to human scorers. Panel (a) and panel (b) show the sensitivity and precision of the detection algorithm with respect to scorer1 and scorer2, respectively. Values are expressed as a function of the amplitude threshold used by the algorithm (10-80% range, 10% steps). In panel (c) are instead reported the mean number of detections per hour for the two human scorers (scr$_1$ and scr$_2$) and for the algorithm.The red shaded area indicates the maximum confidence interval of the mean PDI obtained across the two human scorers and is reported to facilitate the comparison with values obtained for the automated algorithm at different threshold levels. Results are reported for the whole sleep period and separately for NREM and REM stages.*

On the whole, our results indicate that the amplitude threshold of 40% was associated with the bestoverall performance (best compromise between sensitivity and precision values) and resulted in a number of detectionscomparable to thoseof the two scorers. Mean values across scorers were 80.9% and 72.9%, for sensitivity and precision respectively (as reported in detail inTable 2). Similar values were found also when NREM and REM sleep were analyzed separately:mean sensitivity values were in fact 80.4% and 81.8% for NREM and REM respectively, while corresponding mean precision values were 74.0% for NREM and 70.5% for REM. In light of these observations, the 40% amplitude threshold was used forfurtherevaluations aimed at comparing PWA-drops detected by the algorithm with those identifiedupon visual scoring.

	ALL SLEEP	*REM*	*NREM*
Total lenght (min)	382.5 ± 65.3	290.7 ± 47.1	91.8 ± 29.6
N of PWA-drops (Scorer1)	3449 (215.6 ± 160.9)	2292 (143.2 ± 110.7)	1157 (72.3 ± 57.1)
N of PWA-drops (Scorer2)	3053 (190.8 ± 117.9)	2115 (132.2 ± 89.3)	938 (58.6 ± 40.5)
N of PWA-drops (Algorithm)	3696 (231 ± 143.3)	2455 (153.4 ± 111.8)	1241 (77.6 ± 44.6)
Inter-Scorer Agreement (F-score)	70.2 %	71.6 %	67.4 %
Accuracy Vs. Scorer1 (F-score)	78.9 %	78.3 %	80.1 %
Accuracy Vs. Scorer2 (F-score)	73.8 %	75.3 %	70.6 %
Mean Sensitivity	80.9 % Scorer1: TP = 2875, FN = 691 Scorer2: TP = 2510, FN = 589	80.4 % Scorer1: TP = 1896, FN = 474 Scorer2: TP = 1739, FN = 414	81.8 % Scorer1: TP = 979, FN =217 Scorer2: TP = 771, FN = 172
Mean Specificity	96.3 % Scorer1: TN = 18825, FP = 821 Scorer2: TN = 35967, FP = 1186	96.9 % Scorer1: TN = 14741, FP = 559 Scorer2: TN = 28011, FP = 716	94.2 % Scorer1: TN = 4077, FP = 262 Scorer2: TN = 7948, FP = 470
Mean Precision	72.9 % Scorer1: TP =2875, FP = 821 Scorer2: TP = 2510, FP = 1186	74.0 % Scorer1: TP =1896, FP = 559 Scorer2: TP = 1739, FP = 716	70.5 % Scorer1: TP = 979, FP = 262 Scorer2: TP = 771, FP = 470

Table 3. Inter-scorer agreement and detailed performance evaluation of the PWA-drop detection algorithm for T = 40 %.The algorithm-scorer agreement and the inter-scorer agreement levels are reported in terms of F-score. In addition to mean sensitivity and precision, the table also shows mean values of specificity (see section 2.2.3 of main text). For each parameter, the total number of true positive (TP) and false positive (FP) PWA-drops computed with respect to each scorer are reported. Results are divided in "ALL-SLEEP", "NREM" and "REM", according to whether they were computed on the whole sleep period, or specifically within a single sleep stage. The overall length of each considered condition is reported in the first row, while rows from 2 to 5 report the respective number of detections performed by the two scorers and the algorithm.

As shown in Table 2, the overall accuracy of the algorithmwith respect to the human scorers, expressed in terms of F-score, reached values of 78.9% and 73.8% for scorer 1 and scorer 2, respectively. Both these values were greater than the inter-scorer agreement, which corresponded to 70.2%. In other words, the overlaps betweenthe algorithm's detections and the detections of each one of the two scorers were greater than the overlap between the detection of the two scorers. Again, similar results were obtained when NREM and REM sleep were analyzed separately.

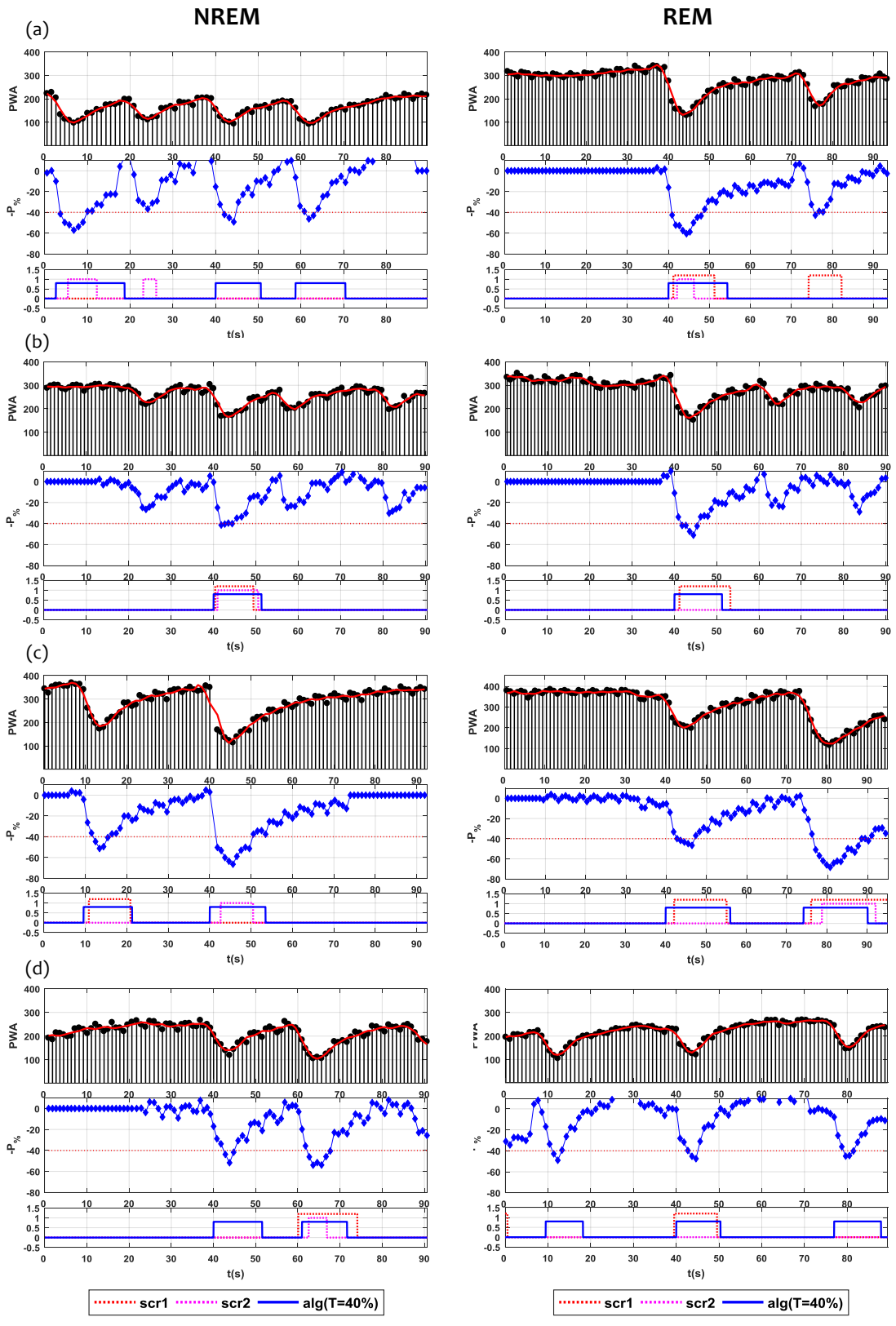


Figure 4.Examples of PWA-drops detected by the algorithm (using T=40%) and by the two human scorers, during NREM (left) and REM (right) sleep periods. Each panel includes three figures: the first row shows the PWA time-series (black dots) and its smoothed version(red line); the second row shows the correspondent time-varying percent signal decrease with respect to the preceding baseline period (blue dots); the third row shows the detections of the two human scorers (in red and magenta respectively) and the onesof the automated algorithm (in blue).

The lack of a full consistency between the algorithm and the scorers, especially in terms of precision, suggested that the detection procedure may have led to a relatively high number of false positive (FP) detections. However, a visual inspection of the PWA-signals allowed us to confirm that quite all detections performed by the algorithm corresponded to 'true' PWA-drops, and that most of the inconsistencies likely emerged from the tendency of the human scorers to detect larger, more evident drops, while neglecting the smaller ones. This is clearly illustrated by the representative examples reported in Figure 4, and by the comparison of the properties of TP and FP PWA-drops shown in Figure 5. Indeed, paired t-tests comparing the propertiesof true positive (TP) and false positive (FP) detectionsacross subjects revealed a significant difference in thearea under the curve (AUC), with greater values for TP relative to FP detections. Moreover, the human scorers tended to detect longer and more abrupt PWA-drops, characterized by a steeper descending slope (slp_1) and a shallower ascending slope (slp_2), as compared with those detected by the algorithm. Of note, however, substantial inter-scorer and inter-stage variability were also observed. For instance, a significant difference in PWA-drop duration between TP and FP detections was found only for scorer 2, while a very similar duration of TP and FP drops was observed for scorer 1.

Finally, Table 4 displays the values of accuracy, sensitivity, specificity and precision of the algorithm obtained after manual re-classification of all PWA-drops that were identified by the algorithm but were missed by at one or both the human scorers. All indices increased considerably and reached values above 90%. Indeed, we observed that 83.4% and 86.1% of algorithm detections missed respectively by scorer1 and scorer2 resulted to represent true PWA-drops. The true false positive cases were mainly caused by artefacts in the PPG signal, mostly in correspondence with movements and arousals

	ALL SLEEP	*REM*	*NREM*
Accuracy Vs. Scorer1 (F-score)	90.3 %	90.1 %	90.5 %
Accuracy Vs. Scorer2 (F-score)	91.3 %	91.0 %	91.9 %
Mean Sensitivity	90.8 %	90.6 %	91.2 %
Mean Specificity	99.6 %	99.7 %	99.3 %
Mean Precision	95.2 %	95.1 %	95.2 %

Table4. table reports performance statistics (the same of Table 3) for the algortithm (T = 40%) re-evaluated after visual reclassification of all algorithm detections marked as wrong by one or both of the two scorers. In order to re-evaluate the statistics, we have added algorithm detections reclassified as "true" to the pool of single scorer detections.

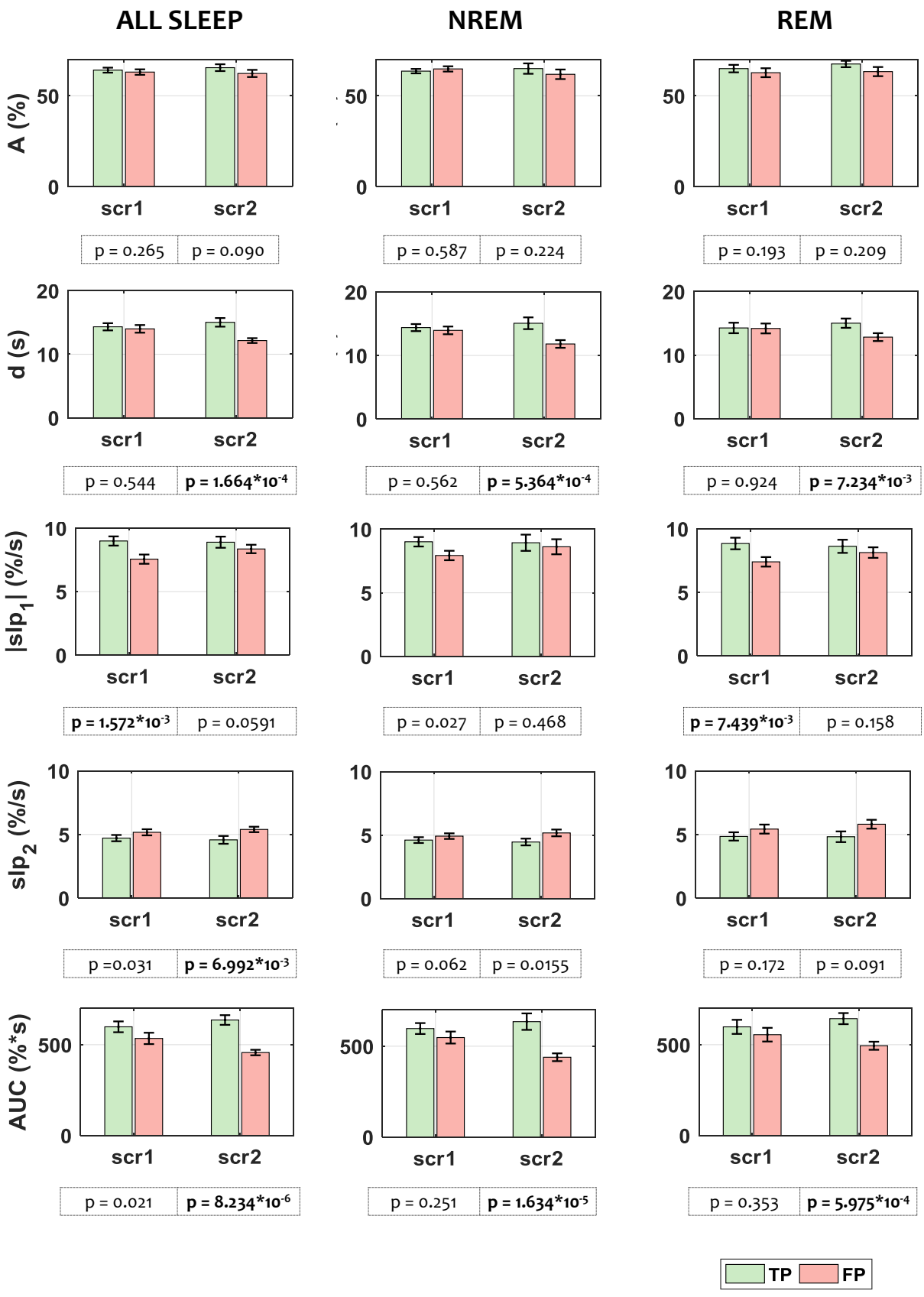


Figure 5. Properties of the PWA-drops corresponding to true positive (TP) and false positive (FP) detections (amplitude threshold T = 40%). Shown properties include amplitude (A), duration (d), descending and ascending slopes ($slp_1$and slp_2), and the area under the curve (AUC). Comparisons were performed using paired t-tests across subjects, separately for the two human scorers (scr1 and scr2, respectively). For each property, mean values across subjects (TP in green and FP in magenta)

are reported in separate bar-plots. P-values obtained for each comparison are reported in boxes below the bar-graphs. The three columns of the graph report results obtained for the whole sleep period as well as for NREM and REM episodes separately. Significant results ($p < 0.01$) are indicated with bold text.

3.2 Analysis of PWA-drops properties and distribution across sleep stages

In order to furtherhighlightpotential applications and advantages of the present automated procedure for PWA-drop detection, we evaluated the properties and distribution of PWA-drops across different sleep stages, investigating the presence of potential differences with a repeated measures (rm)ANOVA performed across subjects. As in previous analyses, PWA-drops were detected using a 40% amplitude threshold. Figure 6 showsin distinct panels the mean number of PWA-drops per hour (PDI), and the mean values of their amplitude (A), duration (d), descending slope (slp_1), ascending slope (slp_2), and area under the curve(AUC),across N1, N2, N3 and REM sleep.

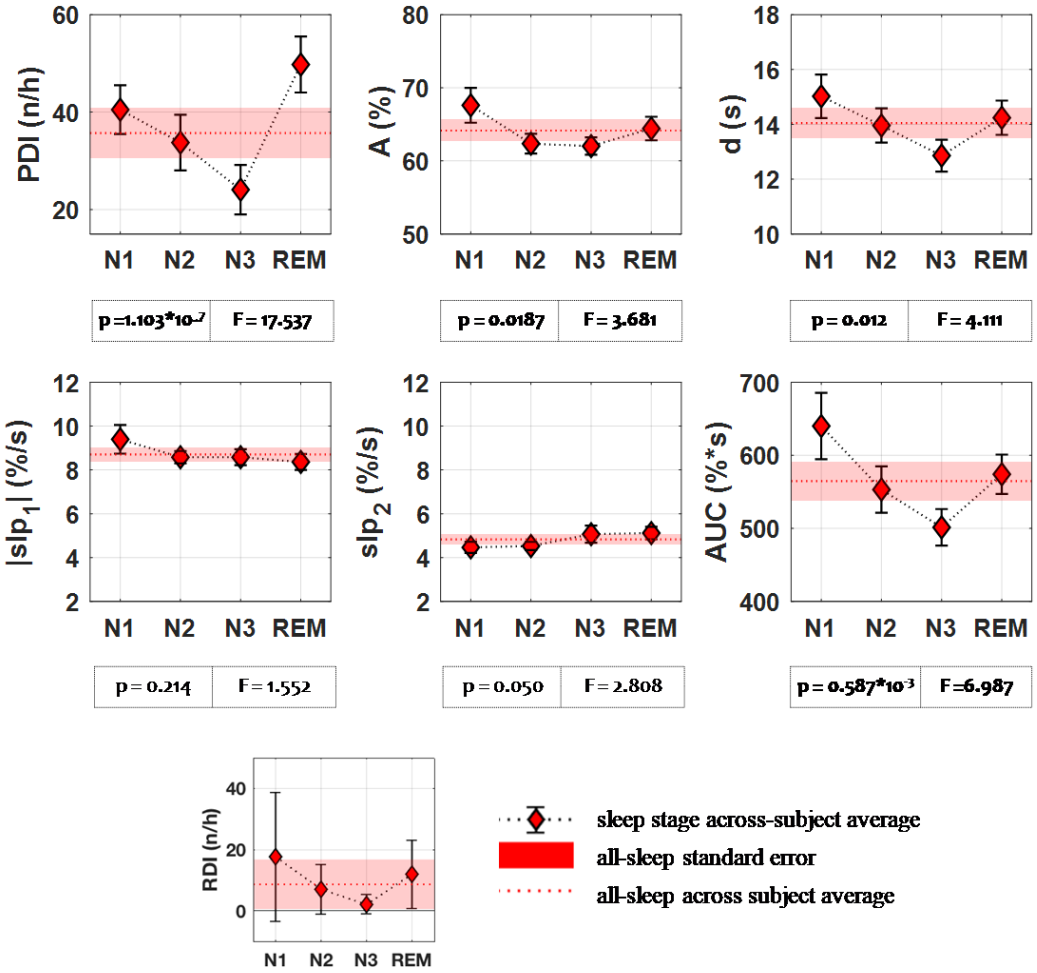


Figure 6. Distribution of PWA-dropindex and maincharacteristics (A, d, slp_1, slp_2 and AUC) across different sleep stages.Red diamonds indicate mean values (across subjects) for N1, N2, N3 and REM episodes.The red shaded areaindicatescorresponding mean values ±SD computed for the whole sleep period. In order to evaluate the effect of the sleep stage on PWA-drop properties anrmANOVA was performed (N=16, STAGE as within subject factor). Correspondent p and F(3,45) values are reported in the boxes below each graph. Significant results ($p < 0.05$) are indicated with bold text. Given the known association between PWA-drops and respiratory events, the bottom panel shows the respiratory disturbance index (RDI) computed for each sleep stage.

The PDI was found to differ significantly across sleep stages (rmANOVA; $p=10^{-7}$, $F_{(3,45)}=17.54$), with the lowest value (24.1 ± 5.1 n/h) in N3 sleep and the highest in REM sleep (49.7 ± 5.7n/h). Significant, but less robust effects, were also observed for amplitude (p=0.02, $F_{(3,45)}=3.68$), duration (p=0.01, $F_{(3,45)}=4.11$) and AUC p=0.0006, $F_{(3,45)}=6.99$). In these cases, the lowest values were again observed for N3, while the highest were found for N1. No significant effects of stage were found for descending (p>0.2) and ascending (p>0.05) slopes, although a statistical trend was observed for this latter parameter. In fact, the ascending slope tended to be lowest in N1 and highest in REM-sleep.

4. DISCUSSION

A growing body of evidence indicates that drops in pulse wave amplitude (PWA)resulting from autonomic vasoconstriction, may represent a sensitive marker of autonomic activations, whose

alterations are in turn associated with severalpathologicalconditions. While PWA-drops can be easily measured non-obtrusivelythrough photo-plethysmography (PPG),their actual detection still largely relies upon human operators, who visually inspect the whole-night PPG-trace and manually mark each event. Importantly, this approach strongly limits reproducibility, due to intra- and inter-scorer variability, and prevents the possibility to investigate the potential predictive role of PWA-related indices in large cohorts of patients.Of note, most of the automated detection tools available in commercial software suites are not scientifically validated, nor freely available, and only allow the computation of simple, basic parameters, such as the PWA-drop index. To overcome these issues, here we developed and validated an automated approach for the detection and characterization of PWA-drops in whole night PPG-signals. We demonstrated that the algorithm reliably detects PWA-drops with an accuracy that appears to be even higher than the onesof expert human operators. Finally, we showed that the algorithm may allow to easily investigate not only the number of events per hour, but also several other parameters with a potential physiological and clinical relevance, including timing, amplitude, duration and slopes. These properties can be used to study how the PWA-drops change as a function of different sleep stages, or in relation to their association with other physiological or pathological events (e.g., cortical arousals, motor or respiratory events). This kind of investigation could gain an important role in clarifying the pathogenesis of many cardiovascular and metabolic diseases, as well as in the definition of tools for the diagnostic and prognostic evaluation of patients with sleep disorders.

Performance of the PWA-drop detection procedure

In order to evaluate the performance of the PWA-drop detection procedure, a comparison was performed with the gold standard represented by the manual scoring of two independent human experts. In line with previous work(Adler et al. 2013; Haba-Rubio et al. 2005), the human scorers were instructed to mark events characterized by a minimum absolute PWA-signal reduction corresponding to the 30%. Of note, however, there is currently no consensus, nor a well-defined pathophysiological knowledge, to guide the selection of thisimportant parameter, that has been in turn indicated as potentially informative about the intensity of the underlying sympathetic activation. Moreover, during visual scoring, the quantification of the actual signal percent variation may be subject to a great within- and between-scorer variability. For these reasons, we incorporated in theautomated detection procedure the possibility for the user to modulate the amplitude threshold and we performed the present validation procedure by testing different values ofthis parameter, between 10% and 80% (with 10% steps).

As expected, sensitivity and precision showed opposite trendsas a function of the amplitude threshold, irrespective of the sleep stage or the human scorer. In particular, the sensitivity of the algorithm was found to reach a maximum in correspondence of the minimum amplitude threshold (T = 10%), when the algorithm detected the greatest number of drops. Vice-versa, values of precision increased with higher amplitude thresholds, while the number of algorithm detections decreased. Among the tested amplitude thresholds, the one corresponding to 40% produced the best compromise between sensitivity (~80%) and precision (~70%) and led to an overall number of detections consistent with those provided by the two human scorers. The detections obtained by the algorithm with a 40% threshold are thus more similar to those obtained based on visual inspection and a (visually estimated) 30% amplitude threshold. This observation suggests that the operators may actually tend to miss PWA-drops for which the amplitude value is higher but relatively close to the selected threshold, probably because of a low precision in visually estimating the drop's relative amplitude. Of note, this effect could be also expected to result in relative variations across scorers, with a negative impact on the reproducibility of results. In line with this, we found that the agreement between the two scorers (expressed in terms of F-score) only reached 70%, a value that

was lower than those expressing the agreement between the algorithm and each of the two scorers (~75%). This result impliesthe existence of a greater variability between the two human scorers, compared to the one between each human scorer and the automated algorithm and points to a greater reliability and reproducibility of the latter.

Importantly, while the sensitivity of the algorithm reached values above 80%, the precision remained relatively low (~70%), indicating the possibility of false positive detections by the algorithm. However, a visual inspection of PWA-drops detected by the algorithm but not by the scorers (apparent false positives), confirmed that these were actual changes in the amplitude of the PWA-signal, very similar in shape to true positive drops, even if sometimes less pronounced and with their amplitude close (but not below) the selected threshold (see examples in Figure 4). Moreover, in many cases, even when a PWA-drop detected by the algorithm was not recognized by a scorer, it was instead identified by the other scorer. A more detailedanalysis of the differences between true positive (i.e., drops marked by both the algorithm and the scorers) and false positive PWA-drops confirmed that such events tended to have similar amplitudes, but different slopesand areas under the curve (AUC).However, these differences were not consistent across the two scorers and could then be driven by subjective criteria.Globally, our results suggest that the human scorers tended to miss relatively small and shallow events, even though they actually passed the selected (30%) amplitude threshold. This happened especially when several PWA-drops occurred in sequence, thus reflecting difficulties in precisely and reproducibly estimatingupon visual inspectionthe maximum amplitude of each dropin relation to preceding baseline values. The same issue related to the withinand betweenscorer variability in amplitude estimation may also contribute to explain the apparent ‘false negative’ detections, corresponding to cases in which the algorithm failed to detect a drop marked by the human scorer(Figure 4).

Overall, while the inclusion of only two human scorers is not sufficient to draw clear conclusions regarding the reproducibility of a human-based evaluation of PWA-drops, our results suggest that this process may be affected by a non-negligible inter-scorer variability. Such variability may alter the reproducibility of results in individual studies, thus limiting the possibility to compare findings obtained in different cohorts and by different research groups. This issue is made even worse by the lack of a consensus regarding the criteria –and in particular the amplitude threshold- that should be applied for the correct identification of clinically relevant PWA-drops.Our results suggest that these limitations could be overcome by using automated approaches based on standardized criteria.

Potential applications in physiological and pathological states

A clear advantage of the automated detection algorithm over manual scoring lies in the possibility to rapidly and easily measure several properties of each PWA-drop, even in large samples. In addition to the timing of each event, which could be used to evaluate the number of drops per hour (PDI) across the whole sleep period or within particular stages, or the association with other specific events (e.g., apnea or hypopnea events), other measured parameters include amplitude, duration, descending slope, ascending slope and area under the curve. Here we showed that these properties of the PWA-drops tend to differ significantly across sleep stages. In particular, the number, amplitude, duration and area under the curve of PWA-drops share a common trend, decreasing from N1 to the deepest stage of NREM sleep (N3). Most properties of PWA-drops occurring during REM-sleep were similar to those observed in N2, with the notable exception of the number of events per hour, which was found to be higher in REM than in any other sleep stage. This observation is consistent with the known intrinsic variability of most autonomic signalsduring REM sleep.While, at present, the possible clinical value of described morphological properties of the PWA-dropsis unknown, future studies in clinical populationswill allow to investigate theirpossible alterations and predictive rolein relation to distinct physiological and pathological processes.

5. CONCLUSION

Evidence indicates that theanalysis of drops inpulse wave amplitudeduring sleep may offer a very promisingmarker ofchanges in autonomic activation, which are potentially associated withsleep fragmentation and disruption in pathological states(Delessert et al. 2010; Grote et al. 2011; Karmakar et al. 2014; Sommermeyer et al. 2014, 2016). Here we developed an automated algorithm for the detection and characterization of PWA-drops that aims at overcoming most of the limitations associated with user-dependent analyses, including long processing time and high intra- and inter-scorer variability. The proposed algorithm is supposed to work in a completely user-independent fashion, starting from all-night PSG-data, and automatically dealing with the potential presence of movement artifacts and/or the possibility of sensor loss/displacement throughout the night, and self-adapting to stage-dependent variability in the PWA-signal.In light of these properties, it may allow to analyze large databases in a relatively short time and provides the opportunity to efficiently evaluate the potential impact of different PWA-drop amplitude thresholds on other parameters of interest. The algorithm, written in MATLAB(The Mathworks Inc 2009), is freely available for downloadin the OSF repository (https://osf.io/c2eup/?view_only=a2890a0f06704cf1a281eee5727d8790).We hope that the use of automated detection approaches by future studies will lead to a better understanding ofphysiological changestriggered by autonomic activations during sleep andof the possible value of PWA-derived parameters in predicting the risk of cardiovascular and metabolic diseases associated with sleep disorders.

BIBLIOGRAPHY

Adler, Dan et al. 2013. "Pulse Wave Amplitude Reduction: A Surrogate Marker of Micro-Arousals Associated with Respiratory Events Occurring under Non-Invasive Ventilation?" *Respiratory Medicine* 107(12):2053–60. Retrieved December 17, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/24169072).

Allen, John. 2007. "Photoplethysmography and Its Application in Clinical Physiological Measurement." *Physiological Measurement* 28(3):R1–39. Retrieved January 6, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/17322588).

Bartels, Wibke, Dana Buck, Martin Glos, Ingo Fietze, and Thomas Penzel. 2016. "Definition and Importance of Autonomic Arousal in Patients with Sleep Disordered Breathing." *Sleep Medicine Clinics* 11(4):435–44. Retrieved April 8, 2019 (https://linkinghub.elsevier.com/retrieve/pii/S1556407X16300728).

Biaggioni, Italo and David A. Calhoun. 2016. "Sympathetic Activity, Hypertension, and The Importance of a Good Night's Sleep." *Hypertension* 68(6):1338–39. Retrieved May 22, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/27698060).

Bosi, Marcello et al. 2018. "Arousal Responses to Respiratory Events during Sleep: The Role of Pulse Wave Amplitude." *Journal of Sleep Research* 27(2):261–69. Retrieved December 17, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/28901049).

British Editorial Society of Bone and Joint Surgery., F. A. R. 1954. *Journal of Bone and Joint Surgery. British Volume.* [Published for the British Editorial Society of Bone & Joint Surgery by Churchill Livingstone]. Retrieved January 3, 2018 (http://bjj.boneandjoint.org.uk/content/36-B/4/706).

Brook, R. D. and S. Julius. 2000. "Autonomic Imbalance, Hypertension, and Cardiovascular Risk." *American journal of hypertension* 13(6 Pt 2):112S–122S. Retrieved January 2, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/10921530).

Carlson, Jan T. et al. 1993. "Augmented Resting Sympathetic Activity in Awake Patients With Obstructive Sleep Apnea." *Chest* 103(6):1763–68. Retrieved February 6, 2018 (http://linkinghub.elsevier.com/retrieve/pii/S001236921542077X).

Catcheside, Peter G., Siau Chien Chiong, Jeremy Mercer, Nicholas A. Saunders, and R. Douglas McEvoy. 2002. "Noninvasive Cardiovascular Markers of Acoustically Induced Arousal from Non-Rapid-Eye-Movement Sleep." *Sleep* 25(7):797–804. Retrieved January 8, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/12405616).

Delessert, Alexandre et al. 2010. "Pulse Wave Amplitude Drops during Sleep Are Reliable Surrogate Markers of Changes in Cortical Activity." *Sleep* 33(12):1687–92. Retrieved April 4, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/21120131).

Dresler, Martin et al. 2012. "Neural Correlates of Dream Lucidity Obtained from Contrasting Lucid versus Non-Lucid REM Sleep: A Combined EEG/FMRI Case Study." *Sleep* 35(7):1017–20. Retrieved June 15, 2016 (http://www.ncbi.nlm.nih.gov/pubmed/22754049).

Esler, M. 2000. "The Sympathetic System and Hypertension." *American journal of hypertension* 13(6 Pt 2):99S–105S. Retrieved January 2, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/10921528).

Fletcher, Eugene C. 2003. "Sympathetic over Activity in the Etiology of Hypertension of Obstructive Sleep Apnea." *Sleep* 26(1):15–19. Retrieved May 22, 2018

(http://www.ncbi.nlm.nih.gov/pubmed/12627727).

Grote, Ludger, Dirk Sommermeyer, Ding Zou, Derek N. Eder, and Jan Hedner. 2011. "Oximeter-Based Autonomic State Indicator Algorithm for Cardiovascular Risk Assessment." *Chest* 139(2):253–59. Retrieved January 6, 2018 (https://www.sciencedirect.com/science/article/pii/S0012369211600639).

Grote, Ludger, Ding Zou, Holger Kraiczi, and Jan Hedner. 2003a. "Finger Plethysmography--a Method for Monitoring Finger Blood Flow during Sleep Disordered Breathing." *Respiratory physiology & neurobiology* 136(2–3):141–52. Retrieved December 30, 2017 (http://www.ncbi.nlm.nih.gov/pubmed/12853006).

Haba-Rubio, José et al. 2005. "Obstructive Sleep Apnea Syndrome: Effect of Respiratory Events and Arousal on Pulse Wave Amplitude Measured by Photoplethysmography in NREM Sleep." *Sleep and Breathing* 9(2):73–81. Retrieved April 4, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/15875228).

Hedner, J., H. Ejnell, J. Sellgren, T. Hedner, and G. Wallin. 1988. "Is High and Fluctuating Muscle Nerve Sympathetic Activity in the Sleep Apnoea Syndrome of Pathogenetic Importance for the Development of Hypertension?" *Journal of hypertension. Supplement : official journal of the International Society of Hypertension* 6(4):S529-31. Retrieved June 26, 2019 (http://www.ncbi.nlm.nih.gov/pubmed/3241251).

Heinzer, R. et al. 2015. "Prevalence of Sleep-Disordered Breathing in the General Population: THE HypnoLaus Study." *The Lancet Respiratory Medicine*.

Janackova, Sona and Emilia Sforza. 2008. "Neurobiology of Sleep Fragmentation: Cortical and Autonomic Markers of Sleep Disorders." *Current Pharmaceutical Design* 14(32):3474–80. Retrieved February 6, 2018 (http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1381-6128&volume=14&issue=32&spage=3474).

Johnson, Laverne C. and Ardie Lubin. 1967. "The Orienting Reflex during Waking and Sleeping." *Electroencephalography and Clinical Neurophysiology* 22(1):11–21. Retrieved January 7, 2018 (http://www.sciencedirect.com/science/article/pii/0013469467900041).

Karmakar, Chandan, Ahsan Khandoker, Thomas Penzel, Christoph Schobel, and Marimuthu Palaniswami. 2014. "Detection of Respiratory Arousals Using Photoplethysmography (PPG) Signal in Sleep Apnea Patients." *IEEE Journal of Biomedical and Health Informatics* 18(3):1065–73. Retrieved January 7, 2018 (http://ieeexplore.ieee.org/document/6606860/).

Korpas, D., J. Hálek, and L. Dolezal. 2009. "Parameters Describing the Pulse Wave." *Physiological research* 58(4):473–79. Retrieved December 17, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/18656997).

Lévy, Patrick and Jean-Louis Pépin. 2003. "Sleep Fragmentation: Clinical Usefulness of Autonomic Markers." *Sleep Medicine* 4(6):489–91. Retrieved February 22, 2019 (https://www.sciencedirect.com/science/article/pii/S1389945703001977?via%3Dihub).

Mark, A. L. 1996. "The Sympathetic Nervous System in Hypertension: A Potential Long-Term Regulator of Arterial Pressure." *Journal of hypertension. Supplement : official journal of the International Society of Hypertension* 14(5):S159-65. Retrieved January 2, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/9120673).

Martin, S. E., P. K. Wraith, I. J. Deary, and N. J. Douglas. 1997. "The Effect of Nonvisible Sleep Fragmentation on Daytime Function." *American Journal of Respiratory and Critical Care Medicine* 155(5):1596–1601. Retrieved January 8, 2018

(http://www.ncbi.nlm.nih.gov/pubmed/9154863).

Nitzan, M., A. Babchenko, B. Khanokh, and D. Landau. 1998. "The Variability of the Photoplethysmographic Signal--a Potential Method for the Evaluation of the Autonomic Nervous System." *Physiological measurement* 19(1):93–102. Retrieved January 8, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/9522390).

Nitzan, M., A. Babchenko, D. Shemesh, and J. Alberton. 2001. "Influence of Thoracic Sympathectomy on Cardiac Induced Oscillations in Tissue Blood Volume." *Medical & Biological Engineering & Computing* 39(5):579–83. Retrieved January 6, 2018 (http://link.springer.com/10.1007/BF02345149).

Nitzan, Meir, Sergei Turivnenko, Adina Milston, Anatoly Babchenko, and Y. Mahler. 1996. "Low-Frequency Variability in the Blood Volume and in the Blood Volume Pulse Measured by Photoplethysmography." *Journal of Biomedical Optics* 1(2):223. Retrieved January 6, 2018 (http://biomedicaloptics.spiedigitallibrary.org/article.aspx?doi=10.1117/12.231366).

Shimizu, Tetsuo et al. 1992. "Muscle Nerve Sympathetic Activity during Sleep and Its Change with Arousal Response." *Journal of Sleep Research* 1(3):178–85. Retrieved January 8, 2018 (http://doi.wiley.com/10.1111/j.1365-2869.1992.tb00035.x).

Silber, Michael H. et al. 2007. "The Visual Scoring of Sleep in Adults." *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 3(2):121–31. Retrieved April 18, 2016 (http://www.ncbi.nlm.nih.gov/pubmed/17557422).

Sinski, M., J. Lewandowski, P. Abramczyk, K. Narkiewicz, and Z. Gaciong. 2006. "Why Study Sympathetic Nervous System?" *Journal of physiology and pharmacology : an official journal of the Polish Physiological Society* 57 Suppl 11:79–92. Retrieved June 26, 2019 (http://www.ncbi.nlm.nih.gov/pubmed/17244940).

Somers, V. K., M. E. Dyken, M. P. Clary, and F. M. Abboud. 1995. "Sympathetic Neural Mechanisms in Obstructive Sleep Apnea." *The Journal of clinical investigation* 96(4):1897–1904. Retrieved February 6, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/7560081).

Somers, Virend K., Mark E. Dyken, Allyn L. Mark, and Francois M. Abboud. 1993. "Sympathetic-Nerve Activity during Sleep in Normal Subjects." *New England Journal of Medicine* 328(5):303–7. Retrieved January 8, 2018 (http://www.nejm.org/doi/abs/10.1056/NEJM199302043280502).

Sommermeyer, Dirk et al. 2014. "The Use of Overnight Pulse Wave Analysis for Recognition of Cardiovascular Risk Factors and Risk." *Journal of Hypertension* 32(2):276–85. Retrieved January 6, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/24248087).

Sommermeyer, Dirk et al. 2016. "Detection of Cardiovascular Risk from a Photoplethysmographic Signal Using a Matching Pursuit Algorithm." *Medical & Biological Engineering & Computing* 54(7):1111–21. Retrieved December 17, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/26538425).

The Mathworks Inc. 2009. "Matlab."

Thompson, R. 1985. "A Note on Restricted Maximum Likelihood Estimation with an Alternative Outlier Model." *Journal of the Royal Statistical Society. Series B (…*. Retrieved May 2, 2016 (http://www.jstor.org/stable/2345543).

Thorp, Alicia A. and Markus P. Schlaich. 2015. "Relevance of Sympathetic Nervous System Activation in Obesity and Metabolic Syndrome." *Journal of diabetes research* 2015:341583. Retrieved January 2, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/26064978).

730 Trinder, J. et al. 2001. “Autonomic Activity during Human Sleep as a Function of Time and Sleep
731 Stage.” *Journal of sleep research* 10(4):253–64. Retrieved January 8, 2018
732 (http://www.ncbi.nlm.nih.gov/pubmed/11903855).

733 Trinder, John, Joanna Waloszek, Michael J. Woods, and Amy S. Jordan. 2012. “Sleep and
734 Cardiovascular Regulation.” *Pflügers Archiv - European Journal of Physiology* 463(1):161–
735 68. Retrieved April 8, 2019 (http://www.ncbi.nlm.nih.gov/pubmed/22038322).

736 Tsioufis, Costas et al. 2011. “Pathophysiology of Resistant Hypertension: The Role of Sympathetic
737 Nervous System.” *International journal of hypertension* 2011:642416. Retrieved January 2,
738 2018 (http://www.ncbi.nlm.nih.gov/pubmed/21331155).

739 Vargas-Pérez, Noel, Kanika Bagai, and Arthur Walters. 2017. “Cardiovascular Comorbidity in
740 Patients with Restless Legs Syndrome: Current Perspectives.” *Journal of Parkinsonism and*
741 *Restless Legs Syndrome*.

742 Vinik, A. I., R. E. Maser, and D. Ziegler. 2011. “Autonomic Imbalance: Prophet of Doom or Scope
743 for Hope?” *Diabetic medicine : a journal of the British Diabetic Association* 28(6):643–51.
744 Retrieved January 3, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/21569084).

745 Whitehurst, Lauren N., Nicola Cellini, Elizabeth A. McDevitt, Katherine A. Duggan, and Sara C.
746 Mednick. 2016. “Autonomic Activity during Sleep Predicts Memory Consolidation in
747 Humans.” *Proceedings of the National Academy of Sciences of the United States of America*
748 113(26):7272–77. Retrieved May 22, 2018 (http://www.ncbi.nlm.nih.gov/pubmed/27298366).

749 Yetton, Benjamin D. et al. 2016. “Automatic Detection of Rapid Eye Movements (REMs): A
750 Machine Learning Approach.” *Journal of Neuroscience Methods*.

751 de Zambotti, Massimiliano, John Trinder, Alessandro Silvani, Ian M. Colrain, and Fiona C. Baker.
752 2018. “Dynamic Coupling between the Central and Autonomic Nervous Systems during Sleep:
753 A Review.” *Neuroscience & Biobehavioral Reviews* 90:84–103. Retrieved April 8, 2019
754 (https://linkinghub.elsevier.com/retrieve/pii/S0149763417306577).

755

Author contributions (CRediT taxonomy)

Conceptualization, M.B., J.H.R., R.H.,G.B., F.S..;
Methodology, M.B.,G.H.;
Formal Analysis, M.B.;
Visualization, M.B.;
Resources, E.R., R.H.;
Software, M. B.;
Supervision, E.R., F.S., R.H., G.B.;
Writing – Original Draft, M.B., G.B.;
Writing – Review & Editing, M.B., G.H., P.P., E.R., J.H.R., F.S., R.H., G.B.