

Research Articles | Behavioral/Cognitive

## Prefrontal Glx reduction during the sleep transition is associated with sleep onset latency

<https://doi.org/10.1523/JNEUROSCI.0372-26.2026>

Received: 27 February 2026

Revised: 2 September 2026

Accepted: 9 September 2026

Copyright © 2026 the authors

---

*This Early Release article has been peer reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.*

**Alerts:** Sign up at [www.jneurosci.org/alerts](http://www.jneurosci.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

**TITLE PAGE**

**Title:** Prefrontal Glx reduction during the sleep transition is associated with sleep onset latency  
(13/50 words)

**Abbreviated title:** Prefrontal glutamate and human sleep onset (42/50 characters)

**Authors:** Takashi Yamada<sup>1,2</sup>, Aaron Cochrane<sup>1</sup>, Theodore LaBonte-Clark<sup>1</sup>, Takeo Watanabe<sup>1</sup>,  
Yuka Sasaki<sup>1</sup>, \*

1: Cognitive and Psychological Sciences, Brown University, Providence, RI 02912, USA

2: Division of Developmental Neuroscience, United Graduate School of Child Development, The  
University of Osaka, Suita 565-0871, Osaka, Japan

\*Corresponding author:

Yuka Sasaki, Ph.D. (E-mail: yuka\_sasaki@brown.edu)

**Number of pages:** 36

**Number of figures:** 9

**Number of tables:** 2

**Abstract word count:** 241

**Introduction word count:** 530

**Discussion word count:** 1440

**Acknowledgements:** This work was supported by the NIH (R01EY031705, R01EY019466,  
R01EY027841), and NSF (NSF-BSF: 2241417).

**Conflicts of interest:** The authors declare no competing financial interests.

JNeurosci Accepted Manuscript

## ABSTRACT

Understanding the neural mechanisms that facilitate the transition from wakefulness to sleep is a fundamental goal in neuroscience, yet the neurochemical dynamics underlying human sleep onset remain poorly understood. Here, we investigated whether changes in inhibitory  $\gamma$ -aminobutyric acid (GABA) and excitatory glutamate plus glutamine (Glx) concentrations in the medial prefrontal cortex (mPFC) are associated with sleep initiation. We leveraged the first-night effect, a transient sleep disruption induced by a novel environment, to manipulate sleep quality within participants. Twenty healthy young adults (12 females) completed two afternoon nap sessions involving simultaneous polysomnography and magnetic resonance spectroscopy of the mPFC, with sleep onset latency (SOL) serving as the primary index of sleep quality. SOL was significantly longer in the first session than in the second, confirming the first-night effect. Glx concentrations decreased during the transition from wakefulness to sleep in the second session, whereas this reduction was absent during disturbed sleep in the first session. Across sessions, greater reductions in Glx were associated with shorter SOL, and mediation analysis indicated that Glx dynamics statistically accounted for session-dependent differences in SOL. In contrast, GABA concentrations did not significantly change from wakefulness to non-rapid eye movement sleep in either session and were not associated with SOL. These findings suggest that downregulation of Glx in the mPFC is closely associated with the facilitation of human sleep onset and that a failure to reduce excitatory neurometabolite levels, rather than altered inhibition, may contribute to acute sleep-onset disturbances. (241/250 words)

## SIGNIFICANCE STATEMENT

The sleep transition is a fundamental neurobiological process, yet the neurometabolic dynamics underlying this process remain elusive. By combining polysomnography with functional magnetic resonance spectroscopy while utilizing the first-night effect to manipulate sleep latency, we demonstrate that successful sleep initiation is closely associated with the downregulation of excitatory Glx, rather than changes in inhibitory GABA, in the medial prefrontal cortical region. Our results suggest that the reduction of excitatory neurometabolites, rather than the recruitment of GABAergic inhibition, is closely linked to a smooth transition into sleep. These findings offer a new neurometabolic perspective for understanding sleep disturbances and may help identify potential therapeutic targets for insomnia. (106/120 words)

## MAIN TEXT

### Introduction

The transition from wakefulness to sleep is a fundamental neurobiological shift. Central to this process is sleep onset latency (SOL), which is defined as the time required to initiate sleep after the intention to do so, such as at lights-off. SOL provides a direct readout of the wake-to-sleep transition, and its prolongation is a core symptom of sleep-onset disturbances (Morin and Benca, 2012). Although existing therapies can relieve symptoms, they have not yet yielded a detailed account of the neurochemical processes that drive sleep initiation in the human brain (Watson et al., 2010). A clearer mechanistic framework for SOL and for the transition into deeper sleep stages would inform basic models of human sleep regulation and improve our understanding of sleep disorders (Saper et al., 2005; Weber and Dan, 2016; Scammell et al., 2017).

Sleep onset is thought to occur when wake-promoting circuits are downregulated. The reticular ascending arousal system and related wake-promoting pathways comprise many subcortical nuclei with widespread projections to the cortex (Moruzzi and Magoun, 1949; Saper et al., 2005; Saper and Fuller, 2017; Jones, 2020; Sulaman et al., 2023). Within these systems, glutamate provides major excitatory drive and gamma-aminobutyric acid (GABA) provides major inhibitory control. Animal studies indicate that both excitatory and inhibitory systems are central to these state transitions (Saper et al., 2005; Saper and Fuller, 2017; Jones, 2020; Sulaman et al., 2023).

However, translating these circuit models into a mechanistic account of human sleep initiation faces two major challenges. First, while insufficient inhibition is frequently invoked as a cause of sleep-onset difficulty (Sherin et al., 1996; Lu et al., 2000; Winkelman et al., 2008; Saper et al., 2010; Watson et al., 2010), the field lacks direct quantification of how changes in inhibition and

excitation each contribute to SOL in humans. Second, SOL is difficult to manipulate experimentally without confounds such as medication exposure or comorbidity.

We approached these challenges by measuring neurochemical dynamics throughout the sleep onset process. We combined continuous magnetic resonance spectroscopy (MRS) with polysomnography (PSG) (Tamaki et al., 2020; Tamaki et al., 2021) to obtain repeated measures of excitatory (Glx) and inhibitory (GABA) neurotransmitter concentrations across wakefulness, the transition to sleep, and subsequent NREM sleep stages. We focused on the medial prefrontal cortical region (mPFC), which receives projections from subcortical wake-promoting systems (Saper et al., 2005; Saper and Fuller, 2017; Jones, 2020; Zhong et al., 2023) and has been linked to the regulation of consciousness and arousal (Nofzinger et al., 2004; Mashour et al., 2022).

To vary SOL while avoiding pharmacological confounds, we leveraged the first-night effect (FNE), a transient sleep disturbance that occurs in a novel environment (Agnew et al., 1966; Carskadon and Dement, 1975; Tamaki et al., 2016). Using a two-session within-subject design, we compared mPFC neurotransmitter concentrations between the first session (transiently disturbed sleep) and the second session (adapted sleep). We tested the hypothesis that longer SOL is accompanied by an altered excitatory-inhibitory profile during the transition to sleep, rather than differences in baseline concentrations during wakefulness. Specifically, we examined whether delayed sleep onset is associated with a reduction in GABA concentrations, an elevation in Glx levels, or a combination of both as individuals transition from wakefulness into sleep. (530/650 words)

## **Methods and materials**

### ***Participants***

In total, 22 young adults (aged 18 to 30 years) were recruited for this study. Two participants were excluded because they failed to complete the sleep experiments, resulting in a final sample of 20 participants (12 females, mean age  $20.9 \pm 2.9$  years). To minimize age-related influences on sleep structure (Ehlers and Kupfer, 1997; Mander et al., 2017), we strictly limited the age range to 18 to 30 years. All participants provided written informed consent, and the study protocol was approved by the Institutional Review Board of Brown University.

We conducted a rigorous screening process consistent with our established methods to ensure participants were healthy and maintained regular sleep-wake patterns (Tamaki et al., 2020). Based on self-reports, we excluded individuals with any history of physical, neurological, or psychiatric illnesses, those currently taking medication, or those with suspected sleep disorders. Eligibility was further verified using a detailed sleep-wake habits questionnaire (Tamaki et al., 2024) and the Munich Chronotype Questionnaire (MCTQ) (Roenneberg et al., 2003). Regularity was defined as an average sleep duration of six to nine hours, with no more than two hours of variation in bedtime or wake-up time between workdays and free days.

Following informed consent, further assessments were conducted using the Morningness-Eveningness Questionnaire (MEQ) (Horne and Ostberg, 1976) and the Pittsburgh Sleep Quality Index (PSQI) (Buysse et al., 1989). These assessments confirmed that no participants exhibited extreme morningness or eveningness, with MEQ scores ranging from 42 to 69. Additionally, all participants were classified as good sleepers, as indicated by PSQI scores of five or less.

### ***Experimental design***

Participants underwent two afternoon nap sessions scheduled between 2:30 and 4:00 PM, approximately one week apart, following established sleep study protocols (Tamaki et al., 2020; Tamaki et al., 2021). Regular sleep-wake schedules were monitored via sleep logs and

actigraphy (GT9X-BT, ActiGraph) for 3–7 days prior to each session. Nap sessions were rescheduled if sleep timing varied by more than two hours between workdays and free days. Participants were instructed to refrain from alcohol, intense exercise, and napping on the day before each session, and from caffeine on the day of the sessions.

On the day of the experiments, participants were asked to sleep inside a 3T magnetic resonance imaging (MRI) scanner (Siemens Prisma). Since the experimental environment was novel to the participants, Session 1 (Day 1) was expected to induce the first-night effect (FNE) (Agnew et al., 1966; Zammit et al., 2009; Tamaki and Sasaki, 2019; Tamaki et al., 2024). We used this inherent response to a novel environment as an experimental manipulation to vary sleep quality within subjects. We anticipated that participants would exhibit a transient reduction in sleep stability during Session 1, whereas Session 2 (Day 2) would reflect an adapted state with improved sleep quality.

In each session, electrodes for polysomnography (PSG) were attached to the participants outside the MRI scanner prior to the nap. PSG recording commenced once participants lay down inside the scanner, followed by the start of MRI acquisition.

### ***Polysomnography and sleep scoring***

During the nap sessions, participants underwent PSG recording simultaneously with MRS. MRI-compatible EEG caps (Brain Products GmbH) recorded EEG, electrooculography (EOG), electromyography (EMG), and electrocardiography (ECG). A 32-channel cap was used for 15 participants, and a 25-channel cap was used for 5 participants. The 32-electrode cap consisted of 23 scalp EEG electrodes, six for EOG, two for EMG, and one for ECG. Horizontal EOGs were obtained from two bipolar electrodes positioned at the outer canthi of both eyes, and vertical

EOGs from four electrodes positioned above and below the left and right eyes. The 25-electrode configuration included 18 scalp EEG electrodes, four EOG electrodes, two EMG electrodes, and one ECG electrode. For this setup, EOGs were captured with bipolar electrodes at the outer canthi of each eye and two additional electrodes above and below the left eye.

Across both cap configurations, EMG was recorded bipolarly from the chin, while ECG was obtained from a site approximately two finger-widths left of the spine at the level of the inferior scapulae. EEG and ECG electrodes were referenced to Fz, with the ground electrode positioned at AFp4. Electrode impedances were maintained at 10 k $\Omega$  for EEG and ECG, and < 15 k $\Omega$  for EOG and EMG. Data acquisition was performed using an MRI-compatible amplifier (BrainAmp MR, Brain Products) and recording software (BrainVision Recorder, Brain Products) at a sampling rate of 5,000 Hz. Scanner and ballistocardiogram artifacts were removed using BrainVision Analyzer 2 (Brain Products). Subsequently, the EEG data were re-referenced to the mastoids (TP9 and TP10).

While we used two EEG caps with different configurations, all electrodes present in the 25-channel cap were also present in the 32-channel cap. To ensure consistency across the two configurations, visual sleep scoring was performed using only the derivations common to both caps. Specifically, sleep was scored from the frontal, central, and occipital EEG electrodes (F3, F4, C3, C4, O1, and O2), each re-referenced to the mastoids (TP9 and TP10) in 30-s epochs according to the standard criteria defined in the American Academy of Sleep Medicine (AASM) Scoring Manual (Berry et al., 2017). The horizontal and vertical EOGs, both common to the two montages, together with the bipolar chin EMG, were also used for scoring. Electrodes present only in the 32-channel cap (namely T7, T8, Cz, Pz, Oz, and the right-eye vertical EOG channels) were excluded from sleep scoring.

Stages were classified as follows: wakefulness (stage W), NREM stages (N1, N2, and N3), and REM sleep (stage REM). Sleep onset was defined by the first appearance of stage N2, consistent with previous research (Tamaki et al., 2020; Tamaki et al., 2021). This criterion was selected to ensure a stable and definitive transition into sleep, thereby minimizing potential contamination from wakefulness-related cortical dynamics, such as residual alpha activity, which frequently persists during stage N1. The time interval from lights-off to sleep onset was defined as sleep onset latency (SOL). Sleep structure data, including SOL, are shown in **Table 1** below.

### ***MRS Data Acquisition and Analysis***

MRS was conducted on a 3T Siemens Prisma scanner with a 64-channel head coil. To ensure participant comfort and minimize head movement during MRI measurements, heads were securely positioned using cushions and gauze. Additional support was provided via back and knee cushions, and blankets were used to maintain body warmth and facilitate sleep onset.

Before MRS acquisition, high-resolution T1-weighted images were obtained using a magnetization-prepared rapid gradient echo (MP-RAGE) sequence (TR = 1.9 s, TE = 3.02 ms, FA = 9 degrees, voxel size = 1 x 1 x 1 mm<sup>3</sup>, 256 slices, no slice gap) to guide precise voxel placement (2.5 x 2.5 x 2.5 cm<sup>3</sup>). The voxel was manually placed in the medial prefrontal cortical region (mPFC) using a standardized, anatomical landmark-based procedure. Placement was performed in real time on the three orthogonal planes (sagittal, coronal, and axial) displayed on the scanner console. On the sagittal midline image, the voxel was centered on the interhemispheric fissure and positioned anterior to the genu of the corpus callosum; it was then adjusted on the coronal and axial planes to maintain left-right symmetry while minimizing inclusion of CSF, skull, frontal sinus, and white matter. The same placement rule was applied on both scan days. Representative voxel locations are shown in sagittal and transverse views in **Figure 1**.

For each session, an MRS voxel mask was generated in individual anatomical space using Gannet v 3.5.3 (Edden et al., 2014), which derives the mask from the MRS voxel geometry and the corresponding T1-weighted image. To quantify between-session overlap, the Day 2 structural image was co-registered to the corresponding Day 1 structural image using SPM12 (Friston, 2007), and the resulting transformation was applied to the Day 2 MRS voxel mask. Overlap was then calculated as the percentage of overlapping image voxels between the Day 1 mask and the transformed Day 2 mask, relative to the total MRS voxel volume (15,625 image voxels, i.e., a  $25 \times 25 \times 25 \text{ mm}^3$  voxel sampled at  $1 \text{ mm}^3$  resolution). The voxel position demonstrated high consistency between sessions, with an overall volume overlap of  $86.6 \pm 1.8\%$  (mean  $\pm$  SEM).

After automated shimming, MRS data were acquired using the MEGA-PRESS sequence. A brief initial scan (TR = 1.25 s, TE = 68 ms, FA = 90 degrees, 32 averages, 2:05 min) confirmed data quality (Cramer-Rao lower bounds [CRLB] < 25) before a longer (~90 min total) scan comprising nine consecutive acquisitions (TR = 1.25 s, TE = 68 ms, FA = 90 degrees, 240 averages, 10:05 min each). Concentrations of Glx, GABA, and N-acetylaspartate (NAA; internal reference) were quantified using LCModel (Version 6.3-1P) (Provencher, 1993) between 1.95 and 4.2 ppm, following water scaling and eddy-current correction (**Supplementary Figure S1**). We used NAA as a control metabolite to normalize Glx and GABA levels, following established procedures (Mullins et al., 2014; Tamaki et al., 2020).

Following temporal co-registration of MRS and PSG data, we quantified the changes in Glx and GABA concentrations during the transition from wakefulness to sleep. Average concentrations during wakefulness, which was defined as the period from lights-off to sleep onset, including any subsequent wakeful periods, served as the baseline. We then assessed the relative changes in these neurotransmitter levels during NREM sleep compared to the baseline. Specifically, the percentage change was calculated by subtracting the wakefulness baseline from the NREM

concentration, dividing the result by the wakefulness baseline, and multiplying by 100, where the wakefulness baseline was set to 0%.

While MRS data were recorded during REM sleep, they were excluded from the primary analysis because REM sleep occurred in only a limited number of participants across both sessions (**Supplementary Table S1**). To maintain a consistent within-subject comparison with the full sample ( $n = 20$ ), we focused our primary analysis on the transition from wakefulness to NREM sleep, as all participants reached the NREM stage in both sessions.

Additionally, we calculated average concentrations for N1, N2, and N3 stages separately to examine the effect of sleep depth on Glx and GABA concentrations. For these stage-specific analyses, data were included only for participants who reached the respective stages in both sessions with successful co-registered MRS-PSG data ( $n = 17$  for N1,  $n = 20$  for N2, and  $n = 6$  for N3; **Supplementary Table S2**).

Notably, we measured Glx (a composite of glutamate and glutamine) rather than isolated glutamate, as linear-combination modeling may not reliably distinguish glutamate from glutamine and GABA in MEGA-PRESS data (Pasanta et al., 2023). Glx is considered a reliable measure (Pasanta et al., 2023), and glutamate is estimated to contribute approximately 85% of the Glx signal (Floyer-Lea et al., 2006). Thus, while results are reported as Glx, they likely predominantly reflect glutamate levels.

### ***Quality control of the MRS data***

First, we used the CRLB to exclude low-quality MRS data using LCModel. The CRLBs indicate the reliability of the quantification of Glx, GABA, and NAA as measures of fitting errors. A threshold

of 25% was used to reject low-quality MRS results (Dydak et al., 2011; Andronesi et al., 2016; Frank et al., 2022). This threshold rejected an average of 1.3% of all MRS data from two sleep sessions, with 1.5% rejected on Day 1 and 1.1% on Day 2.

Next, to comprehensively evaluate and compare the stability and quality of the remaining MRS data between the two sleep sessions across different sleep stages, we calculated four additional parameters for the spectra: shim values, frequency drift, NAA linewidth, and NAA signal-to-noise ratio (SNR). Shim values were defined as the scanner-derived shim settings obtained after automated shimming, reflecting the adjustments applied to optimize B0 field homogeneity within the voxel. Frequency drift, which signifies the instability of the B0 magnetic field, was calculated as the standard deviation of the average-by-average frequency estimates returned by Gannet v3.5.3 (Edden et al., 2014). NAA SNR was determined based on the NAA peak in the edit-OFF spectra using Gannet's standard algorithm, specifically by dividing the maximum amplitude of the modeled NAA peak by the standard deviation of the residuals (noise) in a signal-free baseline region. Additionally, the NAA linewidth was quantified based on the LCModel outputs using an in-house MATLAB script to provide a further measure of signal resolution. To ensure transparency and reproducibility, all acquisition and processing parameters were documented (**Supplementary Table S3**) in accordance with the Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS) consensus guidelines (Lin et al., 2021).

### ***Temporal co-registration of MRS and PSG data***

We followed the established procedure for temporal co-registration of MRS and PSG data (Tamaki et al., 2020; Tamaki et al., 2021). PSG and MRS data were time-aligned in 10-minute blocks to obtain the average neurotransmitter concentration for wakefulness, NREM sleep and

REM sleep. If one sleep stage comprised 80% or more of the twenty 30-s PSG epochs in a 10-min block, the block was labeled with that stage. If no stage was predominant in one block, it was split into five 2-min sub-blocks (4 PSG epochs each). Each sub-block was labeled with the predominant stage (>50%), with ties resolved by a hierarchy (wakefulness > NREM sleep > REM sleep). The same rules were applied for NREM sleep sub-stages (N1, N2, N3), extending the hierarchy to wakefulness > N1 > N2 > N3 > REM sleep (Tamaki et al., 2020; Tamaki et al., 2021). In addition, when we analyzed MRS levels for stages N1, N2, and N3 separately, we applied the same rules but expanded to five categories, prioritizing wakefulness followed by N1, N2, N3, and REM sleep. Due to this temporal co-registration and the hierarchical staging rules, certain sleep stages were prioritized (**Supplementary Table S2**); consequently, the sample sizes in **Supplementary Tables S1 and S2** do not necessarily match. The total duration of analyzed segments for each sleep stage per participant is provided in **Supplementary Table S4**.

### **Statistical analyses**

The normality of all continuous variables was systematically assessed using the Shapiro-Wilk test. Because the majority of our variables deviated significantly from a normal distribution, we primarily conducted comparisons using nonparametric methods, specifically Wilcoxon signed-rank tests and Spearman rank correlations. Furthermore, to rigorously evaluate whether MRS data quality (NAA linewidth and SNR) was affected by sleep stages or measurement order across the nine sessions per participant, we employed linear mixed models (LMMs) with participants treated as a random effect. LMMs were selected because of their robustness against deviations from normality and their capability to handle repeated-measures data structures. IBM SPSS Statistics (version 31.0) was used for these primary analyses and LMMs.

For the mediation analysis (Baron and Kenny, 1986), we examined whether changes in Glx or GABA levels mediated the effect of session (Day 1 vs. Day 2) on SOL. We applied the Yeo-Johnson transformation (Yeo and Johnson, 2000) to SOL, Glx levels, and GABA levels to improve data normality, as the linear regressions used in the mediation analysis assume normality. We defined the "Day" factor as the independent variable, the Glx (or GABA) level as the mediator, and SOL as the dependent variable, with age and sex included as covariates. We employed a structural equation modeling approach (Iacobucci et al., 2007) to incorporate repeated measures, with robust standard errors clustered by participant to account for within-participant correlations. The significance of the mediation effect was tested using a Monte Carlo simulation with 500 replications, as this method does not require the assumption of normality, unlike traditional approaches such as the Sobel test (Baron and Kenny, 1986). We used Stata Standard Edition version 19.0 (StataCorp.) for the mediation analysis.

Statistical analyses were performed using a two-tailed  $\alpha = 0.05$  significance threshold. For all statistical tests, effect sizes were reported alongside p-values. Multiple comparison corrections were not applied.

## Results

### ***Confirmation of the FNE***

As expected, the FNE was observed as a significant sleep disturbance on Day 1 relative to Day 2. Specifically, SOL was significantly longer on Day 1 compared to Day 2 (Wilcoxon signed-rank test:  $Z = -2.67$ ,  $p = 0.008$ , effect size  $r = -0.60$ ; **Figure 2**). This confirms the presence of the FNE in our study, consistent with prior reports (Agnew et al., 1966; Tamaki et al., 2016; Ding et al., 2022; Tamaki et al., 2024).

The sleep architecture for each session is summarized in **Table 1**. Beyond the difference in SOL, no other sleep parameters showed statistically significant differences between Day 1 and Day 2 (all  $p > 0.05$ ). This indicates that while the initiation of sleep was delayed on the first night, the overall structure of sleep, once established, remained comparable across both sessions.

### ***Glx Levels During Sleep Stages***

Glx levels during NREM sleep showed markedly different patterns between sessions, expressed as a percentage change from the wakefulness baseline. On Day 1, mean Glx during NREM sleep was not significantly different from the wakefulness baseline (one-sample Wilcoxon signed rank test:  $Z = -1.05$ ,  $p = 0.296$ ,  $n = 20$ ,  $r = 0.23$ ; **Figure 3**). However, on Day 2, Glx during NREM sleep was significantly lower than the wakefulness baseline ( $Z = -3.51$ ,  $p < 0.001$ ,  $n = 20$ ,  $r = 0.78$ ; **Figure 3**). The reduction in Glx during NREM sleep was significantly greater on Day 2 than on Day 1 (related-samples Wilcoxon signed rank test:  $Z = -2.43$ ,  $p = 0.015$ ,  $n = 20$ ,  $r = 0.54$ ; **Figure 3**).

Considering individual NREM sub-stages (**Figure 4**), on Day 2, Glx significantly decreased from wakefulness during stage N2 ( $Z = -2.20$ ,  $p = 0.028$ ,  $n = 20$ ,  $r = 0.49$ ) and N3 ( $Z = -2.67$ ,  $p = 0.008$ ,  $n = 9$ ,  $r = 0.89$ ), but not during N1 ( $Z = -1.05$ ,  $p = 0.296$ ,  $n = 20$ ,  $r = 0.23$ ). In contrast, there were no significant Glx changes in any NREM stage on Day 1 ( $Z = 0.54$ ,  $p = 0.586$ ,  $n = 17$ ,  $r = 0.13$  for N1;  $Z = -0.19$ ,  $p = 0.856$ ,  $n = 20$ ,  $r = 0.04$  for N2;  $Z = -1.51$ ,  $p = 0.131$ ,  $n = 11$ ,  $r = 0.46$  for N3). The Glx reductions in N2 and N3 were significantly greater on Day 2 than on Day 1 ( $Z = -2.17$ ,  $p = 0.030$ ,  $n = 20$ ,  $r = 0.48$  for N2;  $Z = -1.99$ ,  $p = 0.046$ ,  $n = 6$ ,  $r = 0.81$  for N3). These results indicate that when sleep was improved on Day 2, Glx levels in the mPFC dropped substantially during sleep (especially in deeper NREM stages), whereas no such reduction occurred during the disturbed sleep on Day 1.

### ***Glx and Sleep Onset***

Importantly, we observed that a greater Glx reduction during sleep was associated with a shorter SOL. Across individuals, NREM Glx levels were significantly correlated with SOL on both Day 1 (Spearman  $\rho = 0.60$ ,  $p = 0.005$ ,  $n = 20$ ) and Day 2 (Spearman  $\rho = 0.53$ ,  $p = 0.017$ ,  $n = 20$ ; see **Supplementary Figures 2a, b**). Specifically, participants who exhibited a larger reduction in mPFC Glx during NREM sleep tended to fall asleep faster.

To further explore the role of Glx changes in the prolonged sleep onset observed on Day 1, we performed a mediation analysis (Baron and Kenny, 1986; Iacobucci et al., 2007) as a statistical framework to evaluate the relative pathways among variables. In this model, "Day" (Day 1 vs. Day 2) was entered as the independent variable, the Glx level as the mediator, and SOL as the outcome variable, with age and sex included as covariates. This analysis indicated that the variance in Glx levels statistically accounted for the relationship between the experimental day and SOL. While Day had a significant total effect on SOL because SOL was longer on Day 1, the direct effect of Day on SOL was no longer significant when Glx was included in the model. Instead, there was a significant indirect pathway via Glx, which accounted for approximately 51% of the total Day effect (**Figure 5**). Although the inherent limitations of concurrent measurements preclude definitive causal interpretations, these statistical results suggest that the longer SOL on Day 1 relative to Day 2 was statistically associated with the smaller or absent reduction in Glx during sleep on Day 1.

### ***GABA Levels During Sleep Stages***

In contrast to Glx, GABA did not show a consistent pattern of change during sleep. Overall NREM GABA levels were not significantly different from the wakefulness baseline on either Day 1 or Day 2 (one-sample Wilcoxon signed rank tests:  $Z = 1.61$ ,  $p = 0.108$ ,  $n = 20$ ,  $r = 0.36$  for Day 1;  $Z = -$

1.31,  $p = 0.191$ ,  $n = 20$ ,  $r = 0.29$  for Day 2; **Figure 6**), and there was no significant difference in NREM GABA between Day 1 and Day 2 ( $Z = -1.49$ ,  $p = 0.135$ ,  $n = 20$ ,  $r = 0.33$ ).

However, analysis by each NREM sub-stage suggested a transient GABA elevation during early NREM sub-stages on Day 1. Specifically, on Day 1 (**Figure 7**), GABA was significantly above baseline during N1 ( $p = 0.004$ ,  $n = 17$ ,  $r = 0.71$ ) and showed a similar trend during N2 ( $p = 0.057$ ,  $n = 20$ ,  $r = 0.43$ ), whereas no such increase was observed during N3 ( $p = 0.248$ ,  $n = 11$ ,  $r = 0.35$ ). On Day 2, GABA showed no significant changes from baseline at any NREM stage ( $p = 0.526$ ,  $n = 20$ ,  $r = 0.14$  for N1;  $p = 0.737$ ,  $n = 20$ ,  $r = 0.08$  for N2;  $p = 0.139$ ,  $n = 9$ ,  $r = 0.49$  for N3). Direct comparisons between days revealed that GABA was lower on Day 2 than on Day 1 across all stages, demonstrating marginal trends in N1 ( $p = 0.076$ ,  $n = 17$ ,  $r = 0.43$ ) and N2 ( $p = 0.057$ ,  $n = 20$ ,  $r = 0.43$ ), and a significant reduction in N3 ( $p = 0.028$ ,  $n = 6$ ,  $r = 0.90$ ). Thus, during higher-quality sleep (Day 2), GABA remained essentially unchanged from wakefulness levels, whereas during poorer sleep (Day 1), GABA levels tended to increase in the early NREM stages.

### **GABA and Sleep Onset**

Unlike Glx, the GABA level during NREM sleep showed no significant correlation with SOL on either day (Day 1: Spearman  $\rho = -0.04$ ,  $p = 0.870$ ; Day 2: Spearman  $\rho = 0.12$ ,  $p = 0.607$ ; see **Supplementary Figure S3**). A mediation analysis consistently confirmed that GABA changes did not account for the effect of Day on SOL. The indirect effect of Day via GABA was not significant ( $p = 0.967$ ), contributing less than 2% of the total Day effect (**Figure 8**). Consequently, the direct effect of Day on SOL remained significant after controlling for GABA. Thus, GABA dynamics in the mPFC during NREM sleep appeared to have little influence on SOL in our study.

### **Control Analyses**

To ensure the robustness of our findings, we performed several control analyses (see below). These comprehensive control analyses definitively demonstrate that the observed dynamics in Glx and GABA reflect true neurochemical shifts rather than variations in baseline metabolite levels or data acquisition artifacts.

First, we confirmed the macro-level stability of NAA. NAA concentrations were stable, showing no significant differences between wakefulness and NREM sleep, or between Day 1 and Day 2 (**Figure 9a**). Furthermore, baseline Glx and GABA concentrations during wakefulness were not significantly different between Day 1 and Day 2 (**Figure 9b, c**).

Second, we assessed the quality and stability of the MRS data globally. The macro-level spectral quality parameters (shim values, Glx %SD, GABA %SD, NAA linewidth, and frequency drift) did not differ between Day 1 and Day 2 (**Table 2**; all  $p > 0.05$ ).

Third, we conducted further highly granular quality control analyses at the individual segment level. Specifically, we used linear mixed models (LMMs) to confirm that data quality (NAA linewidth and NAA SNR) was unaffected by sleep stages or measurement order (see below).

#### *Assessment of MRS Data Quality Across Sleep Stages and Day*

As noted in the Methods section, our scanning protocol was designed with 10-minute acquisition blocks. Metabolite spectra were reconstructed and quantified at the 2-minute scale, and the reliability of each 2-minute estimate was verified individually using the CRLB provided by LCModel; only windows with a CRLB (%SD) below 25% were retained. Quantification of individual metabolites is thus safeguarded at the single-window level, independently of the number of averages. Spectral quality metrics such as SNR and linewidth, by contrast, are not accompanied by an equivalent per-fit reliability estimate and become less stable at low average counts. This

constraint is reinforced by the architecture of standard MRS toolkits (including Gannet): because the integrated raw sequence data (the Siemens .dat file containing the reference scans) are stored as a single indivisible package for each complete run, these toolkits derive SNR and linewidth from the full acquisition rather than from arbitrarily subdivided 2-minute windows. We therefore evaluated SNR and linewidth at the level of the complete 10-minute acquisitions, where a fixed and larger number of averages yields robust and comparable quality metrics across sleep states. As a complementary approach to evaluate the potential influence of sleep states under these methodological conditions, we categorized these 10-minute MRS segments into three distinct types based on the sleep architecture within each block: (1) wakefulness-only blocks, (2) NREM sleep-only blocks, and (3) mixed blocks containing multiple sleep stages. We then constructed LMMs with participants treated as a random effect to evaluate whether spectral quality systematically varied across these sleep-stage categories or differed between experimental days (Day 1 vs. Day 2), as well as their interaction.

When NAA SNR was examined as the dependent variable, the mean raw values for Day 1 were  $45.4 \pm 13.4$  for the wakefulness-only blocks,  $46.8 \pm 15.3$  for the NREM sleep-only blocks, and  $37.2 \pm 14.4$  for the mixed blocks. For Day 2, the values were  $43.7 \pm 17.5$ ,  $44.4 \pm 18.1$ , and  $45.6 \pm 20.9$ , respectively. The LMM revealed no significant main effects of the sleep-stage categories ( $F(2, 51.55) = 0.60$ ,  $p = 0.555$ ) or the experimental days ( $F(1, 64.16) = 0.19$ ,  $p = 0.666$ ), and no significant interaction was observed ( $F(2, 51.55) = 1.10$ ,  $p = 0.340$ ).

Next, we evaluated NAA linewidth as the dependent variable. Given the inherent positive skewness of the raw resolution metrics, linewidth values were log-transformed prior to the LMM analysis to meet the statistical assumptions. However, to provide a clear and intuitive

representation of the absolute spectral quality, the raw (untransformed) linewidth values across the conditions are reported here. The mean raw NAA linewidths for Day 1 were  $10.6 \pm 2.2$  Hz for the wakefulness-only blocks,  $11.0 \pm 2.4$  Hz for the NREM sleep-only blocks, and  $10.8 \pm 2.2$  Hz for the mixed blocks. For Day 2, the values were  $10.7 \pm 3.4$  Hz,  $10.4 \pm 2.0$  Hz, and  $10.6 \pm 2.5$  Hz, respectively, reflecting consistently high spectral resolution throughout the entire experiment.

The LMM on the log-transformed NAA linewidth revealed no significant main effects of the sleep-stage categories ( $F(2, 51.21) = 0.11$ ,  $p = 0.900$ ) or the experimental days ( $F(1, 41.75) = 0.22$ ,  $p = 0.643$ ), and no significant interaction between them ( $F(2, 51.21) = 0.18$ ,  $p = 0.836$ ). Although the model's residuals for linewidth exhibited a deviation from strict normality even after log-transformation, we retained this linear mixed-model approach because LMMs are documented to be highly robust against violations of normality in structured multi-level datasets with sufficient sample sizes.

#### *Assessment of MRS Data Quality by measurement order and Day*

To further test whether the time-course or participant fatigue affected spectral quality over repeated measurements, we constructed another LMM evaluating the fixed effects of measurement order (from 1st to 9th), experimental days (Day 1 vs. Day 2), and their interaction. When NAA SNR was examined as the dependent variable, the LMM revealed no significant main effects of the measurement order ( $F(8, 65.73) = 1.33$ ,  $p = 0.243$ ) or the experimental days ( $F(1, 296.09) = 0.03$ ,  $p = 0.857$ ), and no significant interaction between them ( $F(8, 65.73) = 0.80$ ,  $p = 0.603$ ). These results demonstrate that spectral quality remained highly stable and robust against potential confounding factors such as scanner drift or participant movement due to fatigue over the course of the extended recording sessions.

In a similar way, when log-transformed NAA linewidth was entered as the dependent variable, the LMM revealed no significant main effects of measurement order ( $F(8, 68.20) = 0.41, p = 0.910$ ) or the experimental days ( $F(1, 292.64) = 2.93, p = 0.088$ ), and no significant interaction was observed ( $F(8, 68.20) = 0.05, p = 1.000$ ). These results demonstrate that the spectral resolution (measured by linewidth) remained highly consistent across different sleep states and experimental days. This indicates that potential head motion during sleep or wakefulness did not significantly impact the data quality or the reliability of subsequent metabolite quantification.

Fourth, we also verified that neurometabolite changes were not driven by fluctuations in spectral quality. Specifically, Spearman's rank correlation analyses were performed to examine whether Glx or GABA concentrations during overall NREM sleep were linked to NAA SNR or NAA linewidth calculated from NREM sleep-only blocks across both experimental days (**Supplementary Table S5**). The analysis revealed no significant correlations between data quality and metabolite estimations. For instance, for overall NREM Glx on Day 1, concentrations showed no significant correlation with either NAA SNR ( $\rho = -0.04, p = 0.870, n = 19$ ) or NAA linewidth ( $\rho = -0.16, p = 0.523, n = 19$ ). Similarly, no significant correlations were observed for GABA on Day 1 or for any metrics on Day 2 (all  $p$ -values  $> 0.05$ ). These results collectively demonstrate that the quantified concentrations of Glx and GABA were robust against differences in spectral resolution and signal strength, confirming that our neurochemical findings were not driven by underlying data-quality artifacts.

## Discussion

We combined functional MRS with PSG and used the first-night effect (FNE) as a within-subject perturbation to induce variability in sleep-onset latency (SOL). We tracked Glx (glutamate+glutamine) and GABA in the mPFC across two nap sessions that differed in sleep

initiation: Day 1 was characterized by longer sleep-onset latency (SOL), whereas Day 2 showed a significantly shorter SOL. Glx dynamics tracked SOL more closely than GABA dynamics. Mediation analysis further indicated that Glx dynamics was closely linked with the between-session difference in SOL. These results support the hypothesis that insufficient downregulation of excitatory glutamatergic signaling in the mPFC is associated with transient difficulty in initiating sleep.

The decrease in Glx may index the deactivation of wake-promoting systems, given the critical role of excitatory signaling in maintaining wakefulness (Dash et al., 2009; Sulaman et al., 2023; Heiss et al., 2024). We measured Glx in the mPFC, which is considered an integral cortical component of wake-promoting circuits (Sulaman et al., 2023). The mPFC also serves as a bidirectional hub that links subcortical and cortical regions within arousal systems (Pal et al., 2018; Parkar et al., 2020; Huels et al., 2021; Mashour et al., 2022). Accordingly, reduced mPFC excitatory signaling may mark a successful wake-sleep transition and continued downshifting of wake-promoting activity as participants progress into deeper sleep stages. In this view, the Day 2 Glx decrease potentially reflects successful cortical disengagement during sleep initiation.

This FNE-based model may also provide insights into the mechanisms underlying clinical sleep disturbances. Insomnia is often conceptualized as hyperarousal, in which the brain fails to disengage wake-promoting neural activity during attempts to sleep (Levenson et al., 2015). In patients with primary insomnia, positron emission tomography (PET) has shown abnormally small reductions in brain metabolic activity during the wake-to-sleep transition, particularly in frontal cortical regions including the mPFC (Nofzinger et al., 2004). In our study, mPFC Glx did not decrease during disturbed sleep on Day 1, mirroring this pattern. The parallel between Day 1 and clinical PET findings suggests that blunted cortical downregulation of excitation may be a hallmark of difficulty initiating sleep in insomnia.

535

536 In contrast to the clear association between Glx and sleep initiation, the role of GABA in the human  
537 sleep transition appeared less direct than is often assumed. While we found no clear association  
538 between mPFC GABA and SOL, GABA concentrations were not static across states or sessions.  
539 Specifically, GABA levels were significantly higher on Day 1 than on Day 2 during N3, with a trend  
540 during N2 sleep. One interpretation is that this elevated GABA concentration during poor sleep  
541 reflects less efficient or unsuccessful inhibitory signaling. This could manifest as incomplete  
542 GABA binding or receptor engagement required for a successful transition into deeper sleep.

543

544 One may wonder whether alternative explanations could account for the null association between  
545 GABA and SOL. First, the mPFC may not be the optimal cortical region in which to detect the  
546 GABA dynamics that support sleep initiation. However, this seems unlikely because the mPFC is  
547 a bidirectional hub between subcortical and cortical regions that regulate arousal (Mashour et al.,  
548 2022). Second, it could be argued that the increase in inhibitory signaling needed to suppress  
549 wake-promoting systems was below the detection limits of our MRS measurements. However,  
550 our findings argue against the simple explanation based on detectability because, if our  
551 measurements were truly insensitive to GABA changes during sleep, we would not have detected  
552 the significant elevation in GABA during the disturbed sleep of Day 1.

553

554 This paradoxical observation of higher GABA during poorer sleep is particularly intriguing, as it  
555 contrasts with an assumption that higher GABA levels necessarily promote better sleep. Rather  
556 than driving the transition into sleep per se, mPFC GABA concentration may instead track sleep  
557 stage or sleep quality, perhaps reflecting a compensatory but ultimately unsuccessful inhibitory  
558 response to heightened arousal on the first night. This perspective may also reconcile our findings  
559 with the well-established efficacy of GABAergic hypnotics (Bateson, 2004; Sigel and Ernst, 2018).  
560 Importantly, these findings do not imply that GABA is functionally irrelevant to the sleep process;

rather, they suggest that while pharmacological intervention forcefully enhances inhibitory neurotransmission to induce sleep, under natural physiological conditions, the mere endogenous accumulation of GABA may not be sufficient to overcome heightened arousal if the underlying signaling efficiency is compromised. These findings suggest that therapeutic strategies focusing solely on boosting inhibition may be insufficient, as they may overlook the potential importance of excitatory downregulation, which was closely associated with efficient sleep initiation.

A key strength of the present study is the use of the FNE within a within-subject design, which allowed us to examine sleep-initiation-related variability while reducing the impact of individual differences and MRS measurement consistency. By recruiting a rigorously screened sample of healthy young adults and ensuring consistent voxel placement across sessions, we minimized the impact of intersubject heterogeneity and measurement error. These controls reduce the likelihood that the observed differences in Glx and GABA dynamics were driven by intersubject heterogeneity or measurement error.

However, several considerations remain for future research. First, because this study utilized a daytime nap protocol, the generalizability of these findings to nocturnal sleep warrants further investigation. While the neurochemical mechanisms of sleep initiation are likely conserved across different times of day, circadian timing remains an important factor. For instance, a previous study suggested that Glx levels exhibit a diurnal rhythm, increasing during the day and decreasing at night (Yoon et al., 2024). However, since that work did not employ PSG to verify sleep, it remains unclear whether those observed changes reflected circadian oscillation or the sleep process itself. Our PSG-defined data provide clear evidence that the transition into sleep is accompanied by a reduction in mPFC Glx during the day; future work should clarify whether this effect maintains its magnitude during nocturnal sleep onset, when homeostatic sleep pressure and circadian signals are more aligned.

Second, the use of the FNE as an experimental paradigm may involve multifactorial phenomena, including environmental novelty, psychological stress, and heightened vigilance, rather than a specific manipulation of sleep initiation alone. Therefore, the observed session differences in mPFC Glx may also reflect broader changes in arousal-related cortical activity rather than sleep initiation *per se*.

Third, while we interpreted the reduction in Glx as reflecting a downregulation of excitatory signaling, MRS-derived Glx measures total tissue concentrations and cannot directly isolate synaptic glutamatergic activity from broader metabolic processes. Therefore, our findings should be interpreted with caution, as the observed Glx changes may capture shifts in cellular metabolism alongside or instead of direct changes in synaptic neurotransmission.

Fourth, the inherent mismatch in temporal resolution between MRS data and PSG measurements remains a limitation. Although our time-resolved approach subdivided the MRS data into 2-minute windows, this scale is still broader than the finer precision of PSG sleep stage scoring, which is performed in 30-second epochs. Consequently, our data cannot definitively establish whether the observed Glx changes precede, coincide with, or follow the exact transition into sleep. Future longitudinal or higher-resolution studies are required to clarify the precise temporal dynamics driving sleep initiation.

Fifth, we acknowledge that the relatively modest sample size ( $n = 20$ ) constitutes a limitation, particularly for the sleep-stage specific analyses that involved even smaller subsets of participants. Furthermore, multiple comparison corrections were not applied to these exploratory analyses. Importantly, these specific findings exhibited moderate-to-large effect sizes, suggesting that they represent meaningful underlying biological patterns rather than mere noise. However,

given the number of statistical tests, the possibility of false-positive findings arising from uncorrected multiple testing cannot be fully excluded, and these results should therefore be interpreted with appropriate caution. Consequently, future studies with larger cohorts are necessary to verify the generalizability and replicability of these effects.

Sixth, due to its limited occurrence during our daytime nap protocol, REM sleep was excluded from the primary neurochemical analyses. While this exclusion was methodologically necessary to ensure data reliability, it leaves the role of mPFC Glx and GABA dynamics during REM sleep unexamined. Given that REM sleep is characterized by distinct cortical activation and neuromodulatory states, future investigations employing longer nocturnal sleep protocols are warranted to elucidate whether these metabolite trajectories differ during REM sleep compared to NREM sleep.

Finally, while we focused on the balance between glutamate and GABA to understand cortical excitation and inhibition, sleep initiation is a complex process involving numerous neuromodulatory systems beyond these two metabolites (Saper and Fuller, 2017; Scammell et al., 2017; Jones, 2020). Integrating our findings with the dynamics of other neurotransmitters will be essential for a complete understanding of the neurochemical architecture of the human sleep-wake transition.

In conclusion, our findings identify mPFC Glx dynamics as a robust neurochemical correlate of sleep initiation and suggest that cortical excitatory downregulation accompanies successful transition from wakefulness to sleep. By combining time-resolved functional MRS with PSG, this study provides a framework for investigating the neurochemical mechanisms underlying human sleep initiation in vivo.

(1476/1500 words)

## REFERENCES

- Agnew HW, Jr., Webb WB, Williams RL (1966) The first night effect: an EEG study of sleep. *Psychophysiology* 2:263-266.
- Andronesi OC, Loebel F, Bogner W, Marjańska M, Vander Heiden MG, Iafrate AJ, Dietrich J, Batchelor TT, Gerstner ER, Kaelin WG, Chi AS, Rosen BR, Cahill DP (2016) Treatment Response Assessment in IDH-Mutant Glioma Patients by Noninvasive 3D Functional Spectroscopic Mapping of 2-Hydroxyglutarate. *Clin Cancer Res* 22:1632-1641.
- Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 51:1173-1182.
- Bateson AN (2004) The benzodiazepine site of the GABAA receptor: an old target with new potential? *Sleep medicine* 5 Suppl 1:S9-15.
- Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, Troester MT, Vaughn BV (2017) AASM Scoring Manual Updates for 2017 (Version 2.4). *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine* 13:665-666.
- Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ (1989) The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28:193-213.
- Carskadon MA, Dement WC (1975) Sleep studies on a 90-minute day. *Electroencephalogr Clin Neurophysiol* 39:145-155.
- Dash MB, Douglas CL, Vyazovskiy VV, Cirelli C, Tononi G (2009) Long-term homeostasis of extracellular glutamate in the rat cerebral cortex across sleep and waking states. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29:620-629.
- Ding L, Chen B, Dai Y, Li Y (2022) A meta-analysis of the first-night effect in healthy individuals for the full age spectrum. *Sleep medicine* 89:159-165.
- Dydak U, Jiang YM, Long LL, Zhu H, Chen J, Li WM, Edden RA, Hu S, Fu X, Long Z, Mo XA, Meier D, Harezlak J, Aschner M, Murdoch JB, Zheng W (2011) In vivo measurement of brain GABA concentrations by magnetic resonance spectroscopy in smelters occupationally exposed to manganese. *Environ Health Perspect* 119:219-224.
- Edden RA, Puts NA, Harris AD, Barker PB, Evans CJ (2014) Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid-edited MR spectroscopy spectra. *J Magn Reson Imaging* 40:1445-1452.
- Ehlers CL, Kupfer DJ (1997) Slow-wave sleep: do young adult men and women age differently? *Journal of sleep research* 6:211-215.

676 Floyer-Lea A, Wylezinska M, Kincses T, Matthews PM (2006) Rapid modulation of GABA  
677 concentration in human sensorimotor cortex during motor learning. *Journal of*  
678 *neurophysiology* 95:1639-1644.

679 Frank SM, Becker M, Qi A, Geiger P, Frank UI, Rosedahl LA, Malloni WM, Sasaki Y, Greenlee  
680 MW, Watanabe T (2022) Efficient learning in children with rapid GABA boosting during  
681 and after training. *Current biology : CB* 32:5022-5030.e5027.

682 Friston KJ, Ashburner, J., Kiebel, S. J., & Nichols, T. E., Penny, W. D (2007) *Statistical*  
683 *Parametric Mapping: The Analysis of Functional Brain Images*: Elsevier/Academic  
684 Press.

685 Heiss JE, Zhong P, Lee SM, Yamanaka A, Kilduff TS (2024) Distinct lateral hypothalamic  
686 CaMKII $\alpha$  neuronal populations regulate wakefulness and locomotor activity. *Proceedings*  
687 *of the National Academy of Sciences of the United States of America* 121:e2316150121.

688 Horne JA, Ostberg O (1976) A self-assessment questionnaire to determine morningness-  
689 eveningness in human circadian rhythms. *Int J Chronobiol* 4:97-110.

690 Huels ER, Groenhout T, Fields CW, Liu T, Mashour GA, Pal D (2021) Inactivation of Prefrontal  
691 Cortex Delays Emergence From Sevoflurane Anesthesia. *Frontiers in systems*  
692 *neuroscience* 15:690717.

693 Iacobucci D, Saldanha N, Deng X (2007) A Meditation on Mediation: Evidence That Structural  
694 Equations Models Perform Better Than Regressions. *Journal of Consumer Psychology*  
695 17:139-153.

696 Jones BE (2020) Arousal and sleep circuits. *Neuropsychopharmacology : official publication of*  
697 *the American College of Neuropsychopharmacology* 45:6-20.

698 Levenson JC, Kay DB, Buysse DJ (2015) The pathophysiology of insomnia. *Chest* 147:1179-  
699 1192.

700 Lin A et al. (2021) Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy  
701 (MRSinMRS): Experts' consensus recommendations. *NMR in biomedicine* 34:e4484.

702 Lu J, Greco MA, Shiromani P, Saper CB (2000) Effect of lesions of the ventrolateral preoptic  
703 nucleus on NREM and REM sleep. *The Journal of neuroscience : the official journal of*  
704 *the Society for Neuroscience* 20:3830-3842.

705 Mander BA, Winer JR, Walker MP (2017) Sleep and Human Aging. *Neuron* 94:19-36.

706 Mashour GA, Pal D, Brown EN (2022) Prefrontal cortex as a key node in arousal circuitry.  
707 *Trends in neurosciences* 45:722-732.

708 Morin CM, Benca R (2012) Chronic insomnia. *Lancet* 379:1129-1141.

709 Moruzzi G, Magoun HW (1949) Brain stem reticular formation and activation of the EEG.  
710 *Electroencephalogr Clin Neurophysiol* 1:455-473.

711 Mullins PG, McGonigle DJ, O'Gorman RL, Puts NA, Vidyasagar R, Evans CJ, Edden RA (2014)  
712 Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA.  
713 *NeuroImage* 86:43-52.

714 Nofzinger EA, Buysse DJ, Germain A, Price JC, Miewald JM, Kupfer DJ (2004) Functional  
715 neuroimaging evidence for hyperarousal in insomnia. *The American journal of psychiatry*  
716 161:2126-2128.

717 Pal D, Dean JG, Liu T, Li D, Watson CJ, Hudetz AG, Mashour GA (2018) Differential Role of  
718 Prefrontal and Parietal Cortices in Controlling Level of Consciousness. *Current biology :*  
719 *CB* 28:2145-2152.e2145.

720 Parkar A, Fedrigon DC, Alam F, Vanini G, Mashour GA, Pal D (2020) Carbachol and Nicotine in  
721 Prefrontal Cortex Have Differential Effects on Sleep-Wake States. *Front Neurosci*  
722 14:567849.

723 Pasanta D, He JL, Ford T, Oeltzschner G, Lythgoe DJ, Puts NA (2023) Functional MRS studies  
724 of GABA and glutamate/Glx - A systematic review and meta-analysis. *Neuroscience and*  
725 *biobehavioral reviews* 144:104940.

726 Provencher SW (1993) Estimation of metabolite concentrations from localized in vivo proton  
727 NMR spectra. *Magnetic resonance in medicine : official journal of the Society of*  
728 *Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 30:672-  
729 679.

730 Roenneberg T, Wirz-Justice A, Mrosovsky M (2003) Life between clocks: daily temporal patterns of  
731 human chronotypes. *J Biol Rhythms* 18:80-90.

732 Saper CB, Fuller PM (2017) Wake-sleep circuitry: an overview. *Current opinion in neurobiology*  
733 44:186-192.

734 Saper CB, Scammell TE, Lu J (2005) Hypothalamic regulation of sleep and circadian rhythms.  
735 *Nature* 437:1257-1263.

736 Saper CB, Fuller PM, Pedersen NP, Lu J, Scammell TE (2010) Sleep state switching. *Neuron*  
737 68:1023-1042.

738 Scammell TE, Arrigoni E, Lipton JO (2017) Neural Circuitry of Wakefulness and Sleep. *Neuron*  
739 93:747-765.

740 Sherin JE, Shiromani PJ, McCarley RW, Saper CB (1996) Activation of ventrolateral preoptic  
741 neurons during sleep. *Science* 271:216-219.

742 Sigel E, Ernst M (2018) The Benzodiazepine Binding Sites of GABA(A) Receptors. *Trends*  
743 *Pharmacol Sci* 39:659-671.

744 Sulaman BA, Wang S, Tyan J, Eban-Rothschild A (2023) Neuro-orchestration of sleep and  
745 wakefulness. *Nature neuroscience* 26:196-212.

746 Tamaki M, Sasaki Y (2019) Surveillance During REM Sleep for the First-Night Effect. *Front*  
747 *Neurosci* 13:1161.

748 Tamaki M, Watanabe T, Sasaki Y (2021) Coregistration of magnetic resonance spectroscopy  
749 and polysomnography for sleep analysis in human subjects. *STAR Protoc* 2:100974.

- Tamaki M, Bang JW, Watanabe T, Sasaki Y (2016) Night Watch in One Brain Hemisphere during Sleep Associated with the First-Night Effect in Humans. *Current biology : CB* 26:1190-1194.
- Tamaki M, Yamada T, Barnes-Diana T, Wang Z, Watanabe T, Sasaki Y (2024) First-night effect reduces the beneficial effects of sleep on visual plasticity and modifies the underlying neurochemical processes. *Scientific Reports* 14:14388.
- Tamaki M, Wang Z, Barnes-Diana T, Guo D, Berard AV, Walsh E, Watanabe T, Sasaki Y (2020) Complementary contributions of non-REM and REM sleep to visual learning. *Nature neuroscience* 23:1150-1156.
- Watson CJ, Baghdoyan HA, Lydic R (2010) *Neuropharmacology of Sleep and Wakefulness*. *Sleep Med Clin* 5:513-528.
- Weber F, Dan Y (2016) Circuit-based interrogation of sleep control. *Nature* 538:51-59.
- Winkelman JW, Buxton OM, Jensen JE, Benson KL, O'Connor SP, Wang W, Renshaw PF (2008) Reduced brain GABA in primary insomnia: preliminary data from 4T proton magnetic resonance spectroscopy (1H-MRS). *Sleep* 31:1499-1506.
- Yeo IK, Johnson RA (2000) A new family of power transformations to improve normality or symmetry. *Biometrika* 87:954-959.
- Yoon S, Lee S, Joo Y, Ha E, Hong H, Song Y, Lee H, Kim S, Suh C, Lee CJ, Lyoo IK (2024) Variations in Brain Glutamate and Glutamine Levels Throughout the Sleep-Wake Cycle. *Biol Psychiatry*.
- Zammit G, Schwartz H, Roth T, Wang-Weigand S, Sainati S, Zhang J (2009) The effects of ramelteon in a first-night model of transient insomnia. *Sleep medicine* 10:55-59.
- Zhong H, Xu H, Li X, Xie RG, Shi Y, Wang Y, Tong L, Zhu Q, Han J, Tao H, Zhang L, Hu Z, Zhang X, Gu N, Dong H, Xu X (2023) A role of prefrontal cortico-hypothalamic projections in wake promotion. *Cereb Cortex* 33:3026-3042.

## Figure legends

### Figure 1. Voxel placement in the medial prefrontal cortical region.

The volume of interest (VOI) is overlaid on structural T1-weighted images from a representative participant in (a) sagittal and (b) transverse views.

### Figure 2. Comparison of sleep-onset latency (SOL) between sessions.

SOL was significantly longer on Day 1 than on Day 2, reflecting the expected first-night effect on sleep initiation. Error bars represent standard error of the mean (SEM). \*\* denotes  $p < 0.01$ .

**Figure 3. Glx downregulation during NREM sleep.**

Mean percentage change ( $\pm$  SEM) in Glx levels during NREM sleep relative to the wakefulness baseline is shown for Day 1 (white) and Day 2 (red). Glx was significantly reduced during sleep on Day 2, whereas this downregulation was absent on Day 1. Significance levels are indicated as \*  $p < 0.05$  and \*\*\*  $p < 0.001$ .

**Figure 4. Stage-specific Glx dynamics during NREM sleep.**

Mean percentage change ( $\pm$  SEM) in Glx levels from the wakefulness baseline is shown for each NREM sleep stage (N1, N2, and N3). Open squares represent Day 1, and filled red circles represent Day 2. The significant reduction in Glx on Day 2 became increasingly pronounced as sleep deepened, whereas Glx levels remained relatively stable or showed delayed reduction on Day 1. Significance levels are indicated as \*  $p < 0.05$  and \*\*  $p < 0.01$ .

**Figure 5. Statistical pathway linking experimental days and SOL via Glx levels.**

A tripartite analysis model was used to investigate the relationship between Session (Day 1 vs. Day 2), NREM Glx levels, and SOL, with age and sex included as covariates. The analysis shows that changes in Glx levels statistically accounted for the relationship between Session and SOL (indirect pathway  $a \times b = -2.21$ ,  $p = 0.020$ ), explaining 51% of the total effect after adjusting for the covariates. The total effect (path c) represents the overall relationship between Session and SOL, while the direct pathway (path c') represents the relationship after controlling for Glx, age,

and sex. Solid red lines indicate significant paths, and the dashed black line indicates a non-significant direct path. Numerical values represent path coefficients with p-values in parentheses.

**Figure 6. GABA levels during NREM sleep.**

Mean percentage change ( $\pm$  SEM) in GABA levels during NREM sleep relative to the wakefulness baseline is shown for Day 1 (white) and Day 2 (blue). In contrast to Glx, no significant differences in GABA levels were observed between sessions, and the change from baseline did not reach statistical significance in either session (n.s.,  $p > 0.05$ ).

**Figure 7. Stage-specific GABA dynamics during NREM sleep.**

Mean percentage change ( $\pm$  SEM) in GABA levels from the wakefulness baseline is shown for each NREM sleep stage (N1, N2, and N3). Open squares represent Day 1, and filled blue circles represent Day 2. While GABA levels remained higher on Day 1 relative to Day 2 across early sleep stages, these dynamics did not mirror the consistent downregulation observed in Glx. Significance levels are indicated as +  $p < 0.1$ , \*  $p < 0.05$  and \*\*  $p < 0.01$  for between-session comparisons.

**Figure 8. Pathways between experimental days and SOL showing no statistical account by GABA levels.**

A tripartite mediation model was used to investigate the relationship between Session (Day 1 vs. Day 2), NREM GABA levels, and SOL, with age and sex included as covariates. This analysis indicates that changes in GABA levels did not statistically account for the relationship between Session and SOL (indirect effect  $a \times b = -0.021$ ,  $p = 0.967$ ), accounting for only 1.0 percent of the

total effect. The total effect (path c) and the statistical direct pathway (path c', adjusting for GABA and covariates) remained virtually identical. These results, which contrast with those for Glx, suggest that the statistical link between Session and SOL does not implicate GABAergic dynamics. Solid blue lines indicate the tested indirect pathways, and the dashed black line indicates the statistical direct pathway. Numerical values represent path coefficients with p-values in parentheses.

**Figure 9. Stability of control metabolite and baseline concentrations.**

(a) Mean absolute concentrations of N-acetylaspartate (NAA) in institutional units (IU) (+/- SEM) during wakefulness (white) and NREM sleep (gray) for both sessions. (b) Mean baseline Glx levels (expressed as Glx/NAA ratio) during wakefulness on Day 1 and Day 2. (c) Mean baseline GABA levels (expressed as GABA/NAA ratio) during wakefulness on Day 1 and Day 2. No significant differences were observed across sessions or states for any of these baseline measures, confirming the stability of the MRS data.

## Tables

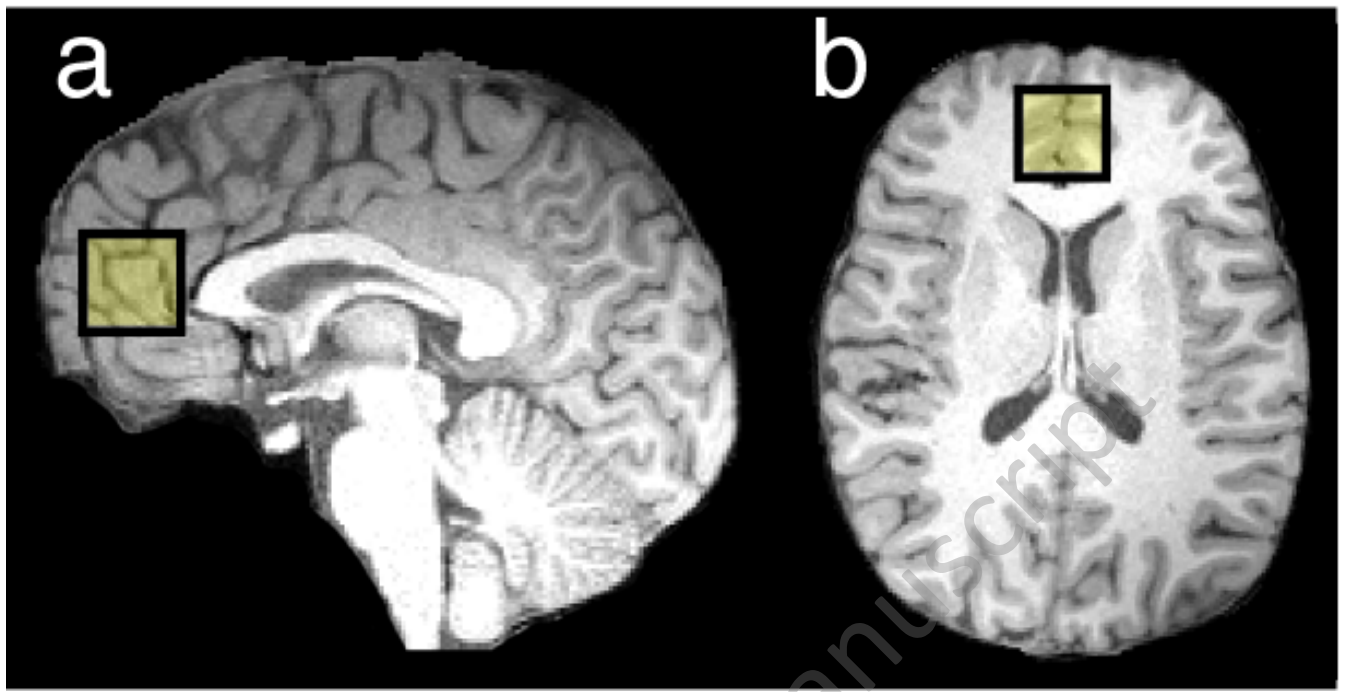
**Table 1: Sleep Architecture**

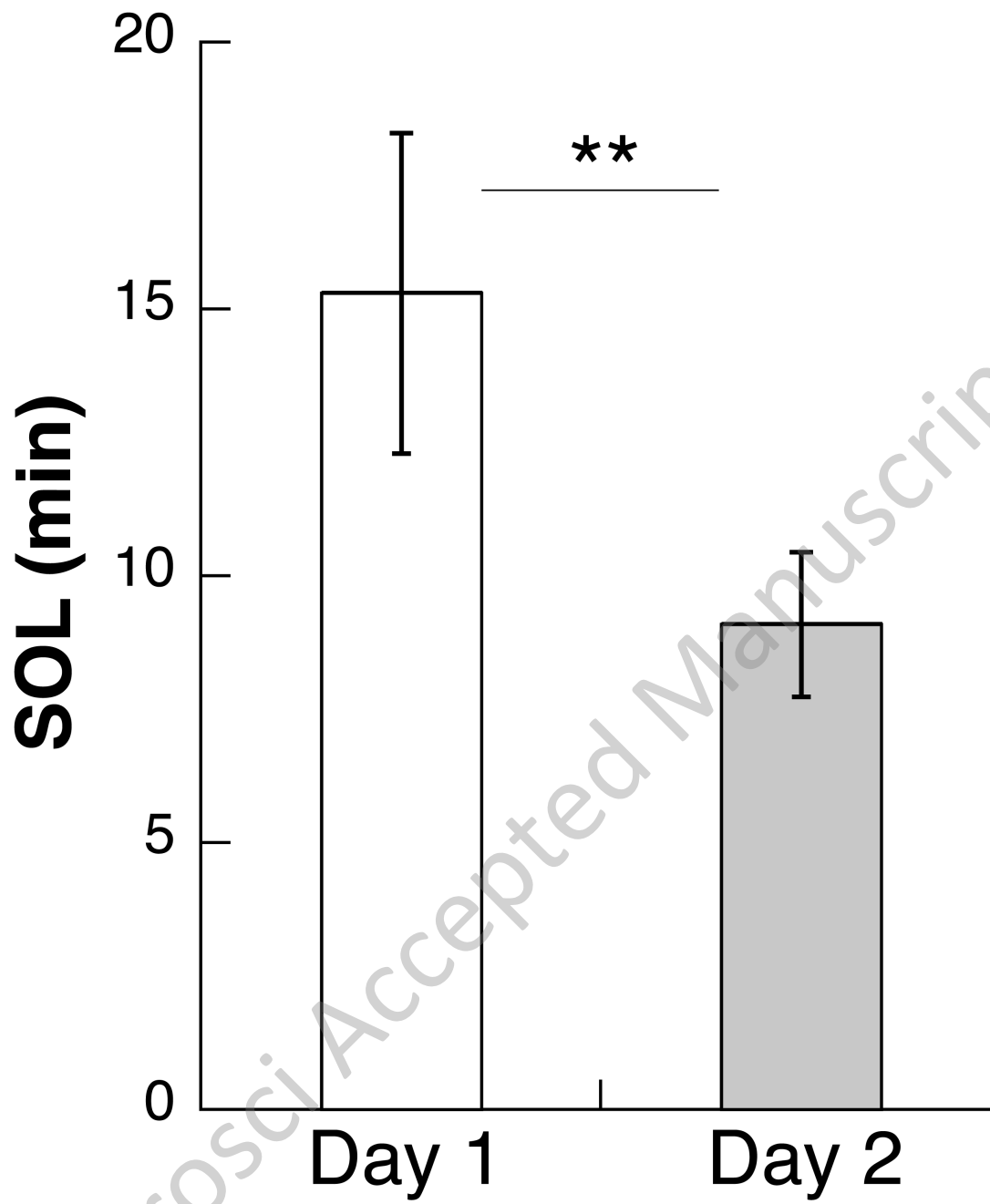
|                      | Day 1       | Day 2       | Wilcoxon signed-rank test |          |          |
|----------------------|-------------|-------------|---------------------------|----------|----------|
|                      |             |             | <i>Z</i>                  | <i>p</i> | <i>r</i> |
| SOL (min)            | 15.3 ± 2.99 | 9.1 ± 1.36  | -2.67                     | 0.008    | 0.60     |
| WASO (min)           | 13.5 ± 3.50 | 11.5 ± 3.17 | -0.56                     | 0.575    | 0.13     |
| Stage N1 (min)       | 25.4 ± 3.39 | 25.0 ± 2.60 | -0.09                     | 0.926    | 0.02     |
| Stage N2 (min)       | 27.3 ± 2.94 | 31.2 ± 3.24 | 1.66                      | 0.096    | 0.37     |
| Stage N3 (min)       | 9.5 ± 2.71  | 13.9 ± 3.67 | 1.05                      | 0.293    | 0.23     |
| Stage R (min)        | 3.3 ± 1.04  | 3.9 ± 0.91  | 0.77                      | 0.441    | 0.17     |
| Sleep efficiency (%) | 76.1 ± 4.57 | 83.7 ± 4.12 | 1.76                      | 0.079    | 0.39     |

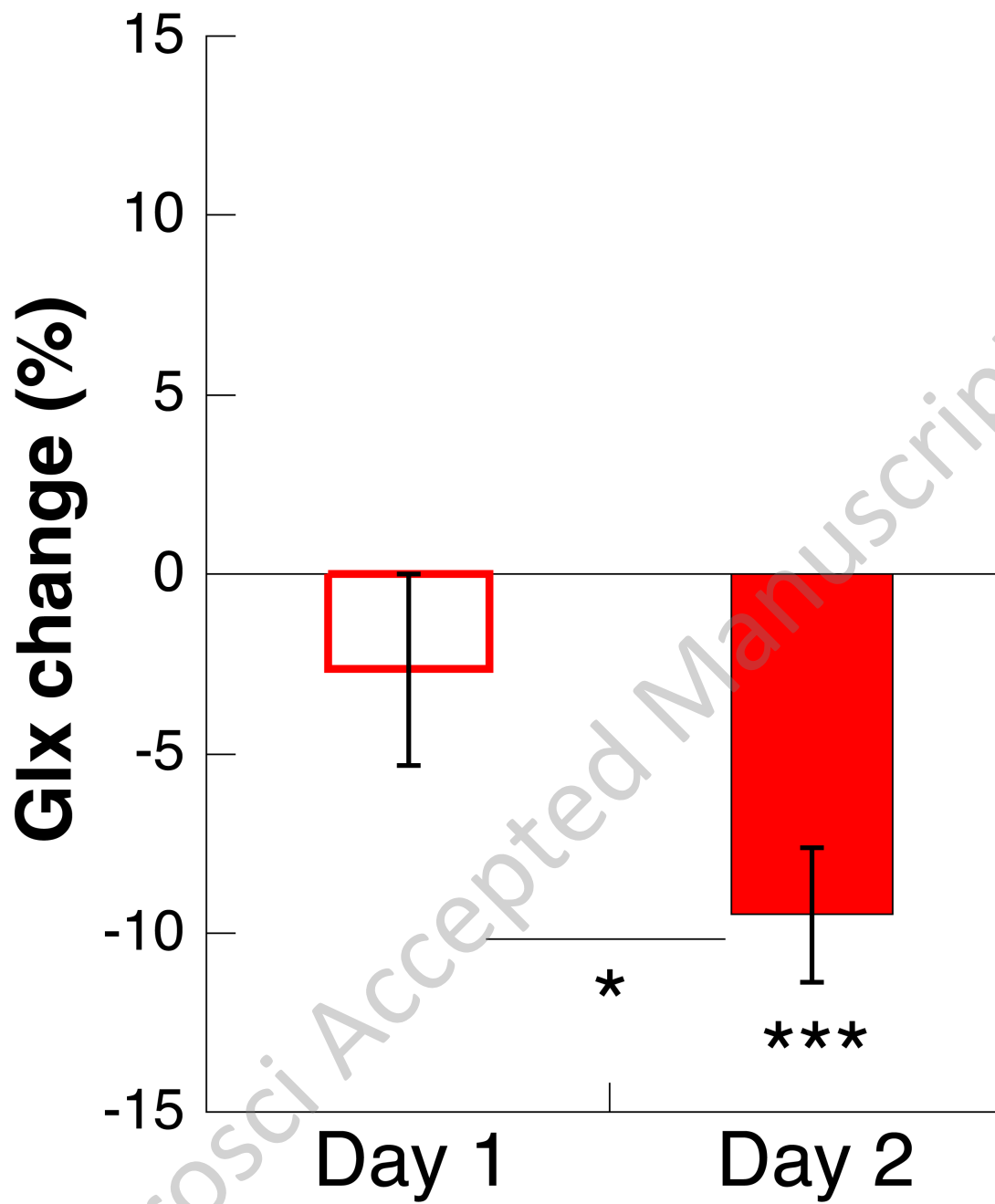
SOL, sleep-onset latency; WASO, wake after sleep onset. Values represent means ± SEM (n = 20). The related-samples Wilcoxon signed-rank test was used to compare parameters between sessions (Day 1 vs. Day 2). For the analysis of sleep architecture, values of zero were assigned to sleep stages that did not occur (e.g., N3 or REM sleep) to allow for a consistent comparison across the entire group. See Supplementary Table S1 for the number of participants who reached each sleep stage.

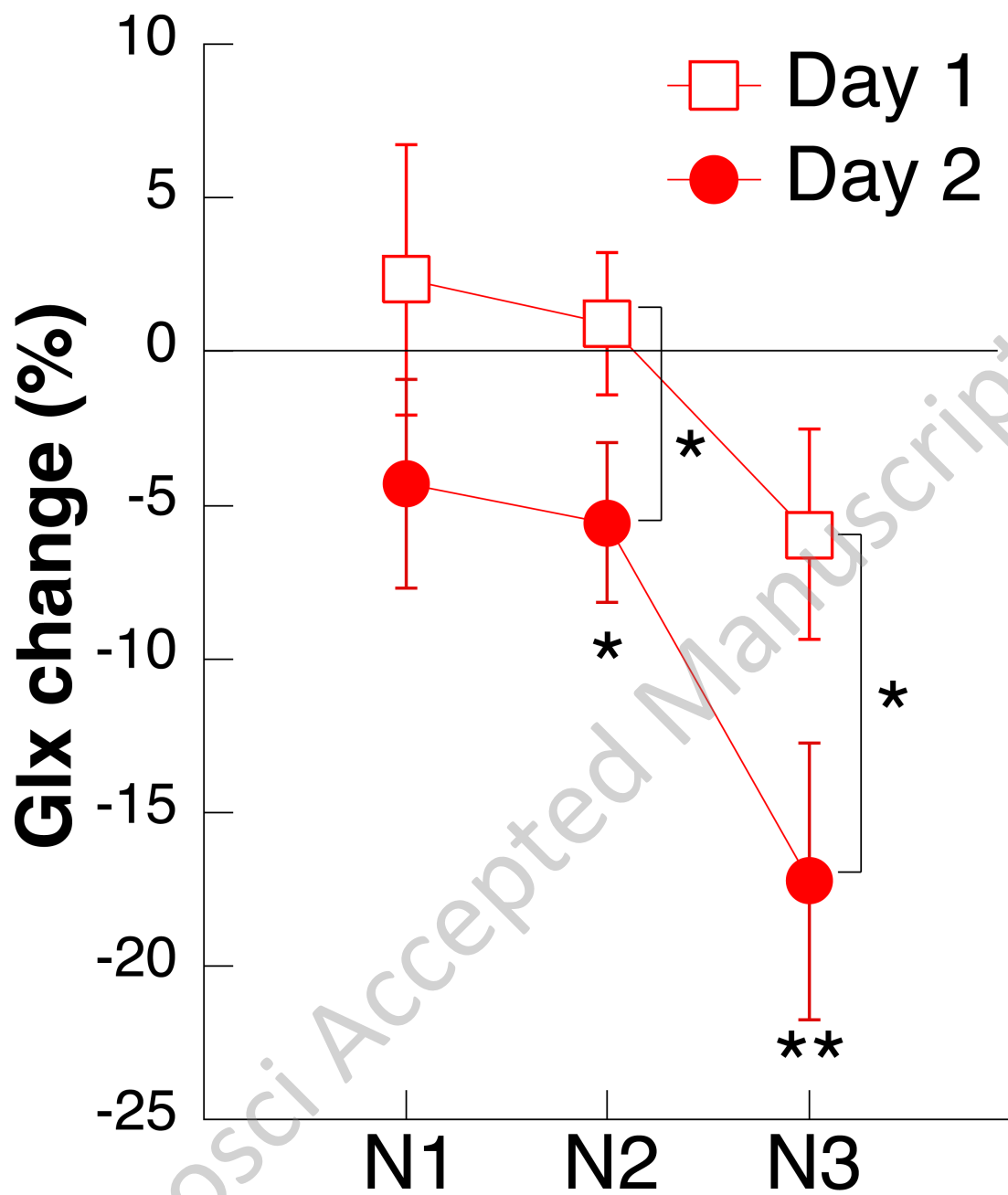
**Table 2. MRS quality test**

|                      | Day 1       | Day 2       | Wilcoxon signed-rank test |          |          |
|----------------------|-------------|-------------|---------------------------|----------|----------|
|                      |             |             | <i>Z</i>                  | <i>p</i> | <i>r</i> |
| Shim values          | 7.9 ± 0.65  | 7.0 ± 0.47  | -1.18                     | 0.240    | 0.26     |
| %SD for GABA         | 8.8 ± 0.50  | 9.4 ± 0.53  | 1.79                      | 0.073    | 0.40     |
| %SD for Glx          | 5.4 ± 0.16  | 5.5 ± 0.17  | 0.43                      | 0.667    | 0.10     |
| NAA linewidth (Hz)   | 11.0 ± 0.48 | 10.1 ± 0.32 | -1.01                     | 0.313    | 0.23     |
| Frequency drift (Hz) | 2.3 ± 0.30  | 1.9 ± 0.22  | -1.27                     | 0.204    | 0.28     |

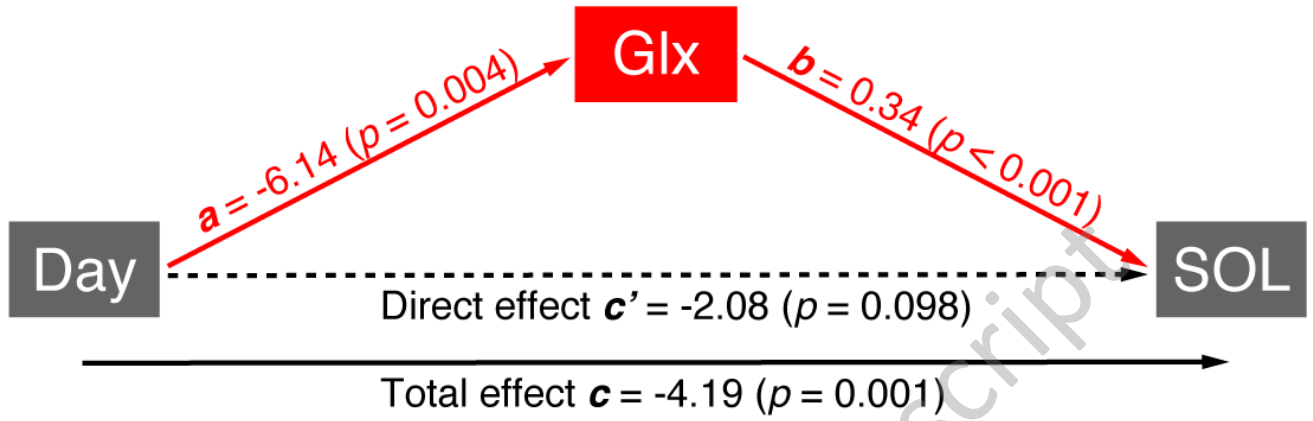


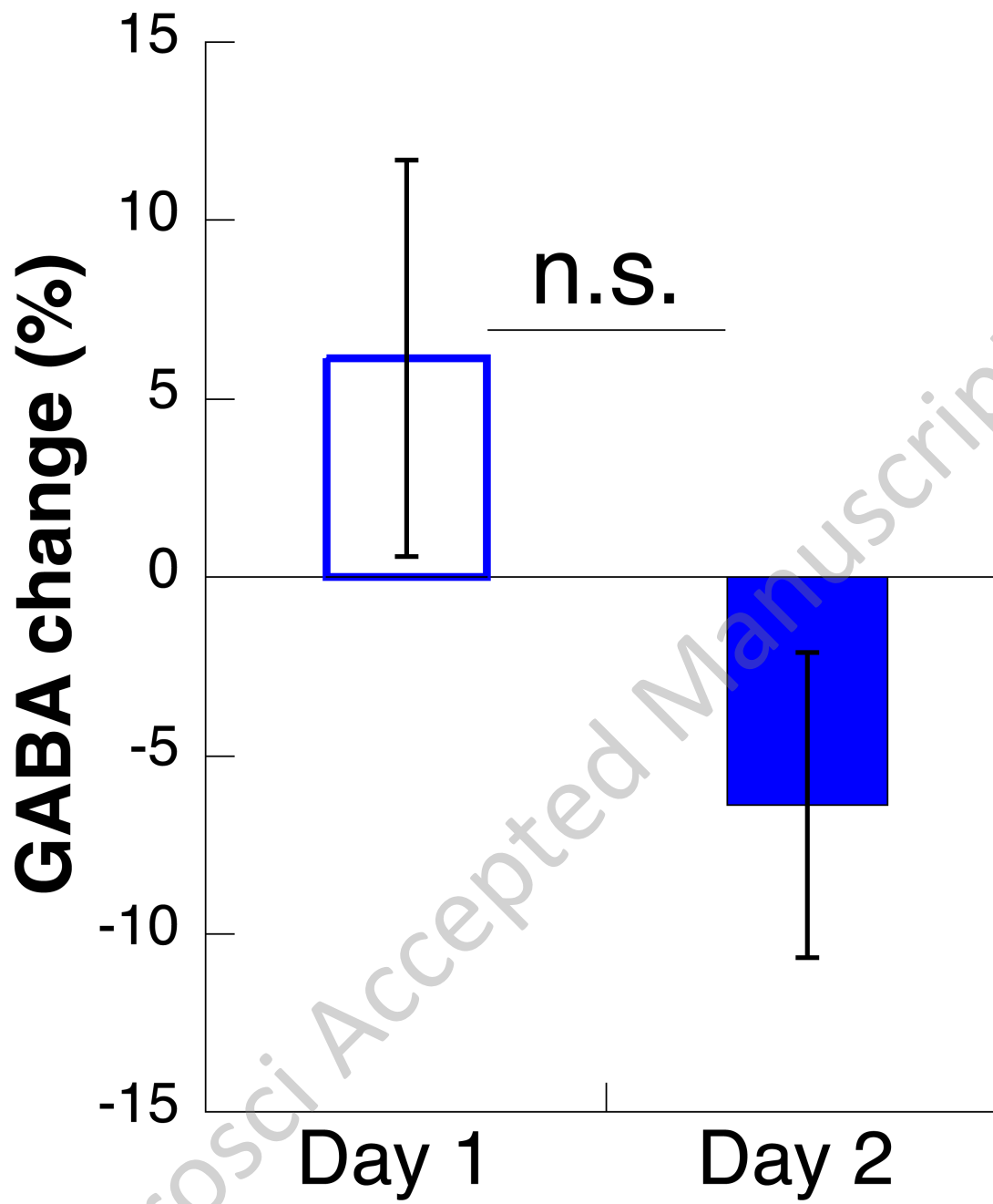


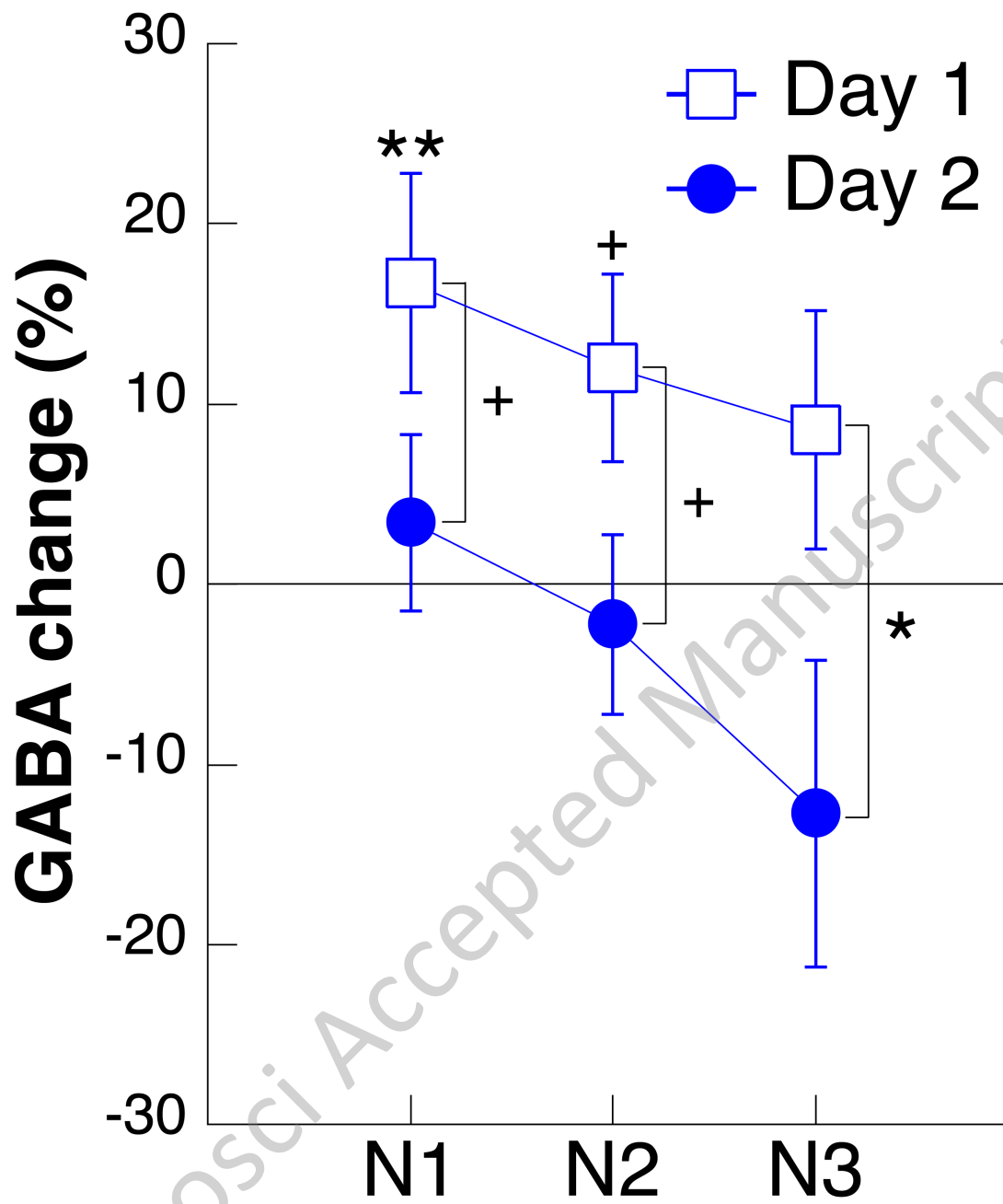




Mediation effect ( $a \times b$ ) = -2.21,  $p = 0.020$   
51% of total effect mediated







Mediation effect ( $a \times b$ ) = -0.021,  $p = 0.967$   
1% of total effect mediated

