Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrating Wearable Sleep Monitoring with Repeated Balance and Cognitive Testing: A Six-Week Longitudinal Observational Cohort and Feasibility Study Among Older Adults

Hongwu Wang

hongwu.wang@ufl.edu

University of Florida

Yuan Li

University of Florida

Raghuveer Chandrashekhar

University of Pittsburgh

Yun Chan Shin

University of Florida

Jason Arias

University of Florida

Isabella F. Tietje

University of Florida

Shrinath Shah

University of Florida

Ishaan Gupta

University of Florida

Eileen Stefaniak

University of Florida

Sheya D. Patel

University of Florida

Lauren R Smith

University of Florida

Article

Keywords:

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11036810/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Preserving functional independence is a primary goal of successful aging, yet the temporal dynamics linking sleep, motor balance, and cognition remain underexplored. This six-week longitudinal study evaluated the feasibility of continuous wearable sleep monitoring and unraveled the "Sleep-Balance-Cognition Triad" in 15 community-dwelling older adults. Sleep was tracked objectively via a consumer-grade wearable (Oura Ring) and subjectively via daily diaries. Motor function (Mini-BESTest, Timed Up and Go) and cognitive performance (NIH Toolbox) were assessed weekly. We evaluated both between-person habitual sleep traits and within-person acute temporal deviations using rolling multi-day windows. Results demonstrated exceptional feasibility for the wearable (100% compliance) compared to daily diaries (89.4%). Bland-Altman analysis revealed severe subjective sleep state misperception: participants systematically underestimated nocturnal awakenings by 64.9 minutes and overestimated total sleep time by 31.3 minutes. Habitual sleep consistency was protective of cognitive health, with longer average sleep strongly correlating with better working memory. Intra-individual analyses revealed a striking dissociation: while cognitive performance remained resilient to acute sleep deviations, motor systems exhibited dose-dependent vulnerability. Extended sleep and lethargy periods of 6 to 10 days or more significantly impaired dynamic balance and slowed functional mobility. In conclusion, sleep exerts a complex, dynamic impact on functional capacity. While long-term consistency buffers cognitive reserve, acute sleep fluctuations unmask progressive motor vulnerabilities. Furthermore, significant discrepancies in subjective recall highlight the clinical necessity of objective tracking. Continuous commercial wearables offer a highly feasible, actionable tool for the remote monitoring of fall risk and functional decline in older adults.

1. INTRODUCTION

Preserving functional independence is the primary goal of successful aging[1]. However, the characteristics of restorative and rejuvenating sleep quality and sleep duration often decline as a fundamental component of the aging process[2]. Poor sleep quality directly impairs the brain’s ability to maintain postural control, steady balance, and gait[3,4]. Alongside this, older adults experiencing sleep deficits often demonstrate impaired executive functioning, memory, and reasoning[5]. Recent efforts to synthesize the literature have highlighted the complex interplay between sleep and functional status. A recent systematic scoping review[6] mapping the relationships among sleep, motor balance, and cognition emphasized that most existing research relies on cross-sectional designs. Correlations between poor sleep and functional decline are significant, while the temporal mechanisms among these three remain underexplored despite their importance[6].

To bridge the gap between traditional laboratory polysomnography (PSG) and real-life function[7], there is a growing need for continuous, noninvasive tracking in community settings. Wearable technologies[8,9], such as the Oura Ring Generation 3[10], have emerged as robust tools for clinical research. Comparisons against gold-standard PSG have demonstrated high inter-device agreement, reporting high sensitivity

(0.96) for sleep detection and strong correlations for Total Sleep Time (TST) and Wake After Sleep Onset (WASO)[11–13]. Despite this validation, there has been limited research mapping continuous data from commercial wearables to meaningful, longitudinal clinical outcomes among older adults.

This study seeks to address this gap by determining if objective sleep characteristics captured in the home environment can serve as practical indicators of functional capacity. Specifically, this research unravels the temporal dynamics of the "Sleep–Balance–Cognition Triad" by evaluating both stable habitual sleep characteristics[14,15] (between-person differences) and acute temporal deviations[16–18] from a personal baseline (within-person fluctuations). By transitioning from traditional laboratory polysomnography to continuous, home-based digital health monitoring, this study seeks to determine if objective data from commercial wearables can serve as practical indicators of functional capacity. Specifically, we are aiming to 1) determine if habitual, stable individual differences in baseline sleep characteristics are associated with overall functional status in motor balance and cognitive performance among healthy older adults; 2) evaluate whether acute, weekly temporal deviations from a participant's personal sleep baseline predict immediate, week-to-week fluctuations in cognitive and motor balance performance; and 3) to establish the study's protocol feasibility, user compliance, and technical practicality of combining continuous consumer-grade wearable monitoring with repetitive weekly clinical testing in a community setting.

2. METHODS

Ethics

The study protocol was approved by the University of Florida Institutional Review Board (IRB202302147). All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki. The study was registered on ClinicalTrials.gov (NCT06584136) on 15/11/ 2024, and all participants provided written informed consent prior to enrollment.

Study Design and Participants

This study utilized a 6-week longitudinal, observational design. Participants were community-dwelling older adults recruited from the Oak Hammock Retirement Community at the University of Florida, between November 2024 and August 2025. To be eligible, individuals had to be aged 65 or older, English-speaking, and able to walk independently or with minimal assistance (e.g., a cane or walker). Exclusion criteria included severe cognitive impairment (e.g., diagnosed dementia), severe sleep disorders (e.g., sleep apnea requiring a CPAP machine), significant neurological conditions (e.g., Parkinson’s disease, stroke), and unstable medical conditions.

Recruitment

We utilized our existing protocols, stakeholder networks, the Institute for Driving, Activity, Participation, and Technology (I-DAPT)'s recruiting infrastructure, and our strong integration with residents at the Oak

Hammock retirement facility to meet our recruitment goals for this study.

A sample of 15 is adequate for a pilot feasibility study based on the literature[19–21] and our recent systematic review[6]. However, to account for drop-out from longitudinal observation and/or attrition, we over-recruited and enrolled 18 participants in the study.

Procedures and Data Collection

Following a baseline screening visit (Week 0) to collect demographic data, medical history, and baseline characteristics, participants were observed over a 6-week continuous tracking period. Continuous tracking included: Objective Sleep Monitoring using an Oura Ring Gen 3, worn continuously (except during charging) to capture TST, Sleep Efficiency (SE), WASO, and sleep stage. Subjective Sleep Diary: Participants completed a daily paper or electronic sleep diary every morning and evening to log perceived sleep patterns using the American Sleep Foundation Sleep Diary[22]. Weekly functional assessments (Weeks 1 through 6) were conducted in-person at the Technology for Occupational Performance (TOP) Laboratory. All assessments were conducted under standardized conditions to minimize variability.

Outcome Measures

Sleep Measures

Objective Sleep Monitoring (Oura Ring Gen 3): Worn continuously to capture objective sleep architecture, including Time in Bed (TIB), Total Sleep Time (TST), Sleep Efficiency (SE), Wake After Sleep Onset (WASO), and sleep staging. The Oura Ring is a highly validated commercial multisensor wearable. Validation studies comparing the Oura Ring to gold-standard laboratory polysomnography (PSG) have demonstrated high inter-device agreement, reporting a sleep detection sensitivity of 0.96 and strong correlations for evaluating TST and WASO[11–13]. Its noninvasive design minimizes the "first-night effect"[23] and eliminates the disruption of lab-based equipment, offering excellent ecological validity and measurement reliability for continuous, longitudinal tracking in the home environment[13].

Subjective Sleep Diary: Completed daily (morning and evening) to record self-perceived sleep parameters, including TIB, Sleep Onset Latency (SOL), WASO, TST, and SE. Daily sleep diaries are widely regarded as the clinical gold standard for assessing subjective sleep continuity and behavior[22]. They demonstrate strong face validity and test-retest reliability for capturing day-to-day variations in a patient's perception of their restfulness[24–26]. Using this subjective measure alongside objective wearables is crucial for capturing a holistic sleep profile and identifying sleep state misperception, a well-documented clinical phenomenon in older adults in which perceived sleep duration differs significantly from actual physiological sleep architecture.

Balance and Mobility

Berg Balance Scale (BBS): Administered at Week 1 and Week 6 to assess overall static and functional balance (scored 0–56). The BBS is widely recognized as a gold-standard clinical measure for assessing fall risk in older adults[27,28], demonstrating excellent inter-rater and test-retest reliability (Intraclass Correlation Coefficients (ICCs) > 0.90) and strong concurrent validity with other functional mobility scales[29–33]. Its inclusion provides a robust, validated baseline and endpoint metric for global postural stability.

Mini-Balance Evaluation Systems Test (Mini-BESTest): Administered weekly to evaluate dynamic balance, anticipatory postural adjustments, and reactive postural control. This 14-item assessment minimizes the ceiling effects frequently observed in traditional balance scales, making it highly responsive to subtle, state-level functional fluctuations[34]. It possesses excellent internal consistency, inter-rater reliability (ICC > 0.90), and robust predictive validity for identifying dynamic instability and recurrent fall risk in community-dwelling populations[35–37].

Timed Up and Go (TUG): Assessed weekly to measure functional mobility and determine Dual-Task Cost when paired with a secondary cognitive challenge[38]. The baseline TUG exhibits exceptional test-retest reliability (ICC > 0.95) and construct validity for assessing basic mobility[39–42]. Calculating the Dual-Task Cost, by introducing a simultaneous cognitive load, significantly increases the test's sensitivity to cognitive-motor interference, providing a psychometrically sound and ecologically valid proxy for real-world fall risk under conditions of divided attention[42].

Cognitive Function

NIH Toolbox Cognition Battery: Administered weekly using brief, standardized tasks to evaluate Working Memory, Processing Speed, Attention, and Executive Function[43]. The NIH Toolbox is a state-of-the-art, extensively validated neurobehavioral assessment platform that demonstrates excellent construct validity when compared against traditional gold-standard neuropsychological batteries[44,45]. It possesses high test-retest reliability (with Intraclass Correlation Coefficients typically exceeding 0.80 for fluid cognition tasks), making it highly sensitive and psychometrically sound for repeated longitudinal testing. Furthermore, it generates demographically corrected normative T-scores that dynamically adjust for age and education[46]. This ensures that the week-to-week cognitive fluctuations captured in our study represent true state-level deviations rather than baseline sociodemographic differences.

Secondary Measures and Potential Confounders

To thoroughly contextualize the findings and control for variables that may influence sleep, balance, or cognition, a comprehensive suite of secondary assessments and surveys was administered.

Morningness-Eveningness Questionnaire (MEQ): Completed at baseline to assess an individual’s circadian preferences (sleep chronotype), determining whether they are a "morning" or "evening" person[46]. The MEQ is a widely validated instrument demonstrating high internal consistency (Cronbach’s

α > 0.80) and robust test-retest reliability[47–49]. It exhibits strong criterion validity[50], evidenced by significant correlations with physiological circadian markers, such as peak melatonin onset and core body temperature fluctuations. Identifying chronotypes is crucial for contextualizing inter-individual variations in sleep timing and daytime functional performance.

Home Falls and Accidents Screening Tool (HOMEFAST): Administered at baseline to evaluate fall risks and environmental safety in participants' home settings, including questions about the usage and perceived benefits of smart home technologies[51]. HOMEFAST is a recognized environmental assessment tool with excellent inter-rater reliability (Cohen's κ > 0.80) and proven predictive validity for identifying future fall risk in community-dwelling older adults[52–54]. Its inclusion allows control of extrinsic environmental fall hazards when isolating the intrinsic motor and cognitive factors that contribute to balance.

Life-Space Assessment (LSA): Administered at baseline to measure the participant's spatial activity patterns, community engagement, and overall mobility movement in different environments over the past month[55]. The LSA demonstrates excellent test-retest reliability (ICC > 0.90) and strong construct validity, correlating significantly with objective measures of physical performance and independence in activities of daily living (ADLs)[56]. It provides a highly ecologically valid metric of how a participant's functional capacity translates into real-world mobility and community participation.

Health and Lifestyle Profiles: A demographic questionnaire (capturing age, gender, race, ethnicity, and marital status), a medical history form (documenting allergies, medications, and relevant conditions), and the International Physical Activity Questionnaire[57] were collected at baseline and the end visit. These standardized self-report instruments exhibit high face and content validity. They are essential for capturing critical covariates, such as polypharmacy, baseline comorbidities, and habitual physical activity levels, that are known to independently influence sleep architecture and neuromuscular aging.

Exit Survey: Administered at the conclusion of the 6-week observation period (Week 6) to evaluate overall protocol acceptability and participant burden. This 17-item survey captured categorical and open-ended feedback regarding wearable device comfort (OURA ring), perceived safety and environmental focus during clinical assessments, and self-reported behavioral or educational benefits resulting from study participation.

Data Collection and Management

Data were collected in a manner consistent with transparent reporting as mandated by the CONSORT guidelines[58]. See Fig. 1 for the CONSORT Chart. Specifically, data were entered by the trained research assistant into the NIH Toolbox-Cognition Battery app and REDCap[59]. All data were scrutinized for quality control by the principal investigator at regular intervals (at least 1x per month). Any discrepancies were resolved, and the research assistant and team members were trained to ensure compliance, accuracy, and consistency of data recording.

Data Analysis

Demographic data are presented using summary statistics appropriate to the level of measurement. To standardize the analysis timeline, study dates were converted to study days, and data were analyzed using advanced statistical methods to address the specific aims. Missing data were handled analytically based on the specific statistical approach. Linear Mixed-Effects models inherently accommodated missing repeated measures using maximum likelihood estimation, while pairwise deletion was applied to all correlational and concordance analyses to ensure that only strictly matched data points were evaluated.

Aim 1 Analysis (Between-Person / Habitual Patterns): Descriptive statistics were calculated for all variables. To determine the between-person relationship, the mean of each participant's sleep metrics (TST, SE, WASO) and sleep variability (Standard Deviation of WASO) across the 6 weeks were calculated. Pearson correlation coefficients were utilized to examine associations between these habitual sleep patterns and average functional performance. Cognitive domain data were analyzed using age- and education-adjusted T-scores to control for known sociodemographic factors. Night-to-night sleep fragmentation was quantified using the standard deviation (SD) of WASO. Overall functional status was calculated as the mean of each participant's weekly motor scores (MiniBEST, TUG, Dual-Task Cost) and age- and education-adjusted cognitive T-Scores (Working Memory, Processing Speed, Attention, Executive Function). Bivariate Pearson correlation coefficients (r) were used to test the associations between the habitual sleep predictors and functional outcomes. To control the risk of Type I errors associated with multiple testing, p-values were adjusted using the Holm-Bonferroni method[60], applied separately to the cognitive and motor clusters.

Aim 2 Analysis (Within-Person / Acute Deviations): To evaluate within-person relationships between acute sleep fluctuations and functional performance, objective TST was dynamically mapped to each clinical assessment. A primary 10-day rolling window, based on a recent systematic review[61], was established to calculate the mean TST from the assessment day and the 9 preceding days.

Sleep volumes were categorized intra-individually to control for baseline inter-individual differences in sleep requirements (trait-level effects) and isolate acute, night-to-night deviations (state-level effects). By converting raw TST into personalized Z-scores, each participant served as their own control. This statistical standardization accounts for natural variations in sleep duration and variance across the cohort, preventing short or long sleepers from biasing clinical outcomes. Similar intra-individual standardization is widely accepted in longitudinal sleep architecture and behavioral medicine research to capture the true impact of acute sleep restriction and extension[62,63]:

- Normal Sleep: Mean TST within ± 1 SD of the personal average.
- Short Sleep: Mean TST < -1 SD below the personal average.
- Long Sleep: Mean TST > + 1 SD above the personal average.

Linear Mixed-Effects (LME) models were used to evaluate the absolute change in functional outcomes, specifying random intercepts for each participant and adjusting for testing week and gender. To map the cumulative threshold of functional impairment without inflating Type I error rates, an exploratory 14-day temporal sensitivity analysis was conducted on representative clinical markers: MiniBEST, Timed Up and Go (TUG), Processing Speed, and Working Memory.

Aim 3 Analysis (Feasibility and Acceptability): Feasibility was evaluated by assessing participant compliance with both subjective and objective data-collection modalities. Objective compliance was calculated as the percentage of valid nightly sleep records generated by the OURA ring out of the total expected study days. Subjective compliance was calculated as the completion rate of the daily REDCap sleep diary across the study duration. To evaluate protocol acceptability and perceived user burden, responses from the end-of-study exit survey were summarized using descriptive statistics for categorical questions (e.g., ease of enrollment, schedule manageability) and thematic review for open-ended feedback.

To evaluate the concordance between participant self-report (sleep diary) and objective sensor data (OURA ring), paired nightly data were analyzed across five primary parameters: TIB, TST, SOL, WASO, and SE. Statistical agreement was evaluated using three methods. First, Pearson correlation coefficients (r) were computed to assess linear relationships. Second, to evaluate absolute consistency beyond simple correlation, a two-way mixed-effects, absolute agreement, single rater Intraclass Correlation Coefficient (ICC 2,1) was utilized. Finally, Bland-Altman analyses were conducted to detect and quantify systematic bias. For each parameter, the mean difference (Bias = Diary - OURA) and the 95% Limits of Agreement (LOA = Bias ± 1.96 standard deviations) were calculated and visually plotted.

All analyses were conducted in R (Version 4.3.0). Statistical significance was set at $\alpha = .05$. Linear Mixed-Effects (LME) modeling was conducted using the *lme4* and *lmerTest* packages. Intraclass Correlation Coefficients (ICC) for absolute agreement were calculated using the *irr* package. Bland-Altman statistics and visualizations were generated using the *BlandAltmanLeh* and *ggplot2* packages, respectively.

3. RESULT

Eighteen community-dwelling older adults from the Oak Hammock Retirement Community at the University of Florida were recruited. Two participants withdrew before data collection, and one was excluded from the final analysis due to an unrelated death. The final sample consisted of 15 participants (mean Age: 78.1 ± 5.6 years; 53% male) and was characterized by a high level of education (mean: 17.7 ± 4.8 years) and a white majority (100%). Details are presented in Table 1. There were no statistically significant differences ($p > 0.05$) in baseline demographic characteristics between completers and dropouts.

Table 1
Baseline characteristics by completion status

		Study Completion Status		
Participant Characteristic	**Overall** N = 18	**Completer** N = 15	**Dropout** N = 3	**p-value**[1]
Age (years), Mean ± SD	78.5 ± 5.3	78.1 ± 5.6	80.7 ± 3.5	0.39
Sex, n (%)				> 0.99
Male	10 (56%)	8 (53%)	2 (67%)	
Female	8 (44%)	7 (47%)	1 (33%)	
Height (inches), Mean ± SD	67.6 ± 3.6	67.9 ± 3.9	65.7 ± 0.6	0.20
Weight (lbs), Mean ± SD	163.7 ± 37.5	166.2 ± 39.0	151.0 ± 31.8	0.64
Race, n (%)				
White	18 (100%)	15 (100%)	3 (100%)	
Marital Status, n (%)				0.31
Never married	1 (5.6%)	1 (6.7%)	0 (0%)	
Married	16 (89%)	14 (93%)	2 (67%)	
Widowed	1 (5.6%)	0 (0%)	1 (33%)	
Household Members, Mean ± SD	1.9 ± 0.3	1.9 ± 0.3	1.7 ± 0.6	0.63
Years of Education, Mean ± SD	17.7 ± 4.4	17.7 ± 4.8	17.7 ± 1.5	0.87
Primary Occupation, n (%)				> 0.99
Paid work	1 (5.6%)	1 (6.7%)	0 (0%)	
Retired	15 (83%)	12 (80%)	3 (100%)	
Unpaid work	0 (0%)	0 (0%)	0 (0%)	
Other	2 (11%)	2 (13%)	0 (0%)	
Household Income, n (%)				> 0.99
$75,000 to $99,999	1 (5.6%)	1 (6.7%)	0 (0%)	
$100,000 and over	13 (72%)	11 (73%)	2 (67%)	
Prefer not to answer	4 (22%)	3 (20%)	1 (33%)	

[1]Wilcoxon rank-sum test for continuous variables and Fisher's Exact Test for categorical variables

		Study Completion Status		
Participant Characteristic	**Overall** N = 18	**Completer** N = 15	**Dropout** N = 3	**p-value**[1]
Number of Comorbidities, Mean ± SD	2.6 ± 1.6	2.5 ± 1.6	3.0 ± 2.0	0.80
HOMEFAST Hazard Score, Mean ± SD	1.0 ± 1.1	0.9 ± 1.2	1.7 ± 0.6	0.20
LSA Composite Score, Mean ± SD	76.8 ± 10.9	74.5 ± 9.4	88.0 ± 13.1	0.12
MEQ Total Score, Mean ± SD	73.6 ± 7.2	73.7 ± 7.8	73.0 ± 3.5	> 0.99
PSQI Global Score, Mean ± SD	7.6 ± 2.3	7.6 ± 2.3	7.3 ± 2.9	0.90
Berg Balance Scale (BBS), Mean ± SD	52.6 ± 3.7	52.5 ± 3.9	53.0 ± 1.4	0.79
Missing	1	0	1	
MiniBESTest, Mean ± SD	18.5 ± 3.4	18.1 ± 3.2	21.5 ± 4.9	0.44
Missing	1	0	1	

[1]Wilcoxon rank-sum test for continuous variables and Fisher's Exact Test for categorical variables

3.1. Aim 1 Results: Between-Person / Habitual Patterns

To determine the between-person relationship, in the cognition domain, we utilized age education-adjusted T-Scores[46], which control for known sociodemographic factors. Table 2 summarizes the descriptive statistics for all variables across the sleep, motor function, and cognition domains. We found a strong, statistically significant positive correlation between habitual TST and Working Memory performance ($r = 0.59,\ p = 0.025$) as shown in Fig. 2 (left). Participants who consistently slept longer demonstrated substantially better scores on the List Sorting Working Memory task. High variability in sleep fragmentation (SD of WASO) showed a strong negative trend with Working Memory ($r = -0.49,\ p = 0.075$), as shown in Fig. 2 (middle), suggesting that highly inconsistent sleep quality may impair cognitive function. Habitual TST also showed a moderate positive association with motor Dual-Task Cost ($r = 0.46,\ p = 0.088$), as shown in Fig. 2 (right). Habitual REM sleep duration showed a moderate positive trend with Working Memory performance ($r = 0.42, p = 0.138$). Since REM sleep is strongly implicated in emotional and procedural memory consolidation, this suggests that REM is a notable subcomponent driving the overall TST benefit. REM sleep duration also showed a moderate positive trend with processing speed ($r = 0.39, p = 0.173$). While not statistically significant, habitual deep sleep (slow-wave sleep, critical for physical restoration) showed a weak negative correlation ($r = -0.27, p = 0.350$) with TUG time (i.e., more deep sleep was associated with *faster* mobility execution).

Table 2
Descriptive Summary Table (N = 15 Completers)

Category	Variable	Mean ± SD	Range (Min - Max)
Sleep Metrics (OURA Ring)	Mean of TST (min)	402.8 ± 23.4	[367.0, 431.6]
	Mean of SE (%)	79.5 ± 4.1	[71.5, 85.1]
	Mean of WASO (min)	106.4 ± 26.1	[67.6, 154.5]
	Mean of Deep Sleep (min)	32.6 ± 14.4	[7.9, 61.4]
	Mean of REM Sleep (min)	68.3 ± 17.8	[41.7, 96.0]
	Mean of Light Sleep (min)	302.0 ± 32.3	[259.1, 365.5]
	SD of TST (min)	50.8 ± 14.6	[29.7, 74.9]
	SD of SE (%)	5.3 ± 2.0	[3.3, 9.9]
	SD of WASO (min)	34.2 ± 15.2	[19.2, 72.8]
	SD of Deep Sleep (min)	16.7 ± 3.7	[8.4, 21.8]
	SD of REM Sleep (min)	22.3 ± 3.8	[18.1, 33.1]
	SD of Light Sleep (min)	45.3 ± 10.6	[30.9, 64.3]
Motor Function	Berg Balance Scale (0–56)	52.9 ± 3.4	[43.0, 56.0]
	MiniBEST Score	18.3 ± 2.5	[10.8, 21.2]
	Timed Up and Go (sec)	10.2 ± 1.4	[8.0, 13.3]
	Dual-Task Cost (%)	24.3 ± 18.3	[8.1, 81.3]
Cognitive Function (NIH Toolbox)	Working Memory (Adj. T-Score)	57.4 ± 4.7	[51.0, 64.8]
	Processing Speed (Adj. T-Score)	64.2 ± 11.8	[45.8, 81.0]
	Attention (Adj. T-Score)	55.7 ± 7.2	[43.7, 66.5]
	Executive Function (Adj. T-Score)	54.6 ± 8.4	[45.1, 74.5]

Note. TST = Total Sleep Time; SE = Sleep Efficiency; WASO = Wake After Sleep Onset; REM = Rapid Eye Movement; SD = Standard Deviation.

3.2. Aim 2 Results: Intra-Individual Effects of Acute Sleep Deviations on Functional Performance

3.2.1. Impact of Sleep Deviations on Motor Function

Intra-individual analyses utilizing the primary 10-day trailing window revealed that acute periods of extended sleep ("Long Sleep") significantly impaired functional mobility, as shown in Table 3. Compared with their personal baseline sleep, participants took nearly 1 full second longer to complete the Timed Up and Go (TUG) assessment during long sleep windows (Estimate = 0.95 seconds, 95% CI [0.27, 1.64], $p = 0.007$). Conversely, at this specific 10-day evaluation mark, dynamic balance and postural control demonstrated statistical resilience: MiniBESTest scores showed no significant decline during either Long Sleep ($p = 0.636$) or Short Sleep ($p = 0.688$). Dual-task motor costs also remained statistically stable regardless of the trailing 10-day sleep duration.

3.2.2 Resilience of Cognitive Function to Sleep Fluctuations

Consistent with the hypothesis of cognitive reserve, cognitive performance demonstrated robust resilience to acute week-to-week sleep deviations. LME models indicated no significant alterations in Processing Speed, Attention, or Executive Function during periods of either Short or Long Sleep (all $p > 0.28$). While a marginal trend suggested a slight elevation in Working Memory scores following 10-day long sleep windows (Estimate = 4.34, 95% CI [-0.79, 9.47], $p = 0.096$), no cognitive domain exhibited statistically significant degradation in response to sleep extension or restriction within the testing timeframe.

Table 3
Functional Change during 10-Day Short vs. Long Sleep Windows (Adjusted for Time and Gender)

Outcome	Sleep Deviation (10-Day)	Estimate (95% CI)	p-value
Cognition: WorkingMem	Short Sleep	1.19 [-3.16, 5.54]	0.587
Cognition: WorkingMem	Long Sleep	4.34 [-0.79, 9.47]	0.096
Cognition: ProcSpeed	Short Sleep	-0.93 [-6.42, 4.55]	0.735
Cognition: ProcSpeed	Long Sleep	-1.79 [-8.27, 4.70]	0.585
Cognition: Attention	Short Sleep	2.08 [-2.40, 6.56]	0.358
Cognition: Attention	Long Sleep	1.23 [-4.06, 6.52]	0.644
Cognition: ExecFunc	Short Sleep	-0.12 [-4.98, 4.73]	0.960
Cognition: ExecFunc	Long Sleep	3.12 [-2.62, 8.86]	0.281
Motor: MiniBEST	Short Sleep	-0.16 [-0.92, 0.61]	0.688
Motor: MiniBEST	Long Sleep	-0.22 [-1.13, 0.69]	0.636
Motor: TUG (sec)	Short Sleep	0.40 [-0.18, 0.97]	0.174
Motor: TUG (sec)	Long Sleep	0.95 [0.27, 1.64]	0.007
Motor: Dual-Task Cost (%)	Short Sleep	-7.51 [-17.15, 2.14]	0.125
Motor: Dual-Task Cost (%)	Long Sleep	-4.09 [-15.48, 7.30]	0.475

Note: Estimates represent absolute change relative to a personal 'Normal Sleep' window. Models are adjusted for testing week (practice effects) and participant gender.

3.2.3 Temporal Dose-Response (Days 1–14)

The daily dose-response tracking confirmed a striking temporal dissociation between motor and cognitive vulnerabilities. Cognitive outcomes maintained state-level stability across all 14 trailing windows, with fluctuations never crossing into statistical significance. Motor function, however, exhibited a cumulative threshold effect. Extended sleep durations began to significantly compromise dynamic balance (MiniBEST) as early as day 6. Functional mobility (TUG) demonstrated a slower, accumulating burden, taking a full 10 days of extended sleep before the slowing effect crossed the threshold of clinical significance and persisted through day 14, as shown in Fig. 3. Temporal sensitivity analyses were also conducted for Executive Function, Attention, and Dual-Task Cost, revealing identical trajectories of state-level stability across all 14 trailing windows (see Supplementary Table S1).

3.3 Aim 3: Feasibility

To assess feasibility, we tracked compliance across the expected 42-day (6-week) study period for the 15 completers (630 total expected days).

Objective Wearable Compliance: Feasibility for the OURA ring was exceptionally high. Participants successfully wore the device and generated valid sleep data for 100% of the expected study days. In fact, many participants continued to wear the device and generate data beyond the 42-day minimum.

Subjective Diary Compliance: Feasibility for the daily REDCap sleep diary was significantly lower. Participants completed the daily diary on exactly 89.4% of the expected days (563 out of 630 expected days).

This yielded a final matched dataset of 544 nights in which a participant had both a valid OURA reading and a completed sleep diary for the same night.

Overall agreement between subjective sleep diaries and objective wearable sensors ranged from poor to moderate depending on the specific sleep parameter. The summary of concordance, including mean biases, Pearson correlations, Intraclass Correlation Coefficients (ICC), and limits of agreement, is detailed in Table 4. Participants systematically overestimated their Total Sleep Time by an average of + 31.3 minutes per night. The primary driver of the TST overestimation is a significant underestimation of intra-night wakefulness. The OURA ring detected an average of 106.2 minutes of awake time, whereas participants reported only 41.4 minutes (a bias of -64.9 minutes). Older adults are highly accurate at estimating their sleep latency, with a mean bias of only − 1.7 minutes.

Table 4
Comparison of Subjective (Diary) vs. Objective (OURA) Sleep Metrics (N = 544 matched nights)

Metric	**Subjective Diary(Mean ± SD)**	**Objective OURA(Mean ± SD)**	**Pearson r**	**ICC**	**Mean Bias(Diary - OURA)**	**Limits of Agreement(95% CI)**
TIB (min)	495.6 ± 64.9	511.4 ± 75.1	0.54	0.52	-15.9	[-148.8, 117.1]
TST(min)	436.5 ± 71.5	405.2 ± 55.3	0.57	0.5	31.3	[-86.7, 149.3]
Sleep Latency (min)	17.6 ± 20.0	19.4 ± 15.7	0.2	0.19	-1.7	[-46.6, 43.1]
Awake Time (min)	41.4 ± 41.9	106.3 ± 46.5	0.28	0.13	-64.9	[-169.0, 39.3]
Sleep Efficiency (%)	88.2 ± 9.6	79.6 ± 6.8	0.3	0.18	8.6	[-10.9, 28.0]

The Bland-Altman analysis (Fig. 4) revealed prominent systematic biases consistent with sleep state misperception in older adults. Participants demonstrated moderate absolute agreement when estimating broad duration metrics, including Time in Bed (ICC = 0.52) and Total Sleep Time (ICC = 0.50). On average, participants slightly underestimated their Time in Bed by roughly 16 minutes (mean bias = -15.9 minutes) and exhibited a negligible underestimation of Sleep Latency (-1.7 minutes).

Despite this moderate baseline tracking of their schedule, the perception of active, high-quality sleep demonstrated severe intra-night discrepancy. Compared to the objective sensor, participants systematically overestimated their Total Sleep Time by an average of 31.3 minutes per night (95% LOA [-86.7, 149.3]). This overestimation of total sleep was directly driven by an inability to accurately perceive nighttime awakenings; participants underreported their Wake After Sleep Onset (WASO) by a substantial 64.9 minutes (ICC = 0.13). Consequently, participants systematically overestimated their overall Sleep Efficiency by 8.6%.

At the conclusion of the 6-week observation period, an exit survey was administered to assess the acceptability and perceived burden of the study protocol. Overall acceptability was exceptionally high: 100% of respondents found the study schedule manageable and would recommend participation to others. The continuous wearable protocol was well-tolerated, with 100% of participants rating their experience with the OURA ring as "Very Positive" or "Somewhat Positive." While participants praised the ease of charging and the usefulness of the objective data feedback, a minority reported minor physical discomforts, such as difficulty removing the ring due to finger swelling. Regarding the clinical assessments, 100% of participants reported feeling physically safe during all dynamic balance and mobility testing. Acceptability was high for the motor assessments (100% positive) but slightly lower for the cognitive assessments (83% positive), with participants noting that cognitive tasks were more susceptible to environmental distractions. Notably, 100% of participants indicated that participating in the study improved their personal understanding of their health, with several users explicitly noting that the objective data helped them recognize the direct correlation between a poor night of sleep and their subsequent balance performance.

4. DISCUSSION

4.1. Summary of the objectives of the study and key findings

This study evaluated the feasibility of using a continuous wearable sensor (Oura Ring) to untangle the complex, dynamic relationships among sleep, motor balance, and cognition in older adults. By assessing both habitual sleep traits and acute temporal deviations, we identified a critical divergence: cognitive performance relies heavily on long-term sleep stability, whereas motor function is acutely vulnerable to short-term sleep fluctuations. Furthermore, our feasibility analysis revealed significant subjective misperception of sleep state, underscoring the clinical necessity of continuous objective monitoring.

Habitual Sleep Patterns and Cognitive Reserve

Our evaluation of between-person differences established that habitual sleep consistency acts as a protective baseline for cognitive health. Specifically, habitual TST was strongly correlated with better Working Memory (r = 0.59), whereas high sleep fragmentation (quantified by the standard deviation of Wake After Sleep Onset) was strongly negatively correlated with memory performance (r = -0.49). The

findings on the relationship between TST and working memory align with the literature[65, 66], suggesting that maintaining regular sleep and wake times could protect working memory. While a meta-analysis focused on sleep-memory associations in older adults did not find WASO to be significantly associated with poorer memory[67], sleep fragmentation has been consistently associated with worse cognition and increased risk of dementia[68–70]. Our moderate negative association between sleep fragmentation and memorizing performance further supports these findings. Participants who consistently maintained longer, uninterrupted sleep across the six-week period demonstrated superior cognitive baseline functioning[66]. Furthermore, habitual REM sleep duration showed positive associations with both Working Memory and Processing Speed, consistent with other studies[71,72]. Because REM sleep is deeply implicated in the consolidation of emotional and procedural memory[73,74], our data suggest that REM preservation is a critical subcomponent driving the cognitive benefits of overall sleep duration.

Acute Motor Vulnerability vs. Cognitive Resilience

The temporal sensitivity analysis uncovered a striking dissociation between cognitive and motor resilience. Consistent with the hypothesis of cognitive reserve, cognitive domains maintained state-level stability across all rolling sleep windows, suggesting that established neural networks can temporarily buffer against short-term sleep disruptions. This resilience contrasts with some broader epidemiological literature[75,76], which often reports an inverted U-shaped relationship in which both excessively short and long sleep durations are associated with cognitive impairment. However, our findings are consistent with recent meta-analytic evidence indicating that objectively measured sleep duration is not significantly associated with short-term cognitive performance in healthy older adults[6,66,77]. This supports the premise that while long-term sleep habits shape cognitive baselines, short-term temporal deviations do not immediately degrade cognitive function.

Motor systems, however, exhibited a dose-dependent, cumulative vulnerability to sleep extension. Dynamic balance (MiniBEST) significantly deteriorated after just 6 days of "Long Sleep," followed by widespread motor slowing on the Timed Up and Go (TUG) by day 10. This acute motor vulnerability aligns with established experimental evidence[61] that sleep disruption negatively affects postural control by altering the metabolism of the thalamus, cerebellum, and basal ganglia, which are critical brain regions for sensory integration and motor coordination.

Crucially, "Long Sleep" periods in older adults are rarely physiological luxuries; they often serve as proxies for underlying fatigue, minor systemic inflammation, or days with markedly reduced physical activity. Because extended time in bed inherently reduces movement, this 10-day accumulation of sedentary or lethargic behavior likely degrades the rapid neuromuscular and vestibular firing required for sharp postural control. This interpretation is supported by recent literature[78] highlighting that inadequate or excessive sleep is closely linked to insufficient physical activity in older adults, and that fatigue and neuroinflammation[79] are key mediators between sleep variations and impaired postural control[6]. This

leads directly to the specific threshold declines in dynamic balance and functional mobility captured by our models.

Feasibility and Sleep State Misperception

The integration of consumer-grade wearables proved highly feasible and well tolerated by our cohort, with participants achieving 100% compliance over the expected 42-day monitoring period. This could be associated with participants' relatively high education levels, high motivation, our on-board workshop and training, and customized guidelines for using the OURA rings[80–82]. However, comparing this continuous objective data with subjective sleep diaries revealed severe intra-night sleep-state misperception. Overall, absolute agreement between the two modalities ranged from poor to moderate.

While older adults were reasonably accurate at estimating broad schedule parameters, such as Time in Bed (ICC = 0.52) and Sleep Latency (Mean Bias = -1.7 minutes), they grossly misjudged their active sleep architecture. On average, participants severely underreported their Wake After Sleep Onset (WASO) by 64.9 minutes. This failure to perceive nighttime awakenings directly led to an artificial overestimation of Total Sleep Time by 31.3 minutes per night and an inflated subjective Sleep Efficiency. This highlights a major limitation in geriatric clinical assessments: relying strictly on self-reported sleep diaries may cause clinicians to miss severe sleep fragmentation that disrupts an older adult's restorative rest.

The high ecological validity and feasibility demonstrated in Aim 3 are further supported by the participants' exit feedback. Despite the demanding nature of a 6-week protocol requiring continuous wearable use, daily diaries, and weekly in-person clinical testing, participant acceptability was overwhelmingly positive. The Oura Ring was well tolerated, though researchers deploying wearables in geriatric populations should carefully monitor ring sizing to account for common issues such as peripheral edema or joint swelling. Interestingly, the discrepancy between subjective diary compliance (89%) and objective wearable compliance (100%) highlights that passive, low-friction digital health tools are significantly less burdensome for older adults than active, recall-based daily diaries. Finally, the exit interviews highlighted a secondary clinical benefit of this protocol: providing older adults with transparent, objective sleep feedback actively improved their health literacy, helping them self-identify the relationship between their sleep fragmentation and daytime fall risk.

Strengths and Limitations

A primary strength of this study is its high ecological validity. By transitioning from traditional laboratory polysomnography to continuous, home-based digital health monitoring with the Oura Ring, we captured naturalistic sleep variations in participants' real-world environments. This methodology not only achieved exceptional user compliance (100% of expected days) but also enabled the continuous, longitudinal tracking necessary to identify specific temporal thresholds of functional decline.

Furthermore, the implementation of a dual-modality sleep assessment that pairs objective sensor data with subjective daily sleep diaries represents a significant methodological strength. This parallel tracking

enabled direct quantification of sleep state misperception among older adults, demonstrating that objective monitoring is essential to capture the severe intra-night sleep fragmentation that self-reports routinely miss.

The analytical framework utilized to evaluate within-person acute deviations also strengthens the clinical relevance of the findings. Moving beyond simple cross-sectional correlations, the use of rolling 10-day and 14-day temporal windows permitted a granular, dose-response analysis of sleep fluctuations. This dynamic mapping successfully isolated the differing resilience thresholds of cognitive and motor systems over time.

Finally, the study thoroughly evaluates the "Sleep-Balance-Cognition Triad" using a comprehensive, highly validated suite of functional assessments. By combining robust clinical motor evaluations (e.g., MiniBESTest, TUG with dual-task costs) with the demographically corrected NIH Toolbox Cognition Battery, the protocol captured sensitive, state-level changes across both physical and neurobehavioral domains while controlling baseline sociodemographic differences.

This study is subject to several limitations that warrant consideration. First, the study relies on a small sample size (N = 15 completers) and a highly homogeneous cohort (100% White, highly educated), which restricts the broad generalizability of the findings to more diverse or socioeconomically varied populations. Additionally, the demanding nature of a 6-week continuous monitoring protocol inherently introduces a healthy-volunteer bias, as the participants who successfully completed the study may possess higher baseline physical and cognitive resilience than the general older adult population.

Second, the rigorous weekly testing cadence likely introduced a practice effect. Participants may have improved on cognitive tasks simply due to repeated exposure, which could have masked subtle, acute sleep-related cognitive declines. While age-adjusted Change-Sensitive Scores from the NIH Toolbox were used to mitigate this, future longitudinal protocols should implement randomized alternate forms of cognitive assessments to further minimize learning bias. Similar familiarization effects may have also stabilized dynamic motor performance scores during the initial weeks of testing.

Third, while the ecological validity of using a consumer-grade wearable is a primary strength, it introduces specific measurement constraints. The Oura Ring is well validated against gold-standard polysomnography (PSG) for broad metrics such as TST and WASO, but its ability to accurately classify specific sleep architecture (such as distinguishing among light, deep, and REM sleep) via photoplethysmography and actigraphy is inherently less precise than traditional EEG-based assessments. Furthermore, the wearable does not explicitly export a discrete count of nocturnal awakenings, which prevented a direct 1:1 comparison with the subjective awakening counts reported in the sleep diaries.

Finally, the observational design of this study precludes definitive causal conclusions. Although our temporal sensitivity analysis, using rolling 10- and 14-day trailing windows, strongly suggests a directional relationship in which extended sleep and lethargy preceded functional motor decline,

unmeasured time-varying confounders could still influence these results. Daily fluctuations in acute joint pain, mood disturbances, or daytime physical exertion could simultaneously drive both an increased need for time in bed and a subsequent slowing of functional mobility on the TUG test. Future studies should incorporate daily tracking of pain, mood, and objective physical activity volumes as time-varying covariates to further isolate the independent mechanisms of the sleep-balance-cognition triad.

CONCLUSION

This study demonstrates that the relationship between sleep and functional capacity in older adults is complex and temporally dynamic, heavily dependent on whether one evaluates long-term habits or short-term fluctuations. While habitual sleep consistency appears to protect cognitive reserve, particularly working memory, acute sleep deviations and accumulated lethargy unmask rapid, dose-dependent vulnerabilities in dynamic balance and functional mobility. Crucially, the significant discrepancies observed between self-reported diaries and objective sensor data underscore the unreliability of subjective recall for capturing intra-night sleep fragmentation and sleep state misperception.

By detecting the hidden sleep disruptions that traditional self-reports routinely miss, this research validates the integration of continuous commercial wearables into geriatric care. These findings establish that passive, home-based digital health sensors are not merely consumer wellness gadgets but highly feasible and clinically actionable tools. Ultimately, deploying these technologies enables continuous remote monitoring of functional status, paving the way for scalable, decentralized early-warning systems for fall risk and mobility decline among community-dwelling older adults.

Declarations

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

AI-Assisted Writing Declaration

During manuscript preparation, the authors used Google Gemini 3.1 Pro to assist with language editing, restructuring selected paragraphs, and reference-formatting support, and used Grammarly (V.1.2.290.1948) for grammar and style suggestions. The authors reviewed, verified, and revised all AI-assisted content and take full responsibility for the accuracy and integrity of the manuscript.

Funding

This work was supported by the University of Florida (UF) College of Public Health and Health Professions (PHHP) Research Innovation Fund (2024), the UF PHHP Artificial Intelligence PhD Fellowship (2023–2024), and the UF PHHP Muthard Research Award (2025). The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement by, UF PHHP.

Author Contribution

H.W.: Conceptualization, funding acquisition, investigation, methodology, resources, supervision, writing—original draft, writing—review and editing. Y.L.: Data collection, Data curation, formal analysis, methodology, validation, and visualization. R.C.: Data collection, data curation, data cleaning, methodology. Y.C.S.: Data collection, data curation, data cleaning, methodology. J.A.: Participant recruitment, study coordination, wearable-device setup and logistics, data collection, and data management and quality control. I.T.: Participant recruitment, study coordination, wearable-device setup and logistics, data collection, and data management and quality control. S.S.: REDCap setup, data management, and quality control. I.G.: Participant recruitment, study coordination, wearable-device setup and logistics, data collection, and data management and quality control. E.S.: Participant recruitment, study coordination, wearable-device setup and logistics, data collection, and data management and quality control. S.P.: Study coordination, wearable-device setup and logistics, data collection, data management, and quality control. L.S.: Participant recruitment, study coordination, and data collection and management. All authors reviewed the manuscript.

Acknowledgement

The authors would like to acknowledge the participants and Emelie Selberg's assistance with data collection. We would also like to thank the coordination and support from staff members working in the UF Smart House.

Data Availability

The participant level wearable data are not publicly available because they contain potentially sensitive longitudinal information and are subject to institutional review board and data-use restrictions. Deidentified analytic data and/or code are available from the corresponding author upon reasonable request and approval of an appropriate data-use agreement.

References

1. Lowry, K. A., Vallejo, A. N. & Studenski, S. A. Successful aging as a continuum of functional independence: lessons from physical disability models of aging. *Aging Dis.* **3** (1), 5–15 (2012).

2. Lo, J. C., Loh, K. K., Zheng, H., Sim, S. K. Y. & Chee, M. W. L. Sleep duration and age-related changes in brain structure and cognitive performance. *Sleep* **37** (7), 1171–1178 (2014).
3. Agmon, M. & Shochat, T. The link between sleep and gait among community dwelling adults. Sleep Medicine [Internet]. ;100((Agmon M.; Shochat T.) University of Haifa, Nursing, Haifa, Israel):S10–1. (2022). Available from: https://www.embase.com/search/results?subaction=viewrecord&id=L2019357624&from=export
4. Robillard, R., Prince, F., Filipini, D. & Carrier, J. Aging worsens the effects of sleep deprivation on postural control. *PLoS One*. **6** (12), e28731 (2011).
5. Siddarth, P. et al. Sleep quality, neurocognitive performance, and memory self-appraisal in middle-aged and older adults with memory complaints. Int Psychogeriatr [Internet]. Jul [cited 2026 Jan 25];33(7):703–13. (2021). Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8004546/
6. Shin, Y. C., Shah, S., Sejpal, H. M. & Wang, H. Mapping the relationships among sleep, motor balance, and cognition in older adults: a systematic scoping review. *Front. Hum. Neurosci.* **19**, 1575155. 10.3389/fnhum.2025.1575155 (2025).
7. Leger, D., Mutti, C., Rouen, A. & Parrino, L. Polysomnography in Transition: Reassessing Its Role in the Future of Sleep Medicine. *J. Sleep. Res.* **34** (6), e70217. 10.1111/jsr.70217 (2025).
8. Yoon, H. & Choi, S. H. Technologies for sleep monitoring at home: wearables and nearables. *Biomed. Eng. Lett.* **13** (3), 313–327. 10.1007/s13534-023-00305-8 (2023).
9. Mathunjwa, B. M., Kor, R. Y. J., Ngarnkuekool, W. & Hsu, Y. L. A Comprehensive Review of Home Sleep Monitoring Technologies: Smartphone Apps, Smartwatches, and Smart Mattresses. *Sens. (Basel)*. **25** (6), 1771. 10.3390/s25061771 (2025).
10. Oura Health Oy. Introducing the New Oura Ring Generation 3. The Pulse Blog. Published October 26, 2021. Accessed August 20. (2026). https://ouraring.com/blog/oura-generation2-vs-generation3/
11. de Zambotti, M., Rosas, L., Colrain, I. M. & Baker, F. C. The Sleep of the Ring: Comparison of the ŌURA Sleep Tracker Against Polysomnography. *Behav. Sleep. Med.* **17** (2), 124–136 (2019).
12. Svensson, T., Madhawa, K., Nt, H., Chung, U. I. & Svensson, A. K. Validity and reliability of the Oura Ring Generation 3 (Gen3) with Oura sleep staging algorithm 2.0 (OSSA 2.0) when compared to multi-night ambulatory polysomnography: A validation study of 96 participants and 421,045 epochs. *Sleep. Med.* **115**, 251–263. 10.1016/j.sleep.2024.01.020 (2024).
13. Khan, S. et al. The Oura Ring Versus Medical-Grade Sleep Studies: A Systematic Review and Meta-Analysis. *OTO Open.* **9** (4), e70181. 10.1002/oto2.70181 (2025).
14. Bliwise, D. L., King, A. C. & Harris, R. B. Habitual sleep durations and health in a 50–65 year old population. *J. Clin. Epidemiol.* **47** (1), 35–41. 10.1016/0895-4356(94)90031-0 (1994).
15. Pienaar, P. R. et al. Associations between habitual sleep characteristics and cardiometabolic disease risk in corporate executives. *Sleep. Health*. **10** (5), 550–557. 10.1016/j.sleh.2024.07.007 (2024).
16. Babson, K. A., Trainor, C. D., Feldner, M. T. & Blumenthal, H. A test of the effects of acute sleep deprivation on general and specific self-reported anxiety and depressive symptoms: an experimental

extension. *J. Behav. Ther. Exp. Psychiatry*. **41** (3), 297–303. 10.1016/j.jbtep.2010.02.008 (2010).

17. Li, X. et al. Acute sleep rebound following sleep restriction is associated with reduced mortality risk. *Nat. Commun.* **17**, 3820. https://doi.org/10.1038/s41467-026-72461-1 (2026).
18. Piltz, S. H., Diniz Behn, C. G. & Booth, V. Habitual sleep duration affects recovery from acute sleep deprivation: A modeling study. *J. Theor. Biol.* **504**, 110401. 10.1016/j.jtbi.2020.110401 (2020).
19. Hertzog, M. A. Considerations in determining sample size for pilot studies. *Res. Nurs. Health*. **31** (2), 180–191. 10.1002/nur.20247 (2008).
20. Goparaju, B., de Palma, G. & Bianchi, M. T. Naturalistic sleep tracking in a longitudinal cohort: Uncertainty and bias in short duration sampling. *PLoS One*. **20** (11), e0334950. 10.1371/journal.pone.0334950 (2025).
21. Norful, A. A. et al. Exploring longitudinal physiologic stress measurement and sleep quality interventions to improve psychological well-being in nurses: a pilot study. *Health Psychol. Behav. Med.* **13** (1), 2503376. 10.1080/21642850.2025.2503376 (2025).
22. National Sleep Foundation. National Sleep Foundation Sleep Diary [Internet]. National Sleep Foundation. 2021 [cited 2024 Jul 8]. Available from: https://www.thensf.org/nsf-sleep-diary/
23. Tamaki, M. et al. First-night effect reduces the beneficial effects of sleep on visual plasticity and modifies the underlying neurochemical processes. *Sci. Rep.* **14**, 14388. 10.1038/s41598-024-64091-8 (2024).
24. Carney, C. E. et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep* **35** (2), 287–302. 10.5665/sleep.1642 (2012).
25. Borba, D. A. et al. How many days are needed for a reliable assessment by the Sleep Diary? Sleep Sci. *2020 Jan-Mar* ;**13**(1):49–53. 10.5935/1984-0063.20190131
26. Silverberg, J. I. et al. Psychometric Validation of the Subject Sleep Diary in Patients with Moderate-to-Severe Atopic Dermatitis. *Dermatol. Ther. (Heidelb)*. **15** (4), 963–995. 10.1007/s13555-025-01385-3 (2025).
27. Berg, K. O., Maki, B. E., Williams, J. I., Holliday, P. J. & Wood-Dauphinee, S. L. Clinical and laboratory measures of postural balance in an elderly population. *Arch. Phys. Med. Rehabil*. **73** (11), 1073–1080 (1992).
28. Berg, K. O., Wood-Dauphinee, S. L., Williams, J. I. & Maki, B. Measuring balance in the elderly: validation of an instrument. *Can. J. Public. Health*. **83** (Suppl 2), S7–11 (1992 Jul-Aug).
29. Major, M. J., Fatone, S. & Roth, E. J. Validity and reliability of the Berg Balance Scale for community-dwelling persons with lower-limb amputation. *Arch. Phys. Med. Rehabil*. **94** (11), 2194–2202. 10.1016/j.apmr.2013.07.002 (2013).
30. Downs, S., Marquez, J. & Chiarelli, P. The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. *J. Physiother*. **59** (2), 93–99. 10.1016/S1836-9553(13)70161-9 (2013).

31. Wong, C. K. Interrater reliability of the Berg Balance Scale when used by clinicians of various experience levels to assess people with lower limb amputations. *Phys. Ther.* **94** (3), 371–378. 10.2522/ptj.20130182 (2014).
32. Downs, S. The Berg Balance Scale. *J. Physiother*. **61** (1), 46. 10.1016/j.jphys.2014.10.002 (2015).
33. Park, S. H. & Lee, Y. S. The Diagnostic Accuracy of the Berg Balance Scale in Predicting Falls. *West. J. Nurs. Res.* **39** (11), 1502–1525. 10.1177/0193945916670894 (2017).
34. King, L. & Horak, F. On the mini-BESTest: scoring and the reporting of total scores. *Phys. Ther.* **93** (4), 571–575. 10.2522/ptj.2013.93.4.571 (2013).
35. Godi, M. et al. Comparison of reliability, validity, and responsiveness of the mini-BESTest and Berg Balance Scale in patients with balance disorders. *Phys. Ther.* **93** (2), 158–167. 10.2522/ptj.20120171 (2013).
36. Yingyongyudha, A., Saengsirisuwan, V., Panichaporn, W. & Boonsinsukh, R. The Mini-Balance Evaluation Systems Test (Mini-BESTest) Demonstrates Higher Accuracy in Identifying Older Adult Participants With History of Falls Than Do the BESTest, Berg Balance Scale, or Timed Up and Go Test. *J. Geriatr. Phys. Ther.* **39** (2), 64–70 (2016 Apr-Jun).
37. Phyu, S. N., Peungsuwan, P., Puntumetakul, R. & Chatchawan, U. Reliability and Validity of Mini-Balance Evaluation System Test in Type 2 Diabetic Patients with Peripheral Neuropathy. *Int. J. Environ. Res. Public. Health*. **19** (11), 6944. 10.3390/ijerph19116944 (2022).
38. Mathias, S., Nayak, U. S. & Isaacs, B. Balance in elderly patients: the get-up and go test. *Arch. Phys. Med. Rehabil*. **67** (6), 387–389 (1986).
39. Morris, S., Morris, M. E. & Iansek, R. Reliability of measurements obtained with the Timed Up & Go test in people with Parkinson disease. *Phys. Ther.* **81** (2), 810–818. 10.1093/ptj/81.2.810 (2001).
40. Lin, M. R. et al. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. *J. Am. Geriatr. Soc.* **52** (8), 1343–1348. 10.1111/j.1532-5415.2004.52366.x (2004).
41. Ng, S. S. & Hui-Chan, C. W. The timed up & go test: its reliability and association with lower-limb impairments and locomotor capacities in people with chronic stroke. *Arch. Phys. Med. Rehabil*. **86** (8), 1641–1647. 10.1016/j.apmr.2005.01.011 (2005).
42. Christopher, A., Kraft, E., Olenick, H., Kiesling, R. & Doty, A. The reliability and validity of the Timed Up and Go as a clinical tool in individuals with and without disabilities across a lifespan: a systematic review. *Disabil. Rehabil*. **43** (13), 1799–1813. 10.1080/09638288.2019.1682066 (2021).
43. HealthMeasures. Cognition measures: NIH Toolbox Cognition Batteries. (2017).
44. Zelazo, P. D. et al. NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. *J. Int. Neuropsychol. Soc.* **20** (6), 620–629. 10.1017/S1355617714000472 (2014).
45. Akshoomoff, N. et al. The NIH Toolbox Cognition Battery: results from a large normative developmental sample (PING). *Neuropsychology* **28** (1), 1–10. 10.1037/neu0000001 (2014).

46. Casaletto, K. B. et al. Demographically Corrected Normative Standards for the English Version of the NIH Toolbox Cognition Battery. *J. Int. Neuropsychol. Soc.* **21** (5), 378–391 (2015).
47. Horne, J. A. & Ostberg, O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *Int. J. Chronobiol*. **4** (2), 97–110 (1976).
48. Rhee, M. K., Taillard, J., Philip, P. & Bioulac, B. Morningness/eveningness and the need for sleep. *J. Sleep. Res.* **8** (4), 291–295. 10.1046/j.1365-2869.1999.00176.x (1999).
49. Lee, H. J., Rex, K. M. & Kripke, D. F. Evaluation of two circadian rhythm questionnaires for screening for the delayed sleep phase disorder. *Psychiatry Investig*. **9** (3), 236–244. 10.4306/pi.2012.9.3.236 (2012).
50. Panjeh, S. et al. What are we measuring with the morningness-eveningness questionnaire? Exploratory factor analysis across four samples from two countries. *Chronobiol Int.* **38** (2), 234–247. 10.1080/07420528.2020.1815758 (2021).
51. Mackenzie, L., Byles, J. & Higginbotham, N. Designing the Home Falls and Accidents Screening Tool (HOME FAST): selecting the items. *Br. J. Occup. Ther.* **63** (6), 260–269 (2000).
52. Mackenzie, L., Byles, J. & Higginbotham, N. Professional perceptions about home safety: cross-national validation of the Home Falls and Accidents Screening Tool (HOME FAST). *J. Allied Health*. **31** (1), 22–28 (2002 Spring).
53. Mackenzie, L., Byles, J. & Higginbotham, N. Reliability of the Home Falls and Accidents Screening Tool (HOME FAST) for identifying older people at increased risk of falls. *Disabil. Rehabil*. **24** (5), 266–274. 10.1080/09638280110087089 (2002).
54. Mackenzie, L., Byles, J. & D'Este, C. Longitudinal study of the Home Falls and Accidents Screening Tool in identifying older people at increased risk of falls. *Australas J. Ageing*. **28** (2), 64–69. 10.1111/j.1741-6612.2009.00361.x (2009).
55. Peel, C. et al. Assessing mobility in older adults: the UAB Study of Aging Life-Space Assessment. *Phys. Ther.* **85** (10), 1008–1119 (2005).
56. Kuspinar, A. et al. Assessing the measurement properties of life-space mobility measures in community-dwelling older adults: a systematic review. *Age Ageing*. **52** (Suppl 4), iv86–iv99. 10.1093/ageing/afad119 (2023).
57. Craig, C. L. et al. International physical activity questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* **35** (8), 1381–1395. 10.1249/01.MSS.0000078924.61453.FB (2003).
58. Cuschieri, S. The CONSORT statement. *Saudi J. Anaesth.* **13** (Suppl 1), S27–S30. 10.4103/sja.SJA_559_18 (2019).
59. Harris, P. A. et al. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inf.* **42** (2), 377–381. 10.1016/j.jbi.2008.08.010 (2009).
60. Aickin, M. & Gensler, H. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. *Am. J. Public. Health*. **86** (5), 726–728. 10.2105/ajph.86.5.726 (1996).

61. Umemura, G. S., Furtado, F., Santos, F. C., Gonçalves, B. S. B. & Forner-Cordero, A. Is Balance Control Affected by Sleep Deprivation? A Systematic Review of the Impact of Sleep on the Control of Balance. *Front. Neurosci.* **16**, 779086. 10.3389/fnins.2022.779086 (2022).
62. Fischer, D., Klerman, E. B. & Phillips, A. J. K. Measuring sleep regularity: theoretical properties and practical usage of existing metrics. *Sleep* **44** (10), zsab103. 10.1093/sleep/zsab103 (2021).
63. Windred, D. P. et al. Sleep regularity is a stronger predictor of mortality risk than sleep duration: A prospective cohort study. *Sleep* **47** (1), zsad253. 10.1093/sleep/zsad253 (2024).
64. Intra-individual variability. in behavior: links to brain structure, neurotransmission and neuronal activity: Trends in Neurosciences [Internet]. [cited 2026 Jan 25]. Available from: https://www.cell.com/trends/neurosciences/abstract/S0166-2236(06)00125-1
65. Newbury, C. R., Crowley, R., Rastle, K. & Tamminen, J. Sleep deprivation and memory: Meta-analytic reviews of studies on sleep deprivation before and after learning. *Psychol. Bull.* **147** (11), 1215–1240. 10.1037/bul0000348 (2021).
66. Qin, S., Leong, R. L. F., Ong, J. L. & Chee, M. W. L. Associations between objectively measured sleep parameters and cognition in healthy older adults: A meta-analysis. *Sleep. Med. Rev.* **67**, 101734. 10.1016/j.smrv.2022.101734 (2023).
67. Hokett, E., Arunmozhi, A., Campbell, J., Verhaeghen, P. & Duarte, A. A systematic review and meta-analysis of individual differences in naturalistic sleep quality and episodic memory performance in young and older adults. *Neurosci. Biobehav Rev.* **127**, 675–688. 10.1016/j.neubiorev.2021.05.010 (2021).
68. Kaneshwaran, K. et al. Sleep fragmentation, microglial aging, and cognitive impairment in adults with and without Alzheimer's dementia. *Sci. Adv.* **5** (12), eaax7331. 10.1126/sciadv.aax7331 (2019).
69. Lim, A. S., Kowgier, M., Yu, L., Buchman, A. S. & Bennett, D. A. Sleep Fragmentation and the Risk of Incident Alzheimer's Disease and Cognitive Decline in Older Persons. *Sleep* **36** (7), 1027–1032. 10.5665/sleep.2802 (2013).
70. Manousakis, J. E., Scovelle, A. J., Rajaratnam, S. M. W., Naismith, S. L. & Anderson, C. Advanced Circadian Timing and Sleep Fragmentation Differentially Impact on Memory Complaint Subtype in Subjective Cognitive Decline. *J. Alzheimers Dis.* **66** (2), 565–577. 10.3233/JAD-180612 (2018).
71. Scullin, M. K. & Gao, C. Dynamic Contributions of Slow Wave Sleep and REM Sleep to Cognitive Longevity. *Curr. Sleep. Med. Rep.* **4** (4), 284–293. 10.1007/s40675-018-0131-6 (2018).
72. Suh, S. W. et al. Short Average Duration of NREM/REM Cycle Is Related to Cognitive Decline in an Elderly Cohort: An Exploratory Investigation. *J. Alzheimers Dis.* **70** (4), 1123–1132. 10.3233/JAD-190399 (2019).
73. Cairney, S. A., Durrant, S. J., Power, R. & Lewis, P. A. Complementary roles of slow-wave sleep and rapid eye movement sleep in emotional memory consolidation. *Cereb. Cortex*. **25** (6), 1565–1575. 10.1093/cercor/bht349 (2015).
74. Yuksel, C. et al. Both slow wave and rapid eye movement sleep contribute to emotional memory consolidation. *Commun. Biol.* **8**, 485. 10.1038/s42003-025-07868-5 (2025).

75. Ma, Y. et al. Association Between Sleep Duration and Cognitive Decline. *JAMA Netw. Open.* **3** (9), e2013573. 10.1001/jamanetworkopen.2020.13573 (2020).
76. Guo, G. et al. Associations Between Sleep Duration and Cognitive Function Among Older Adults: Cross-Sectional Study. *JMIR Hum. Factors*. **12**, e72886. 10.2196/72886 (2025).
77. Hyndych, A., El-Abassi, R. & Mader, E. C. Jr The Role of Sleep and the Effects of Sleep Loss on Cognitive, Affective, and Behavioral Processes. *Cureus* **17** (5), e84232. 10.7759/cureus.84232 (2025).
78. Alnawwar, M. A. et al. The Effect of Physical Activity on Sleep Quality and Sleep Disorder: A Systematic Review. *Cureus* **15** (8), e43595. 10.7759/cureus.43595 (2023).
79. Mullington, J. M., Simpson, N. S., Meier-Ewert, H. K. & Haack, M. Sleep loss and inflammation. *Best Pract. Res. Clin. Endocrinol. Metab.* **24** (5), 775–784. 10.1016/j.beem.2010.08.014 (2010).
80. Meinders, M. J. et al. Optimizing wrist-worn wearable compliance with insights from two Parkinson's disease cohort studies. *NPJ Parkinsons Dis.* **11**, 152. 10.1038/s41531-025-01016-w (2025).
81. Pergolizzi, J. V. Jr, LeQuang, J. A. K., El-Tallawy, S. N. & Varrassi, G. What Clinicians Should Tell Patients About Wearable Devices and Data Privacy: A Narrative *Rev. Cureus* ;**17**(3):e81167. doi: 10.7759/cureus.81167 (2025).
82. Ding, H. et al. Assessment of Wearable Device Adherence for Monitoring Physical Activity in Older Adults: Pilot Cohort Study. *JMIR Aging*. **7**, e60209. 10.2196/60209 (2024).
83. Supplementary Table, S. Functional Change during 14-Day Short vs. Long Sleep Windows (Ref: Normal Sleep).

Figures

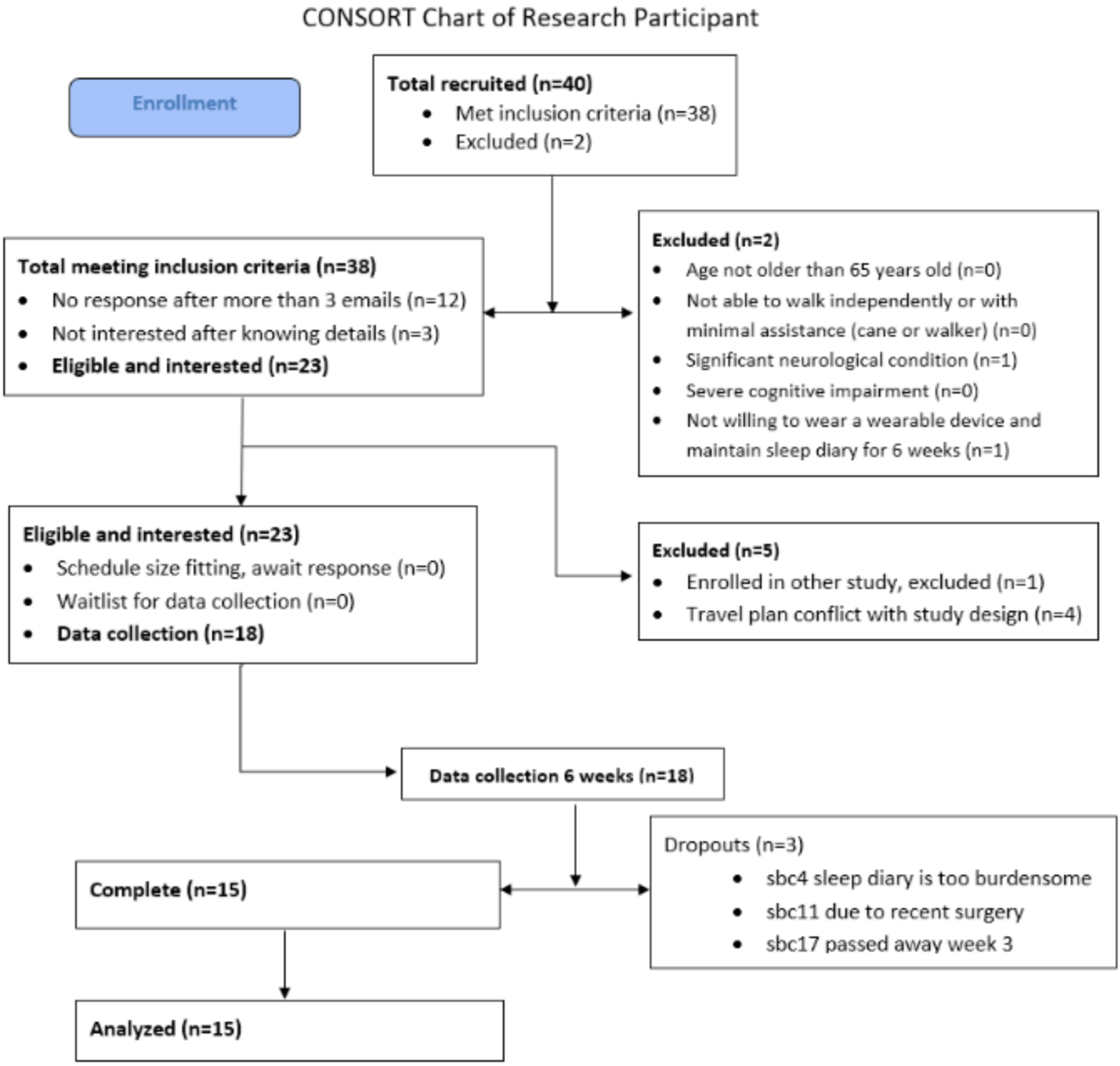

Figure 1

CONSORT Chart.

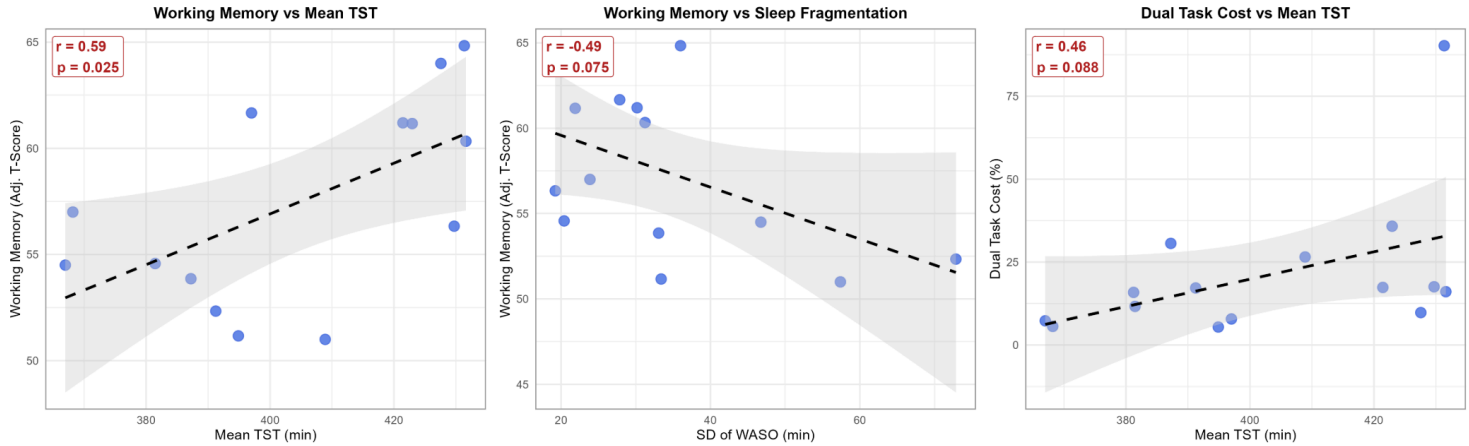

Figure 2

Scatter plots show associations between sleep characteristics and functional Outcomes.

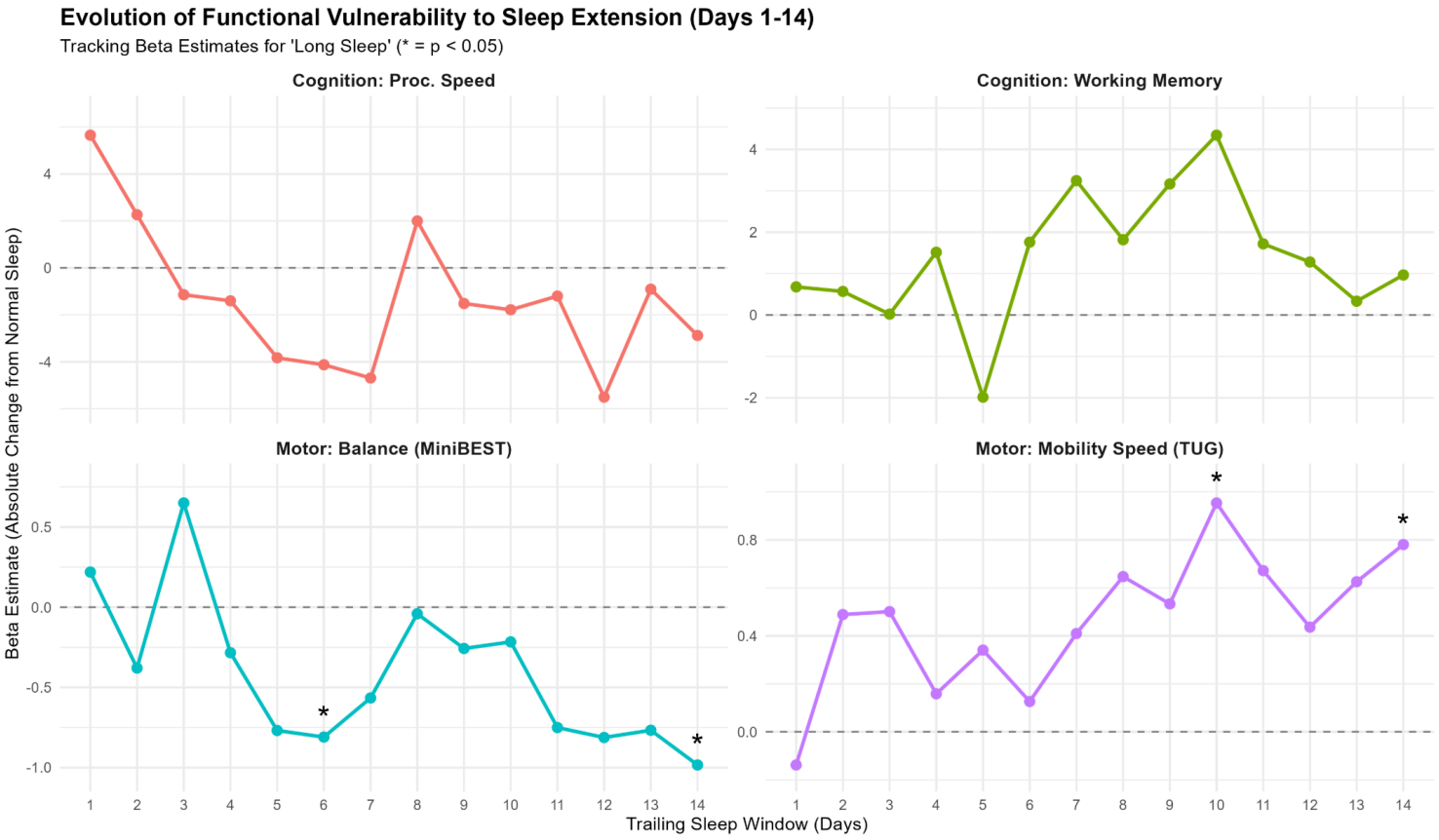

Figure 3

The temporal sensitivity analysis revealed that while cognitive function remains highly stable under state-level sleep fluctuations, motor control is heavily dose-dependent. Extended sleep durations begin to significantly compromise dynamic balance (MiniBEST) as early as day 6, with widespread motor impairment, including significantly slowed functional mobility (TUG), emerging by day 10 and persisting through day 14. This timeline indicates that prolonged sleep extension, likely a proxy for chronic lethargy

or reduced physical activity, might be a progressive, cumulative stressor on the neuromuscular systems that govern fall risk in older adults.

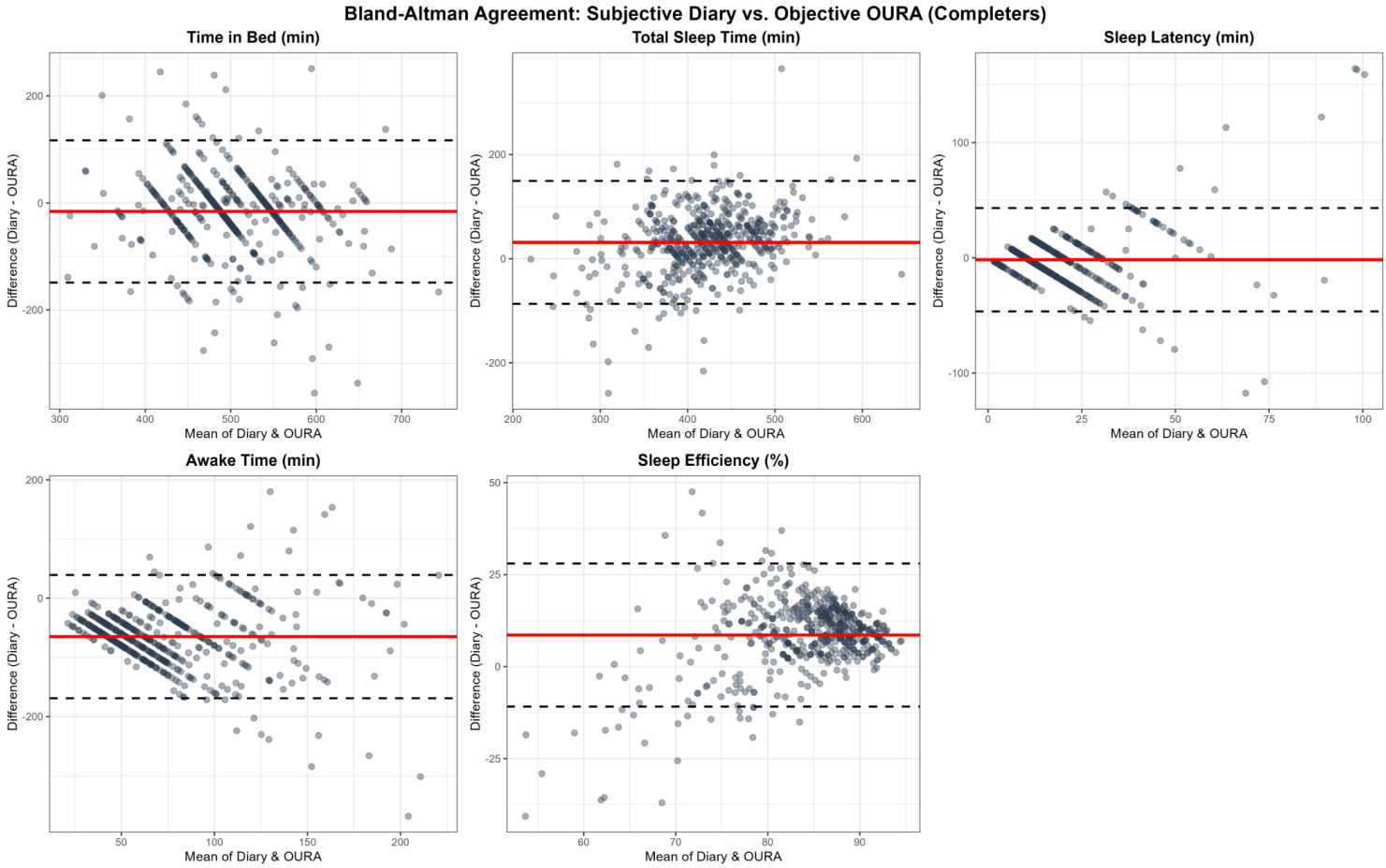

Figure 4

Bland-Altman plots assessing the agreement between subjective daily sleep diaries and objective continuous wearable monitoring (OURA ring) across 544 matched nights. The x-axis displays the mean of the subjective and objective measurements, while the y-axis displays the absolute difference (Subjective Diary – Objective OURA). Solid red lines denote the mean systematic bias, and dashed black lines indicate the upper and lower 95% Limits of Agreement (± 1.96 SD). The panels illustrate that while older adults reasonably estimated broader schedule metrics such as Time in Bed, they exhibited significant misperception of sleep states during the night. Notably, participants systematically overestimated their Total Sleep Time (Mean Bias = +31.3 min) and severely underestimated their Awake Time/Wake After Sleep Onset (Mean Bias = -64.9 min), leading to an artificially inflated subjective Sleep Efficiency.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SBCCONSORT2025checklist.docx