SCIENTIFIC INVESTIGATIONS

Sleep-wake parameters can be detected in patients with chronic stroke using a multisensor accelerometer: a validation study

Elie Gottlieb, BA[1,2]; Leonid Churilov, PhD[2]; Emilio Werden, PhD[1,2]; Thomas Churchward, BAppSc, RPSGT, RST[3,4]; Matthew P. Pase, PhD[5,6]; Natalia Egorova, PhD[1,7]; Mark E. Howard, MBBS, PhD[2,3,4,*]; Amy Brodtmann, MBBS, PhD[1,2,*]

[1]The Florey Institute of Neuroscience and Mental Health, Melbourne, Victoria, Australia; [2]University of Melbourne, Melbourne, Victoria, Australia; [3]Institute for Breathing and Sleep, Melbourne, Victoria, Australia; [4]Austin Health, Heidelberg, Victoria, Australia; [5]Turner Institute for Brain and Mental Health, School of Psychological Sciences, Monash University, Victoria, Australia; [6]Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts; [7]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Victoria, Australia; *Co-senior authors

Study Objectives: Sleep-wake dysfunction is bidirectionally associated with the pathogenesis and evolution of stroke. Longitudinal and prospective measurement of sleep after chronic stroke remains poorly characterized because of a lack of validated objective and ambulatory sleep measurement tools in neurological populations. This study aimed to validate a multisensor sleep monitor, the SenseWear Armband (SWA), in patients with ischemic stroke and control patients using at-home polysomnography.

Methods: Twenty-eight radiologically confirmed patients with ischemic stroke (aged 69.61 ± 7.35 years; mean = 4.1 years poststroke) and 16 control patients (aged 73.75 ± 7.10 years) underwent overnight at-home polysomnography in tandem with the SWA. Lin's concordance correlation coefficient and reduced major axis regressions were employed to assess concordance of SWA vs polysomnography-measured total sleep time, sleep efficiency, sleep onset latency, and wake after sleep onset. Subsequently, data were converted to 30-second epochs to match at-home polysomnography. Epoch-by-epoch agreement between SWA and at-home polysomnography was estimated using crude agreement, Cohen's kappa, sensitivity, and specificity.

Results: Total sleep time was the most robustly quantified sleep-wake variable (concordance correlation coefficient = 0.49). The SWA performed poorest for sleep measures requiring discrimination of wakefulness (sleep onset latency; concordance correlation coefficient = 0.16). The sensitivity of the SWA was high (95.90%) for patients with stroke and for control patients (95.70%). The specificity of the SWA was fair-moderate for patients with stroke (40.45%) and moderate for control patients (45.60%). Epoch-by-epoch agreement rate was fair (78%) in patients with stroke and fair (74%) in controls.

Conclusions: The SWA shows promise as an ambulatory tool to estimate macro parameters of sleep-wake; however, agreement at an epoch level is only moderate-fair. Use of the SWA warrants caution when it is used as a diagnostic tool or in populations with significant sleep-wake fragmentation.

Keywords: validation, accelerometer, stroke, aging, behavioral sleep medicine, scoring, instrumentation, sleep/wake physiology

Citation: Gottlieb E, Churilov L, Werden E, et al. Sleep-wake parameters can be detected in patients with chronic stroke using a multisensory accelerometer: a validation study. *J Clin Sleep Med.* 2021:17(2):167–175.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Longitudinal and ambulatory measurement of sleep in patients with stroke is vital to further establish the bidirectional impact of sleep-wake dysfunction in stroke. It remains unclear whether noninvasive and multisensor ambulatory sleep-wake monitors can accurately and reliably detect sleep-wake parameters in stroke populations.

Study Impact: The SenseWear Armband shows promise as an ambulatory sleep-wake measurement tool. However, caution is warranted when using the SenseWear Armband as a diagnostic tool or in clinical populations with significant sleep-wake fragmentation.

INTRODUCTION

Stroke is a leading cause of long-term disability and death worldwide.[1] Sleep-wake dysfunction is bidirectionally associated with the pathogenesis and evolution of stroke and may be a key modifiable risk factor for stroke.[2,3] Excessively long sleep duration characterized by > 8 hours of sleep is an independent risk factor and consequence of ischemic stroke.[4,5] Inversely, sleep efficiency is significantly reduced in patients with stroke and negatively correlated with poststroke recovery.[6] Despite the rise in interest in sleep-wake dysfunction as a modifiable risk factor for stroke, the tools used to measure sleep-wake after

stroke are heterogeneous.[2] Poor validation methodologies of existing wearable sleep technology have limited their use in larger prospective studies conducted in ambulatory settings.[7]

Polysomnography (PSG), despite being recognized as the gold standard for sleep-wake detection, is intrusive and is generally conducted for only 1–5 nights.[8] The use of PSG is therefore unsuitable for prospective cohort studies requiring longitudinal sleep-wake measurement. Sleep diaries and validated sleep-wake questionnaires are self-reported and potentially unreliable in populations with neurological disease, where sleep-state misperception may be a feature of sleep-wake pathology.[9] The development and validation of ambulatory,

objective, and noninvasive sleep-wake monitors is warranted. Recent guidelines have established a set of best practices for validation studies.[8] Whether ambulatory sleep-wake monitors accurately and reliably detect sleep-wake states in stroke populations is unclear, particularly given that patients with stroke often display abnormal sleep architecture and a higher prevalence of sleep disorders.[6]

Traditional sleep-wake wearables solely utilize accelerometry to estimate sleep via micromovement detection. Recent advances in wearable technologies have enabled the use of multisensor arrays to robustly quantify sleep-wake parameters. For example, the BodyMedia SenseWear Armband (SWA; BodyMedia Inc, Pittsburgh, PA) utilizes a dual-axis accelerometer, galvanic skin response sensors, heat flux sensors, skin temperature sensors, and a near-body ambient temperature sensor to provide a more comprehensive breadth of physiological observations to enhance the accuracy of sleep estimation.[10] Previous validation of the SWA vs PSG has been conducted in healthy participants,[11,12] adolescents,[13] and patients with OSA,[14] but not in clinical populations with neurological disorders. The present validation study was designed to assess the clinical utility of the SWA as a sleep monitor vs ambulatory at-home polysomnography (home-PSG) in participants with stroke and healthy control patients by addressing the agreement between the SWA and home-PSG at both the patient and epoch-by-epoch level.

METHODS

Participants

Participants were recruited from the Cognition and Neocortical Volume After Stroke (CANVAS) study cohort. The study includes participants with first-ever or recurrent ischemic stroke within any circulation[15] or etiology[16] and no history of dementia or another neurodegenerative condition. The protocol and inclusion criteria for control participants were identical to those for participants with stroke, except that control participants did not have a history of stroke. Patients were recruited from 3 hospitals in Melbourne, Australia: Austin Health, Eastern Health, and Melbourne Health. The study was approved by each hospital's human research ethics committee in line with the Declaration of Helsinki.[17] Patients with untreated or previously diagnosed OSA were excluded. Patients with narcolepsy and those taking medications with a primary effect on sleep architecture, as measured by electroencephalogram (EEG) during the time of the study, were excluded. Ischemic stroke was a clinical diagnosis and radiologically confirmed by clinical computed tomography or magnetic resonance imaging during acute hospitalization for the event. Patients were excluded if they were unable to undergo a 3T magnetic resonance imaging scan, were unable to provide informed consent because of severe aphasia, were diagnosed with transient ischemic attack or failed to have their stroke confirmed by computed tomography or magnetic resonance imaging, or had significant medical comorbidities precluding participation in cognitive-behavioral testing. Participants were recruited, on average, 4.10 years poststroke (standard deviation = 0.91).

Sleep-wake protocol

The protocol for the present study was split into 2 timepoints. During timepoint 1, participants wore the SWA and completed a 16-item sleep diary for at least 1 week. Timepoint 1 sleep data provided an "average" sleep-day used to schedule participants' home-PSG examination. During timepoint 2, participants were fitted with an ambulatory home-PSG device (Somté PSG version 2.0; Compumedics Limited, Abbotsford, Victoria, Australia) and wore the SWA in tandem for 1 night at home. Upon returning from timepoint 2, participants completed an array of demographic, mood, and sleep-circadian questionnaires.

The SWA was attached to participants' upper right arm according to manufacturer recommendations. The SWA utilizes a combination of accelerometry, near-body ambient skin temperature, heat flux, and galvanic skin response to measure sleep-wake. The accelerometer includes a microelectromechanical sensor, which has a scale of ±2 g and a sensitivity of 167 mV/g. With respect to sleep duration, raw data output is binary (ie, 1 = sleeping, 0 = awake). Sleep-wake averages from timepoint 2 were extracted using BodyMedia software (InnerView; BodyMedia, Pittsburgh, PA) and used as primary validation outcome variables.

Home-PSG was conducted by a trained sleep scientist in accordance with *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications,* version 2.6[18] and the 2020 Australia and New Zealand Sleep Science Association (ANZSSA) and Australasian Sleep Association (ASA) guidelines.[19] Profusion PSG Sleep Software (V4.5 Build 468; Compumedics Limited, Abbotsford, Victoria, Australia) was used to analyze sleep. Six-lead EEG placement was used in accordance to the international 10–20 system (F4-M1, C4-M1, O2-M1; F3-M2, C3-M2, O1-M2). For patients with stroke, ipsilesional and contralesional hemisphere EEG were scored independently because of the potential effects of infarction and regional brain volume loss on unihemispheric sleep.[5,20]

Sleep and respiratory scoring were completed by an experienced research-grade sleep scientist. The scientist remained blinded to the healthy vs lesioned hemispheres until both hemispheric sleep-wake analyses were completed. Upon completion of the independent bihemispheric sleep-wake analyses, the scientist was unblinded to the lesioned hemisphere and utilized the contralesional hemisphere's sleep-wake states and arousal events for scoring respiratory events that were used in subsequent calculations, such as the AHI. Thus, sleep-wake variables were scored (blindly) in the ipsilesional and contralesional hemispheres by the sleep scientist to allow for comparisons based on lesion location, whereas respiratory parameters were scored contralesionally to allow for comparison between healthy control patients and participants with stroke, in which case the health-state of the hemisphere used could be a confounder for respiratory events.[21,22] In addition to EEG, the following were recorded via home-PSG: electrooculogram (left and right placement), electromyogram (mentalis/submentalis), electrocardiogram (modified lead II), pulse oximetry (oxygen saturation), respiration (oronasal airflow using a nasal cannula), respiratory inductance plethysmography (chest and abdominal wall movement), and leg electromyogram.

Outcome measures

Patient macrolevel outcomes of sleep-wake included total sleep time, sleep onset latency, wake after sleep onset, and sleep

efficiency (the ratio of total sleep time to time in bed). Total sleep time and sleep efficiency were automatically computed using the SWA proprietary software (BodyMedia InnerView Research Software). Sleep onset latency was manually calculated using patient-measured "lights-out" time assessed from the sleep diary time, which was confirmed with the SWA's detection of the first "lay" time. Wake after sleep onset was manually calculated as periods of wakefulness after the first defined sleep onset.

Epoch-by-epoch microlevel outcomes were measured as follows. Epoch-by-epoch analyses and harmonization of SWA and home-PSG data were conducted in accordance with previously published SWA validation methodologies.[12,23,24] The binary (1 = sleeping, 0 = awake) SWA sleep-wake estimations were computed in 1-minute epochs across a single main nighttime sleep period. Home-PSG sleep-wake data were analyzed in 30-second segments (epochs) and were used as the reference comparator for the SWA. PSG sleep epochs (NREM 1–3 and REM) and wake epochs were converted to binary (sleep = 1, wake = 0) to mirror the SWA output binary data. Because the SWA is limited to 1-minute epoch sampling, each 1-minute SWA epoch was divided into two 30-second epochs to mirror PSG epochs, as previously reported.[12,23,24] For example, a 1-minute SWA sleep epoch was converted to two 30-second sleep epochs to match home-PSG epochs. SWA and home-PSG data exports were synchronized on the same network of computers to ensure that time zone outputs were matched between measures. Secondary manual checks of time synchronization for each epoch were conducted when binary SWA and home-PSG data were aligned.

Statistical analysis

Statistical analyses were conducted in SPSS Version 26 (IBM Corp, Armonk, NY) and Stata 15IC (Stata Corp, College Station, TX). Reduced major axis regressions and concordance statistics (ie, Bland-Altman, Lin's concordance correlation coefficients [CCC]) were conducted in Stata. Epoch-by-epoch agreements, Cohen's kappa coefficients, and demographic and clinical comparisons were computed in SPSS. All analyses were 2-tailed, and a critical P value of .05 was used. Continuous variables were summarized by median and interquartile ranges. Categorical variables were summarized by number (n) and percent. Mann-Whitney U tests were used for comparisons between continuous measures because normality testing can be severely underpowered in small samples. Fisher exact tests were used for 2 × 2 categorical variables.

For the patient macrolevel analysis, the agreement between SWA and PSG in total sleep time, sleep onset latency, sleep efficiency, and wake after sleep onset was estimated using CCC and further investigated using Bland-Altman limits of agreement and reduced major axis regression (RMAR).[25] The RMAR yields a slope and an intercept. A slope different from 1 is indicative of the presence of proportional bias, when the magnitude of disagreement between 2 methods increases or decreases proportionally to the measured values. Under the absence of proportional bias, an intercept different from 0 is indicative of the presence of fixed bias, when the magnitude of disagreement between 2 methods is constant across the range of measured values. The results of RMAR are reported as scatterplots that include both the line of perfect concordance and the fitted RMAR line.

For the epoch-by-epoch microlevel analysis, the sensitivity and specificity of SWA to correctly identify individual epochs as "sleep" or "wake" on PSG were estimated. Epoch-by-epoch agreement between home-PSG and the SWA were estimated for individual patients using Cohen's kappa coefficient, where k = 1 shows perfect agreement and k = 0 shows agreement based on chance alone. Viera and Garrett's[26] kappa scoring interpretation for agreement between measures was used.

RESULTS

A total of 44 participants (28 patients with stroke, 16 healthy control patients) underwent simultaneous overnight home-PSG and SWA monitoring. One patient with stroke was excluded because of SWA data loss. Polysomnographically measured total sleep time, sleep efficiency, sleep onset latency, and wake after sleep onset were not statistically different between groups (ipsilesional-stroke vs control). Patients with stroke had a higher prevalence of moderate-to-severe OSA (defined as AHI > 15 events/h; 57% vs 38%; P = .09) and a higher arousal index compared with the control patients (20 vs 12.50 events/h; P = .03). Participant demographics, stroke characteristics, and sleep metrics for patients with stroke and healthy control patients are summarized in **Table 1.** Stroke characteristics are listed in **Table S1** in the supplemental material.

Patient macrolevel outcomes

In the patients with stroke with total sleep time scored both ipsilesionally and contralesionally, we found moderate-to-fair agreement (CCC = 0.49) between the SWA and home-PSG. Reduced major axis results for the concordance of all sleep metrics in patients with stroke between the SWA and home-PSG are summarized in **Table 2** and plotted in **Figure 1.** There was no evidence of proportional bias (slope = 0.99). We observed a fixed bias for total sleep time (intercept = –76.67); the SWA systematically overestimated sleep time by approximately 77 minutes at a consistent rather than a changing amount (ie, proportional bias) across magnitude. We observed poor agreement (CCC < 0.40) between the SWA and home-PSG for all other measured sleep metrics, including sleep efficiency, wake after sleep onset, and sleep onset latency. The SWA overestimated sleep efficiency by 14.78 with evidence of proportional bias (RMAR slope = 1.41; intercept = –49.58). The SWA underestimated wake after sleep onset by approximately 63 minutes with evidence of proportional bias (RMAR slope = 1.36; intercept = 40.44). The SWA underestimated sleep onset latency by 3.51 minutes, and we observed evidence of proportional bias (RMAR slope = 1.78; intercept = –2.16).

Comparisons of concordance in healthy control patients for the SWA and home-PSG are summarized in **Table 3** and plotted in **Figure S1** in the supplemental material. We observed poor agreement (CCC < 0.40) between the SWA and home-PSG for all sleep metrics in the healthy control patients, including total sleep time, sleep efficiency, wake after sleep onset, and sleep onset latency. Similar to the comparison

Table 1—Participant demographics and sleep-respiratory characteristics of all participants with stroke and healthy control patients.

	Stroke (n = 28)	Control (n = 16)	*P* Value
Demographic characteristics			
Age, y, median (IQR)	70 (63.25–77)	76 (71.50–77.50)	.07[a]
Sex, male, n (%)	22 (79)	10 (63)	.30[b]
Education, y, median (IQR)	12.50 (10.75–15.25)	16 (10.75–18)	.19[a]
BMI, median (IQR)	28.02 (26–30.53)	25 (23–26)	.87[a]
NART-FSIQ, median (IQR)	112.30 (104.30–120.80)	119.30 (111.05–124.15)	.19[a]
Family history of stroke, n (%)	9 (32)	5 (31)	> .99[b]
Family history of dementia, n (%)	6 (21)	6 (38)	.30[b]
Depression diagnosis, n (%)	3 (11)	2 (13)	> .99[b]
Hyperlipidemia diagnosis, n (%)	10 (36)	8 (50)	.53[b]
Hypertension diagnosis, n (%)	13 (46)	6 (38)	.75[b]
Ischemic heart disease, n (%)	2 (7)	2 (13)	.61[b]
Atrial fibrillation diagnosis, n (%)	3 (11)	0 (0)	.29[b]
Type 2 diabetes mellitus, n (%)	5 (18)	0 (0)	.14[b]
High alcohol intake (> 14 standard drinks per wk), n (%)	2 (7)	2 (13)	.61[b]
ApoE_e4 (≥ 1 allele), n (%)	7 (25)	1 (6)	.22[b]
PHQ-9, median (IQR)	3 (1.25–5.75)	1 (1–2)	.039[a]
GAD-7, median (IQR)	0.50 (0–3.5)	0 (0–1)	.16[a]
PSG measured sleep-respiratory characteristics			
Total sleep time, min, median (IQR)	319.50 (280.50–375.75)	319 (278.37–341.75)	.59[a]
Sleep efficiency, ratio, median (IQR)	70.50 (60.75–78.62)	72.75 (57.37–78.00)	.73[a]
Sleep onset latency, min, median (IQR)	8.00 (3.62–16.37)	14.25 (9.37–18.00)	.08[a]
Wake after sleep onset, min, median (IQR)	119.50 (94.25–168.00)	110.50 (90.12–189.87)	.75[a]
AHI, events/h, median (IQR)	25.00 (9.25–33.25)	10.00 (7.0–21.25)	.09[a]
Arousal index, events/h, median (IQR)	20.00 (13.00–27.75)	12.50 (10.25–19.75)	.03[a]
Moderate-severe OSA, n (%)	16 (57)	6 (38)	.09[b]

[a]Mann-Whitney *U* test. [b]Fisher exact test. ApoE = apolipoprotein E allele, BMI = body mass index, GAD-7 = Generalized Anxiety Disorder 7-item questionnaire, IQR = interquartile range (25th, 75th quartiles), NART-FSIQ = National Adult Reading Test-Full Scale, PHQ-9 = Patient Health Questionnaire-9, PSG = polysomnography.

of concordance between the SWA and home-PSG in patients with stroke, we observed fixed bias in the healthy control patients for total sleep time (slope = 0.99; intercept = –94.04). Proportional bias was observed for sleep efficiency (slope = 2.05; intercept = –115.74), wake after sleep onset (slope = 2.07; intercept = 33.35), and sleep onset latency (slope = 1.13; intercept = 6.71).

Epoch-by-epoch microlevel analysis

Group summaries of average within-participant epoch-by-epoch agreement characteristics for patients with stroke and healthy control patients are presented in **Table 4.** No significant differences in within-patient epoch-by-epoch agreement between patients with stroke and healthy control patients were identified. Crude agreement between home-PSG and the SWA was 78.21% for patients with stroke and 74.20% for control patients.

Median sensitivity of the SWA (ie, agreement for sleep detection) was high (95.90%) for patients with stroke and for control patients (95.70%). Average specificity for the SWA (ie, agreement for wake detection) was fair-moderate for patients with stroke (40.45%) and moderate (47.12%) for control patients. Cohen's kappa coefficients revealed fair agreement (0.31) between the SWA and home-PSG in patients with stroke and moderate agreement (0.42) in control patients.

DISCUSSION

Stroke is a common condition with high morbidity and mortality. Disordered sleep in stroke is linked to important clinical outcomes.[1] This is the first study to undertake a robust evaluation of ambulatory and noninvasive sleep monitoring devices that will enable the objective measurement of the impact of sleep-related interventions to improve sleep in stroke. We assessed the concordance, specificity, and sensitivity of the SWA against ambulatory home-PSG in patients with ischemic stroke and healthy control patients using robust validation

Table 2—Comparison of concordance between PSG and SWA in ipsilesionally and contralesionally scored participants with stroke.

Sleep Variable	Mean Difference, in Min (SD)	95% CI Limits of Agreement (Bland-Altman)	CCC (95% CI)	RMAR Slope	RMAR Intercept
Ipsilesional					
Total sleep time	–77.22 (71.28)	–216.93 to 62.48	0.49 (0.28 to 0.71)	0.99	–76.67
Sleep efficiency	–14.78 (14.33)	–42.88 to 13.32	0.23 (0.01 to 0.44)	1.41	–49.58
Wake after sleep onset	63.59 (60.79)	–55.56 to 182.74	0.30 (0.08 to 0.52)	1.36	40.44
Sleep onset latency	3.51 (10.23)	–16.53 to 23.57	0.16 (–0.13 to 0.46)	1.78	–2.16
Contralesional					
Total sleep time	–80.15 (70.16)	–217.67 to 57.37	0.48 (0.26 to 0.70)	0.97	–68.71
Sleep efficiency	–15.15 (14.74)	–44.03 to 13.74	0.21 (–0.01 to 0.42)	1.42	–50.87
Wake after sleep onset	65.04 (60.83)	–54.19 to 184.26	0.30 (0.09 to 0.52)	1.38	40.42
Sleep onset latency	2.93 (10.18)	–17.02 to 22.87	0.03 (–0.30 to 0.36)	1.58	–1.31

CCC = Lin's concordance correlation coefficients, CI = confidence interval (lower–upper limit), PSG = polysomnography, RMAR = reduced major axis regression, SD = standard deviation, SWA = SenseWear armband.

statistics. Despite an acceptable (> 78%) average within-patient epoch-by-epoch agreement in patients with stroke, the SWA did not exceed moderate-to-fair agreement (CCC > 0.50) for any measured sleep variables. Interestingly, the variability in agreement between the SWA and home-PSG was consistent between the ipsilesionally vs contralesionally scored sleep of patients with stroke. Unihemispheric sleep EEG changes over the ipsilesional hemisphere have previously been described after acute stroke.[20] Our null hemispheric findings suggest that irrespective of stroke localization, the SWA's sleep detection is unlikely to be influenced by stroke hemisphere in the chronic stages of stroke.

In addition, we observed proportional bias according to the RMAR for all sleep variables excluding total sleep time. In line with previous validation studies, total sleep time was the most robustly quantified sleep variable computed by the SWA. We observed no evidence of proportional bias for total sleep time; however, a fixed bias was evident and the SWA systematically overestimated total sleep time by 77 minutes. The variability in agreement between home-PSG and the SWA for all sleep variables was extensive, and the SWA performed poorest for sleep measures requiring discrimination of wakefulness. For example, accurate measurement of sleep onset latency, an index of time elapsed from lights-out to sleep onset, requires differentiation of mere lying down/quiet restfulness from sleep initiation.

Unlike the more frequently used Bland-Altman analysis, the RMAR allowed us to quantify both fixed and proportional biases between the SWA and home-PSG.[27,28] By separating the fixed and proportional biases of the SWA, we were able to determine whether the SWA gave differing values relative to home-PSG across an entire range of measurement and whether the SWA estimation diverged progressively. Our results suggest that the measurement of total sleep time using the SWA diverge from home-PSG across an entire measurement range in a consistent amount across magnitude. Thus, the SWA's overestimation of total sleep time may be easier to rectify (ie,

transform) when compared with sleep-wake variables with proportional bias such as sleep efficiency, wake after sleep onset, and sleep onset latency.[29]

Although the SWA includes additional biosensors (eg, galvanic scale, near-body ambient skin temperature, heat flux), which may enhance sleep-wake discrimination beyond micromovement detection via traditional accelerometer, we observed a relatively low specificity (ie, wake detection) of 40% for the SWA in patients with stroke.[23] Whereas previous validation work comparing triaxial vs uniaxial accelerometers showed a higher epoch-by-epoch agreement in the triaxial device, a uniaxial device exhibited higher concordance for wake after sleep onset and total sleep time.[30] The proprietary sleep-wake measurement algorithms of the SWA prevented us from exploring biosensors that may be contributing to poor specificity (ie, wakefulness) detection.

The armband nature of the SWA may provide a novel solution to sleep measurement in clinical populations with upper limb injuries resulting from stroke. Relative to wrist-worn monitors, placement of the SWA on the upper arm minimizes noise associated with small micromovements of the wrist.[31] Although the SWA has previously been shown to outperform other consumer wrist-worn accelerometers such as the Fitbit (Fitbit Inc, San Francisco, CA, USA) and Jawbone Up (Jawbone Inc, San Francisco, CA, USA), it remains unclear whether the armband vs wrist-worn nature of the SWA or its additional sensors are responsible for enhanced sleep-wake detection because of the device's proprietary black-box algorithms.[32] Furthermore, no studies to date have assessed the SWA as a sleep monitor in patients with spinal cord injury or stroke where upper limb injury is pervasive. However, the SWA has been used to assess energy expenditure in patients with stroke with conflicting results. Manns and Haennel[33] found good agreement (CCC = 0.70) in a small sample of 12 patients with chronic stroke when comparing energy expenditure measured by the SWA and the StepWatch Activity Monitor. In the acute stages of stroke (within 7 days of stroke), however,

Figure 1—RMAR plots of polysomnographically scored contralesional and ipsilesional sleep variables vs SWA scores.
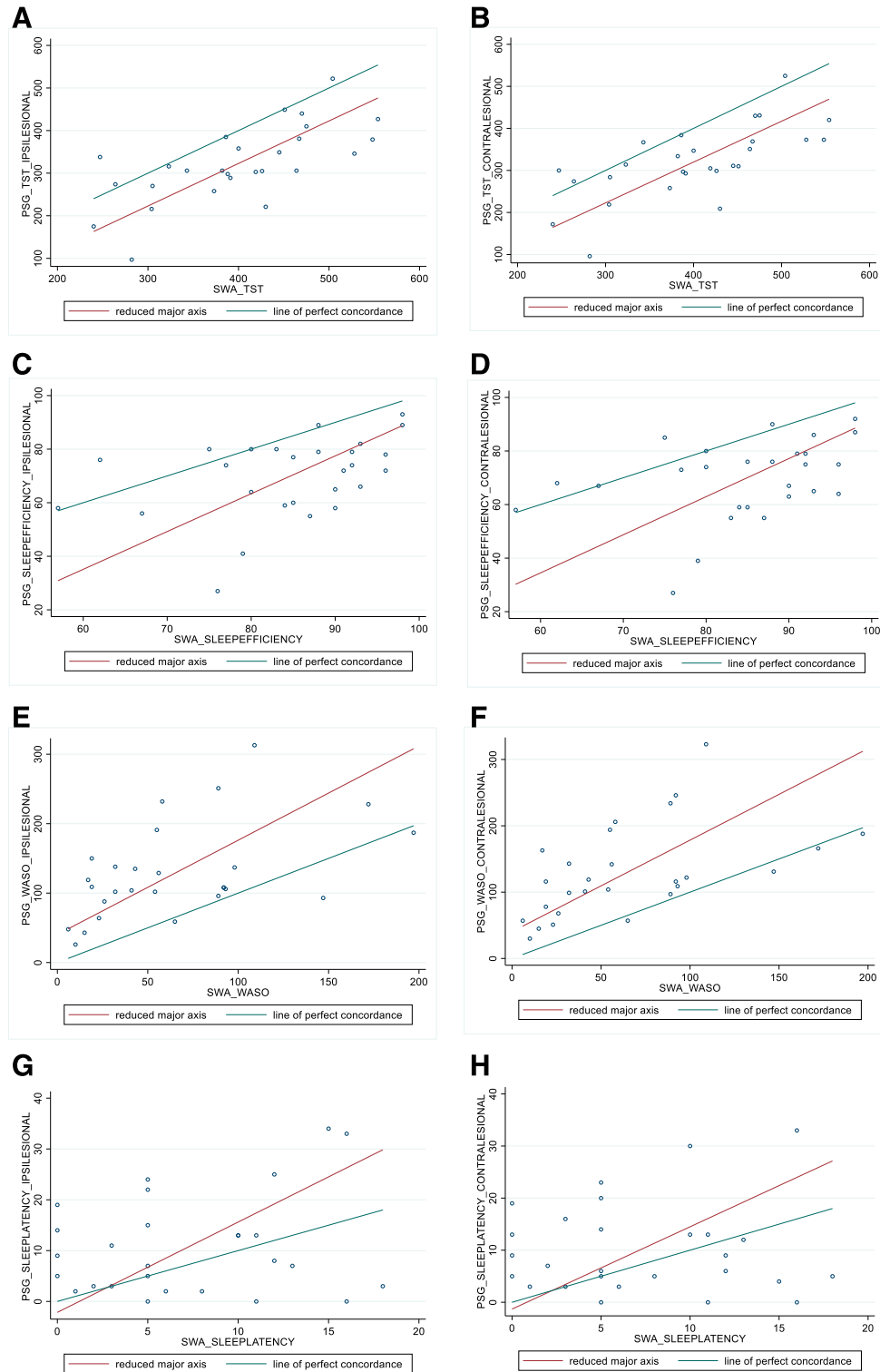
Total sleep time (**A, B**), sleep efficiency (**C, D**), wake after sleep onset (**E, F**), and sleep onset latency (**G, H**). PSG = polysomnography, RMAR = reduced major axis regression, SWA = SenseWear armband, TST = total sleep time, WASO = wake after sleep onset.

Kramer and colleagues[34] found poor agreement (CCC < 0.40) between the SWA placed on the unaffected arm and the metabolic cart. Neither study compared the concordance of the SWA between the unaffected vs the affected arm. It is therefore unclear whether upper limb impairment contributed to the SWA's performance. Future studies systematically assessing the SWA's sleep-wake detection in populations with upper limb injuries are warranted.

Table 3—Comparison of concordance between PSG and SWA in healthy control participants.

Sleep Variable	Mean Difference in Min (SD)	95% CI Limits of Agreement (Bland-Altman)	CCC (95% CI)	RMAR Slope	RMAR Intercept
Total sleep time	–99.88 (56.52)	–210.65 to 10.90	0.10 (–0.07 to 0.26)	0.99	–94.04
Sleep efficiency	–21.06 (11.89)	–44.37 to 2.25	0.03 (–0.07 to 0.14)	2.05	–115.74
Wake after sleep	80.63 (55.94)	–29.01 to 190.26	0.04 (–0.10 to 0.19)	2.07	33.35
Sleep onset latency	7.68 (7.94)	–7.87 to 23.24	0.37 (0.06 to 0.69)	1.13	6.71

CCC = Lin's concordance correlation coefficients, CI = confidence interval (lower–upper limit), PSG = polysomnography, RMAR = reduced major axis regression, SD = standard deviation, SWA = SenseWear armband.

Table 4—Within-participant epoch-by-epoch agreement measures in patients with stroke and healthy control patients.

	Stroke (n = 27)	Control (n = 16)	*P* Value
Crude agreement, median (IQR)	78.21 (73.21–81.45)	74.20 (70.90–77.96)	.25[a]
Kappa coefficient, median (IQR)	0.31 (0.10–0.44)	0.42 (0.33–0.49)	.16[a]
Sensitivity, %, median (IQR)	95.90 (84.98–97.35)	95.70 (88.15–97.90)	.87[a]
Specificity, %, median (IQR)	40.45 (18.97–54.57)	45.60 (34.05–55.25)	.18[a]

[a]Mann-Whitney *U* test. IQR = interquartile range (25th, 75th quartiles).

Reduced agreement and increased variability between PSG and the SWA generally occurred in unison with poorer sleep and increased wakefulness. For example, large variability and lack of concordance was particularly evident when sleep efficiency was < 80% (**Figure 1C, Figure1D**), and a similar trend was observed for > 100 minutes of wake after sleep onset (**Figure 1E, Figure 1F**). Previous validation studies of the SWA have been undertaken in a range of populations including healthy adults[10] and patients with OSA.[14] We are the first to examine the validity of the SWA in patients with a neurological disorder in which sleep architectural disruption because of infarction and de novo sleep disorders such as OSA are common. Unexpectedly, we found a high prevalence of previously undiagnosed moderate-to-severe OSA in both cohorts. Despite excluding participants with previously diagnosed OSA, we found that over half of all patients with stroke exhibited undiagnosed moderate-to-severe OSA, defined as AHI > 15 events/h. The high prevalence of OSA in our stroke cohort is unsurprising; poststroke OSA prevalence ranges from 44%–72%.[35] In addition to our validation findings, our respiratory PSG-derived finding highlights the importance of formal sleep studies in patients with stroke to identify undiagnosed OSA.

Interestingly, both our neurologically intact control patients and patients with stroke exhibited relatively poor sleep-wake quality (ie, sleep efficiency < 70%) and comparable SWA-PSG concordance—possibly because of a lack of habituation period for a single night of PSG. Nonetheless, our epoch-by-epoch agreement and systematic biases are comparable to prior validation work. For example, Soric et al[13] and Roane et al[23] reported high individual variability in the SWA's estimation of total sleep time and an overestimation of total sleep time by > 60 minutes, respectively. However, this is in sharp contrast to Sharif and Bahammam,[14] who revealed near-perfect agreement (whole-sample intraclass correlation coefficient [ICC] = 0.92)

for all sleep variables measured in a heterogeneous OSA sample of patients. They did not assess a binary epoch-by-epoch agreement rating. However, similar to our findings, they also generally observed lower agreement and widening of variability between the SWA and PSG in participants with sleep efficiency < 60%.[14] The significant differences in agreement between our findings may therefore result from our cohort's poorer sleep efficiency (70% vs 73%–79%) and longer total sleep time (319 minutes vs 187–290 minutes). These findings suggest that caution is warranted when utilizing the SWA in populations with excessive nighttime awakenings and fragmentation mediating poor sleep efficiency.

Limitations

An analytical limitation of the present study was that the SWA and home-PSG recorded sleep at differing epochs, which required potentially imperfect matching. Specifically, 30-second epochs were recorded for the home-PSG, whereas the SWA recorded sleep in 1-minute epochs. Thus, in our epoch-by-epoch analysis, we divided the SWA outputs to correspond with each 30-second PSG epoch. This procedure may have introduced bias into our SWA's epoch-by-epoch agreement rate, including an overestimation of discordance with home-PSG. Additional limitations lie in our cohort, which was relatively small. However, our sample size of 44 participants was comparable to or larger than previous published SWA validation studies that included epoch-by-epoch analyses.[12,23,24] Our sample also included patients with mild stroke severity (baseline National Institutes of Health Stroke Scale score, mean = 2.96). Because sleep-wake dysfunction after stroke parallels stroke severity, the SWA may perform worse in a larger, more heterogeneous sample of patients with stroke with moderate-severe stroke severity. Our study protocol only included a single sleep night with no PSG habituation period. Although our use of

at-home ambulatory PSG may have remediated the deleterious environmental stressors associated with in-lab PSG studies, multinight validation protocols should be prioritized when viable. Furthermore, the limiting nature of proprietary black-box algorithms common in consumer devices prevented us from identifying which biosensors were responsible for the SWA's poor wake detection. Finally, a potential limitation was our use of a single scorer for sleep staging. To minimize scorer bias, we utilized an experienced research-grade sleep scientist with an extensive track record of satisfactory participation in an external analysis quality assurance program. In addition, the use of a single scorer has the advantage of avoiding interscorer variability. We also limited the confounding effects of variably classified events by assessing epochs binarily as sleep or wake rather than comparing sleep subtypes/stages, which are prone to interscorer variability and errors (ie, differentiating awake/NREM-1, NREM-1/NREM-2, and NREM-2/NREM-3 sleep).[36]

CONCLUSIONS

Overall, the SWA shows promise as an ambulatory tool to measure sleep-wake at a group level. Specificity at an individual level, however, is only moderate-fair and warrants caution when used as a diagnostic tool or in populations with significant sleep-wake fragmentation. This is especially the case if sleep-wake variables beyond total sleep time (eg, sleep onset latency, sleep efficiency, and wake after sleep onset) are used as primary outcome measures. It is important to assess sleep-wake parameters beyond total sleep time in stroke populations; although patients with stroke often exhibit long sleep duration, sleep may be fragmented with increased wake after sleep onset and poor sleep efficiency.[37,38] Similarly, sleep architecture may be disturbed after stroke.[38] Earlier work suggested that patients with stroke exhibit alterations to NREM-3/slow-wave sleep, which may be mediated by sleep disorders such as OSA.[20] Nonetheless, there are several advantages of the SWA over traditional PSG,[8,39] which generally mirror those of actigraphy and include: its provision of longitudinal-prospective monitoring in a naturalistic environment; its ambulatory, noninvasive, and cost-effective technology; its ability to discriminate periods when the device is not worn; and its limited effect on sleep architecture (ie, a habituation period is not needed). Future validation studies are warranted in heterogeneous stroke populations to examine the effects of variable stroke severities (and stroke topographies) on device performance. In addition, future studies should examine the SWA in tandem with more recently developed and open-source (nonproprietary) multisensor ambulatory sleep-wake technologies to establish which array of biosensors and software contributes to optimal (> 90%) specificity and sensitivity detection.[40]

ABBREVIATIONS

CCC, Lin's concordance correlation coefficient
EEG, electroencephalography
home-PSG, at-home polysomnography

PSG, polysomnography
RMAR, reduced major axis regression
SWA, SenseWear armband

REFERENCES

1. Campbell BCV, De Silva DA, Macleod MR, et al. Ischaemic stroke. *Nat Rev Dis Primers*. 2019;5(1):70.

2. Gottlieb E, Landau E, Baxter H, Werden E, Howard ME, Brodtmann A. The bidirectional impact of sleep and circadian rhythm dysfunction in human ischaemic stroke: a systematic review. *Sleep Med Rev*. June 2019;45:54–69.

3. Redline S, Foody J. Sleep disturbances: time to join the top 10 potentially modifiable cardiovascular risk factors? *Circulation*. 2011;124(19):2049–2051.

4. Jike M, Itani O, Watanabe N, Buysse DJ, Kaneita Y. Long sleep duration and health outcomes: a systematic review, meta-analysis and meta-regression. *Sleep Med Rev*. June 2018;39:25–36.

5. Gottlieb E, Egorova N, Khlif MS, et al. Regional neurodegeneration correlates with sleep-wake dysfunction after stroke. *Sleep*. 2020;43(9):zsaa054.

6. Baglioni C, Nissen C, Schweinoch A, et al. Polysomnographic characteristics of sleep in stroke: a systematic review and meta-analysis. *PLoS One*.2016;11(3): e0148496. *Erratum appears in PLoS One*. 2016;11(5):e0155652.

7. Scott H, Lack L, Lovato N. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. *Sleep Med Rev*. February 2020;49: 101227.

8. Depner CM, Cheng PC, Devine JK, et al. Wearable technologies for developing sleep and circadian biomarkers: a summary of workshop discussions. *Sleep*. 2020; 43(2):zsz254.

9. Merino-Andreu M, Arnulf I, Konofal E, Derenne JP, Agid Y. Unawareness of naps in Parkinson's disease and in disorders with excessive daytime sleepiness. *Neurology*. 2003;60(9):1553–1554.

10. Andre D, Pelletier R, Farringdon J, et al. The Development of the SenseWear® Armband, a Revolutionary Energy Assessment Device to Assess Physical Activity and Lifestyle. BodyMedia, Inc. http://www.dotfit.com/sites/63/templates/ categories/images/1783/dev_sensewear_article.pdf. Updated 2006. Accessed October 1, 2020.

11. Peterson BT, Chiao P, Pickering E, et al. Comparison of actigraphy and polysomnography to assess effects of zolpidem in a clinical research unit. *Sleep Med*. 2012;13(4):419–424.

12. Shin M, Swan P, Chow CM. The validity of Actiwatch2 and SenseWear armband compared against polysomnography at different ambient temperature conditions. *Sleep Sci*. 2015;8(1):9–15.

13. Soric M, Turkalj M, Kucic D, Marusic I, Plavec D, Misigoj-Durakovic M. Validation of a multi-sensor activity monitor for assessing sleep in children and adolescents. *Sleep Med*. 2013;14(2):201–205.

14. Sharif MM, Bahammam AS. Sleep estimation using BodyMedia's SenseWear™ armband in patients with obstructive sleep apnea. *Ann Thorac Med*. 2013;8(1):53–57.

15. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke*. 1993;24(1):35–41.

16. Bamford J, Sandercock P, Dennis M, Burn J, Warlow C. Classification and natural history of clinically identifiable subtypes of cerebral infarction. *Lancet*. 1991; 337(8756):1521–1526.

17. World Medical Association. WMA Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects. https://www.wma.net/policies-post/ wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving- human-subjects/. Updated July 9, 2018. Accessed October 1, 2020.

18. Berry RB, Quan SF, Abreu AR, et al; for the American Academy of Sleep Medicine. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Version 2.6. Darien, IL: American Academy of Sleep Medicine; 2020.

19. Jorgensen G, Downey C, Goldin J, Melehan K, Rochford P, Ruehland W. An Australasian commentary on the AASM Manual for the Scoring of Sleep and Associated Events. *Sleep Biol Rhythms*. 2020;18:163–185.

20. Poryazova R, Huber R, Khatami R, et al. Topographic sleep EEG changes in the acute and chronic stage of hemispheric stroke. *J Sleep Res*. 2015;24(1):54–65.

21. Stevens D, Martins RT, Mukherjee S, Vakulin A. Post-stroke sleep-disordered breathing—pathophysiology and therapy options. *Front Surg*. 2018;5:9.

22. Stahl SM, Yaggi HK, Taylor S, et al. Infarct location and sleep apnea: evaluating the potential association in acute ischemic stroke. *Sleep Med*. 2015;16(10):1198–1203.

23. Roane BM, Van Reen E, Hart CN, Wing R, Carskadon MA. Estimating sleep from multisensory armband measurements: validity and reliability in teens. *J Sleep Res*. 2015;24(6):714–721.

24. O'Driscoll DM, Turton AR, Copland JM, Strauss BJ, Hamilton GS. Energy expenditure in obstructive sleep apnea: validation of a multiple physiological sensor for determination of sleep and wake. *Sleep Breath*. 2013;17(1):139–146.

25. Smith RJ. Use and misuse of the reduced major axis for line-fitting. *Am J Phys Anthropol*. 2009;140(3):476–486.

26. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360–363.

27. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. *Clin Exp Pharmacol.Physiol*. 2002;29(7):527–536.

28. Ludbrook J. Confidence in Altman-Bland plots: a critical review of the method of differences. *Clin Exp Pharmacol Physiol*. 2010;37(2):143–149.

29. Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: but which regression? *Clin Exp Pharmacol Physiol*. 2010;37(7):692–699.

30. Cellini N, McDevitt EA, Ricker AA, Rowe KM, Mednick SC. Validation of an automated wireless system for sleep monitoring during daytime naps. *Behav Sleep Med*. 2015;13(2):157–168.

31. Sunseri M, Liden CB, Farringdon J, et al. The SenseWear Armband as a sleep detection device. BodyMedia, Inc. internal white paper. 2009:1–9. https://pdfs.semanticscholar.org/e36b/f00cac00e094eb605bbf2bcc5a384e96874c.pdf+&cd=1&hl=en&ct=clnk&gl=au. Accessed June 2, 2020.

32. Gruwez A, Bruyneel A-V, Bruyneel M. The validity of two commercially-available sleep trackers and actigraphy for assessment of sleep parameters in obstructive sleep apnea patients. *PLoS One*. 2019;14(1):e0210569.

33. Manns PJ, Haennel RG. SenseWear armband and stroke: validity of energy expenditure and step count measurement during walking. *Stroke Res Treat*. 2012;2012:247165.

34. Kramer SF, Johnson L, Bernhardt J, Cumming T. Validity of multisensor array for measuring energy expenditure of an activity bout in early stroke survivors. *Stroke Res Treat*. 2018;2018:9134547.

35. Pasic Z, Smajlovic D, Dostovic Z, Kojic B, Selmanovic S. Incidence and types of sleep disorders in patients with stroke. *Med Arh*. 2011;65(4):225–227.

36. Younes M, Raneri J, Hanly P. Staging sleep in polysomnograms: analysis of inter-scorer variability. *J Clin Sleep Med*. 2016;12(6):885–894.

37. Duss SB, Seiler A, Schmidt MH, et al. The role of sleep in recovery following ischemic stroke: a review of human and animal data. *Neurobiol Sleep Circadian Rhythms*. November 2016;2:94–105.

38. Terzoudi A, Vorvolakos T, Heliopoulos I, Livaditis M, Vadikolias K, Piperidou H. Sleep architecture in stroke and relation to outcome. *Eur Neurol*. 2009;61(1):16–22.

39. Roberts DM, Schade MM, Mathew GM, Gartenberg D, Buxton OM. Detecting sleep using heart rate and motion data from multisensor consumer-grade wearables, relative to wrist actigraphy and polysomnography. *Sleep*. 2020;43(7):zsaa045.

40. Kolla BP, Mansukhani S, Mansukhani MP. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Rev Med Devices*. 2016;13(5):497–506.

SUBMISSION & CORRESPONDENCE INFORMATION

DISCLOSURE STATEMENT