

STUDY PROTOCOL

Validating a novel driving simulation-based MWT against the standard MWT in an OSA-cohort challenged by CPAP-withdrawal (DS-MWT2) – Protocol for a monocentric, controlled, randomized, crossover trial

Veronika Gambin¹, Nan Li¹, Esther Irene Schwarz², Kristina Keller¹, Stefan Lakämper^{1*}

1 Department of Traffic Medicine, Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland, **2** Department of Pulmonology and Sleep Disorders Centre, University Hospital Zurich, Zurich, Switzerland

* stefan.lakaemper@irm.uzh.ch



Abstract

Background

Excessive daytime sleepiness (EDS) is a major yet under-recognized contributor to road traffic accidents. Traditional diagnostic tools, such as the Maintenance of Wakefulness Test (MWT), assess wakefulness under passive conditions but may not accurately reflect real-world driving risks. To address this gap, we have piloted a Driving Simulation-based MWT (DS-MWT), designed to evaluate sleepiness in an ecologically valid driving scenario. The present study aims to validate the novel DS-MWT against the classical MWT in a functionally relevant cohort of patients with obstructive sleep apnoea (OSA).

Methods

The present monocentric, randomized, controlled, within-subject crossover trial will include 54 participants: 36 patients with OSA undergoing ≥ 7-day CPAP withdrawal (W) or continuation (C), and 18 healthy controls. The study employs a well-established CPAP-withdrawal model in patients with prior optimal treatment adherence to transiently induce EDS under controlled conditions. A healthy control group is included to enable between-group comparisons. The primary expected outcome is the difference in mean latencies between MWT and DS-MWT, determined during four standardized test sessions per condition. Secondary exploratory outcomes are defined as the presence, direction, and magnitude of differences or correlations between treatment status (CPAP withdrawal vs. continuation) and driving performance metrics (e.g., lateral position, speed, lane departures, etc.), EEG and eye-tracking features, subjective sleepiness scores, at-home polysomnography

OPEN ACCESS

Citation: Gambin V, Li N, Schwarz EI, Keller K, Lakämper S (2026) Validating a novel driving simulation-based MWT against the standard MWT in an OSA-cohort challenged by CPAP-withdrawal (DS-MWT2) – Protocol for a monocentric, controlled, randomized, crossover trial. PLoS One 21(7): e0341997. <https://doi.org/10.1371/journal.pone.0341997>

Editor: Christian Veauthier, Charité Medical University Berlin, GERMANY

Received: January 16, 2026

Accepted: June 3, 2026

Published: July 9, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0341997>

Copyright: © 2026 Gambin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/),

which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: No datasets were generated or analysed during the current study. All relevant data from this study will be made available upon study completion.

Funding: Predominant funding is provided by internal funding (personnel, UZH). Dr Stefan Lakämper received the grant from Emma-Louise-Kessler Fonds (Zurich, ZRH, CH), ID: DS-MWT2, URL: N/A. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: AE: Adverse Event; ASR: Annual Safety Report; CRF: Case Report Form; GCP: Good Clinical Practice; ICH: International Council for Harmonisation; HRA: Human Research Act (in German: HFG, in French: LRH, in Italian: LRUM); ClinO: Ordinance on Clinical Trials in Human Research (in German: KlinV, in French: OClin, in Italian: OSRUm); EEG: Electroencephalogram; ET: Eye-tracking; DS: driving simulation M; MWT: maintenance of wakefulness test D; DS-MWT: driving simulated maintenance of wakefulness test; IDS: integrated driving score; SDLP: standard deviation of lane position; SDS: standard deviation of velocity; SDD: distance to the leading vehicle in the car-following task; SDSW: steering wheel angle; V, Visit; ESS: Epworth Sleepiness scale; KSS: Karolinska Sleepiness scale; EDS: excessive daytime sleepiness; OSA: Obstructive sleep apnea; CPAP: continuous positive airway pressure; WCPAP: withdrawal; CCPAP: continuation; SA: Sleep Accident; AASM: American Academy of Sleep Medicine; PID: unique participant ID; AIC: Akaike Information Criterion.

(PSG) parameters, and metabolomic biomarkers (saliva, exhaled breath and dried blood spots). Data will be analyzed using linear mixed models, repeated-measures ANOVA, and predictive modeling with cross-validation.

Discussion

This trial addresses a critical limitation in sleep and traffic medicine by introducing a realistic, supposedly more ecologically valid alternative to standard sleepiness assessment tools. The DS-MWT may enhance clinical decision-making regarding fitness to drive (FTD) and provide a framework for identifying physiological and behavioral markers of sleepiness in realistic conditions.

Trial registration

ClinicalTrials.gov Identifier: [NCT06872593](https://clinicaltrials.gov/study/NCT06872593), released on 12.03.2025, <https://clinicaltrials.gov/study/NCT06872593>. Swiss National Clinical Trial Portal SNCTP000006301, released on 19.03.2025, <https://www.humanforschung-schweiz.ch/en/trial-search/study-detail/66469>.

Introduction

Background and rationale

Sleepiness at the wheel – frequently termed drowsy driving – is one of the major contributors to road accidents [1] with increasing need for an adequate legal regulation [2,3]. Recent work foregrounds the necessity for appropriate tools to assess sleepiness behind the wheel [4]. The proposed study aims to validate a recently developed naturalistic driving simulation-based test [5], which seeks to quantify the extent of sleepiness during a standardized monotonous night-time driving task.

In medical terms, sleepiness is often referred to as excessive daytime sleepiness (EDS), a symptomatic condition characterized by an increased propensity to fall asleep during waking hours. EDS may result from insufficient or compromised sleep, caused by psychosocial factors (e.g., stress), voluntary sleep debt, shift work, or medical conditions such as narcolepsy or obstructive sleep apnoea (OSA) [6]. Among the medical causes of EDS, OSA is one of the most prevalent and well-characterized, it is marked by recurrent episodes of upper airway obstruction during sleep, resulting in fragmented sleep and intermittent hypoxia [7]. Although the primary treatment of OSA with continuous positive airway pressure (CPAP) can alleviate EDS [8], a substantial subset of patients continues to report EDS despite receiving an optimized therapy [9]. Untreated EDS poses a substantial risk to traffic safety and is considered one of the highest-ranking causes of motor vehicle accidents [1,10,11].

Sleepiness and its dangers at the wheel might subjectively not be registered by the affected drivers [12], where the subjective perception of sleepiness frequently fails to align with objective measures of wakefulness [13].

Given the limitations of subjective perception, objective tools are essential for the assessment of sleepiness. There exists a variety of partially complementary tools to evaluate the extent of EDS. Mean sleep latencies obtained in the maintenance of wakefulness tests (MWT) are commonly used in assessing EDS. MWT latency refers to the time it takes for a person to fall asleep during a 40-minute-long sitting session. It is defined as the interval (in minutes) from the start of the session (lights out) until the start of the first epoch of any stage of sleep (an epoch of N1, N2, N3, or R) [14]. However, there is inconsistent or insufficient evidence for MWT to reliably predict the fitness-to-drive (FTD) in general, potentially due to the test's susceptibility to individual motivational factors [15,16]. Moreover, the cost and time demands of MWTs limit their feasibility for widespread application in driving license certification [17].

Currently debated is whether the MWT should be replaced or supplemented by emerging tools [4], such as future roadside metabolomic tests for detecting sleepiness in traffic [18,19], or whether the MWT itself should be adapted to more effectively communicate the risks of EDS for road safety and its relevance in determining FTD [20,21].

In line with these arguments, a driving simulation-based MWT (DS-MWT) was developed to complement the classical MWT by providing an additional assessment approach with greater ecological relevance for FTD evaluation [5]. Specifically, the operational structure of MWT was retained, including electroencephalography (EEG) and video monitoring across four 40-minute soporific sessions. The drowsiness-inducing conditions were achieved through a highly monotonous, nighttime car-following scenario, which directly mimics real-life driving environments in which a considerable proportion of vehicle accidents occur [22]. In addition to the solely EEG-based criteria used to define the latency in MWT, the DS-MWT latency was complemented by a "naturalistic" criterion, an accident with closed eyes ("sleep accident," SA). The pilot resulted in about 23% of driving sessions (runs) terminating with SA in the unstructured cohort, yielding a run latency, and thus inherently displayed sufficiently soporific conditions of DS-MWT [5]. Participant feedback indicated high user acceptance and increased ecological face validity.

While the pilot served as a successful feasibility study, offering proof of concept with encouraging outcomes, these findings warrant a systematic comparison of MWT and DS-MWT based on their resulting latencies. It also remains to be determined whether these latencies reliably reflect clinical parameters of EDS associated with underlying medical conditions such as OSA. Assessing the extent of EDS in the OSA-affected population may be further complicated by suboptimal CPAP adherence in CPAP-naïve patients, potentially leading to inconsistent and inconclusive results. A previously established experimental model suggests overcoming these drawbacks by employing a restricted pool of patients effectively treated and compliant with CPAP-therapy [23,24]. Short-term CPAP withdrawal in patients with previously optimal adherence has been shown to reactivate OSA and induce associated consequences, such as a deterioration of daytime symptoms and psychomotor performance [24,25]. However, it remains unclear whether EDS itself is consistently and measurably reactivated in this model, as no systematic data on its recurrence are currently available. Overall, this trial may contribute valuable insight into whether short-term CPAP withdrawal can reliably induce EDS and whether this effect can be captured through objective measures.

This constellation of unresolved questions constitutes an important evidence gap for both medical experts in pulmonology and traffic medicine, but also for affected drivers. This gap will be addressed through a systematic comparison of classical and simulation-based MWTs by means of their resulting latencies. In a within-subject design involving 36 effectively treated and highly CPAP-adherent OSA patients, latency variations will be analyzed in response to controlled CPAP therapy withdrawal (W) and continuation (C), each maintained for a period of at least 7 days. A healthy control group (N = 18) is included to enable comparison of MWT- and DS-MWT- latencies across clinical and non-clinical populations.

Objectives

The primary objective of this trial is to determine whether there is a statistically significant difference in latency between the classical MWT and the DS-MWT. The primary hypothesis assumes no statistically significant difference between MWT and DS-MWT latencies. This hypothesis will be tested exclusively using mean latency values, independent of secondary

and exploratory analyses. Latency differences will be measured in relation to the intervention CPAP withdrawal (OSA-affected only) and in relation to the test sequence (both OSA-affected and healthy control groups).

Secondary, exploratory objectives are designed to evaluate the effects of the intervention (CPAP withdrawal vs. CPAP continuation) and to relate these effects to those observed in a healthy control group. The following secondary objectives are based on a range of behavioral, physiological, and performance-related parameters.

The study will investigate potential associations between CPAP treatment status and driving performance indicators, including the standard deviation of lateral position (SDLP), driving speed, following distance, and steering wheel angle variability. These driving parameters will also be analyzed in relation to subjective perception of sleepiness, measured by the Epworth Sleepiness Scale (ESS) and the Karolinska Sleepiness Scale (KSS).

The relationship between CPAP condition and eye-tracking parameters, including blink rate, blink duration, eye openness, gaze direction and duration, and attention distribution across predefined areas of interest (e.g., windshield, side mirrors, rear mirror), will also be explored. In parallel, EEG-derived features will be examined as supplementary indicators that may reflect sleepiness, including microsleep-like episodes (MSEs) and individual electrophysiological signatures of sleep–wake transitions.

In addition, based on at-home polysomnography (PSG) recordings preceding each MWT and DS-MWT, night-to-night variability in the sleep quality will be examined in both groups, as well as key OSA parameters in the patient group, such as the apnoea–hypopnea index (AHI), oxygen desaturation index (ODI) and cumulative time with oxygen saturation below 90% [26].

Finally, the potential utility of noninvasive biological samples, including saliva, exhaled breath, and dried blood spots, for identifying proteomic biomarkers associated with sleepiness will be explored [18,19]. These biomarkers will be analyzed in relation to sleep quality, CPAP treatment condition and subjective sleepiness metrics in the OSA group, and will be compared to baseline values from healthy participants.

Materials and methods

Trial setting

This study will be conducted solely on the premises of the University of Zurich, Zurich, Switzerland. Data collection will be conducted at the Traffic Medicine Department of the Institute of Forensic Medicine, University of Zurich.

Trial status

The trial is recruiting. The original study protocol was first approved in version 1.1 by the ethics committee and amended twice, with the current protocol version being 1.3 (dated 08/07/2025). The planned overall duration of the study is 24 months, with 01/05/2025 being First-Participant-in (FPFV, anticipated date of the screening/enrollment of the first patient) and 30/04/2027 being Last-Participant-Out (LPLV, anticipated date of the last participant's final visit). A published manuscript or submission of the final report will be handed in by 30/04/2028.

Ethics

Research ethics approval. This study is conducted in accordance with the current version of the Declaration of Helsinki, the International Council for Harmonisation Good Clinical Practice (ICH-GCP) guidelines, the Human Research Act (HRA), and all other locally applicable regulatory requirements. The trial protocol and related study materials were reviewed and approved by the Kantonale Ethikkommission Zürich (Zurich Cantonal Ethics Committee), with the reference number 2024-01948. No participant recruitment or data collection activities commenced prior to obtaining this ethics approval.

Protocol amendments. Any substantial amendments to the protocol, including changes to study objectives, design, methodology, study site(s), sponsor, or key study personnel, will be submitted for review and approval by the Ethics Committee prior to implementation. Exceptions may only occur under emergency circumstances to ensure participant safety and well-being; such deviations shall be documented and reported to the Ethics Committee as soon as possible.

Substantial amendments are defined in accordance with ClinO (Ordinance on Clinical Trials in Human Research), Art. 29, and include changes affecting participant safety, rights, or obligations. A list of substantial changes is maintained on www.swissethics.ch. Non-substantial amendments will be compiled and submitted annually to the Ethics Committee together with the Annual Safety Report (ASR).

Confidentiality. Participant confidentiality will be maintained throughout the trial in compliance with Swiss data protection regulations. Identifiable information is stored separately in a subject identification log, linking names to PIDs assigned in sequence of enrolment. This log also records reasons for exclusion, early withdrawal, or non-adherence.

Access to identifiable data is restricted to authorized personnel involved in the study. Any disclosure to third parties requires prior written permission from the data owner. Study data may be accessed by monitors or the Ethics Committee solely for verification purposes.

Harms. This study is classified as category A, involving only minimal risk to participants. The procedures conducted across the seven study visits consist predominantly of standardized, non-invasive assessments, performed under continuous supervision by trained research personnel.

The use of a driving simulator may occasionally cause dizziness, nausea, or discomfort due to simulator sickness. These symptoms typically subside promptly upon test completion or leaving the simulator environment.

EEG monitoring and blood collection via a peripheral venous catheter are performed by qualified personnel. Minor adverse events, such as hematoma at the venipuncture site or skin irritation from EEG electrodes, may occur but are generally transient and resolve without complications. The total blood volume collected (9 mL) poses no significant health risk. The collection of saliva, urine, DBSs, and exhaled breath is entirely non-invasive and carries no known risks.

Participants assigned to the intervention arm will discontinue their nightly CPAP therapy at home for a period of ≥ 7 days, during which a temporary reactivation of OSA symptoms is expected. This CPAP-withdrawal model is clinically validated as a safe and effective method for studying the pathophysiology of OSA [24,27]. Only patients deemed low risk for adverse outcomes by physicians at the Sleep Disorders Center, based on clinical judgment and prior screening, will be included. Participants will resume their established long-term CPAP therapy immediately after the study.

All procedures, with the exception of at-home PSG recordings, will be conducted in controlled environments by trained staff, with medical oversight provided as needed. Participants are closely monitored for any adverse effects, and appropriate support will be made available. Individuals retain the right to withdraw from the study at any point without consequences for their ongoing medical care.

Participation in the study offers no direct therapeutic benefit to the individuals involved. However, the findings may contribute to the advancement of diagnostic and safety tools in traffic medicine and sleep disorder management, potentially benefiting future patients and public health initiatives.

Trial design

This study is a monocentric, controlled, within-subjects, randomized, crossover clinical trial in a cohort of OSA patients and in a healthy control group. The study consists of two equivalent arms designed to assess the impact of the test sequence based on the initial condition. The protocol will be reported following the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guideline [S1 File \[28\]](#).

Eligibility criteria

Inclusion criteria (OSA patients):

- adult driver
- diagnosed OSA

- established CPAP-treatment regime, highly adherent and compliant within the last 6 months (>5h, >80% of days)
- at impaired eyesight with more than +/- 5 diopter or astigmatism, contact lenses are required (for eye tracking (ET))

Inclusion criteria (healthy comparison group):

- adult driver
- no declared psychiatric disorders
- no declared sleep-related diagnosis
- at impaired eyesight with more than +/- 5 diopter or astigmatism, contact lenses are required (for ET)

Exclusion criteria (for both groups):

- sensibility to motion sickness (kinetosis, dizziness, etc. in 5 min screening drive)
- professional drivers (if working during the study period)
- inability to understand the study procedure for linguistic or cognitive reasons

Based on the inclusion criteria and the preselected pool for recruitment, no vulnerable participants, as defined by the Swiss HRA (chapter 3, i.e., children, adolescents, adults lacking capacity, or prisoners), will be included. However, participation in the experimental procedure does not pose any additional risks for pregnant women, nor for patients, healthy adults or the unborn child. Therefore, the participation of pregnant individuals is possible, and a pregnancy test is not required.

Recruitment

The total sample size for the study is $N_{\text{total}} = 54$, comprising an intervention group of $N_I = 36$ patients with OSA and a healthy comparison group of $N_H = 18$. Details of the sample size calculation are presented in section [Sample size calculation](#).

Intervention group (OSA patients). Participants in the intervention group will be recruited from a restricted pool of patients treated at the Sleep Disorders Center and Pulmonary Division of the University Hospital Zurich. This population is characterized by high adherence to CPAP therapy and regular clinical follow-up. Eligible individuals are preselected by the partner institution based on clinical criteria, including the absence of substantial comorbidities (e.g., major affective, somatic, or neurological disorders), cognitive impairments, or increased risk associated with CPAP withdrawal.

Although gender balance is desired, it cannot be guaranteed in the intervention group due to the significantly higher prevalence of OSA in males. Given the limited sample size and inter-individual variability of outcome measures, retrospective stratification by gender is not planned at this stage.

Healthy comparison group. Healthy participants will be recruited from the general population. Recruitment will be facilitated through institutional communication channels of the University of Zurich, including digital platforms and email distribution. Additionally, advertisements will be placed in local schools and public areas throughout Zurich to inform about the open recruitment status.

Recruitment process. First contact will be offered by a flyer containing contact information to the investigators. Interested individuals can reach out to the study team via phone, email, video call (optional), or in-person at the study site. Investigators will provide detailed information about the study objectives, design, procedures, duration, risks, benefits, and any potential discomforts. These explanations can be delivered individually or in groups, using the participant's preferred communication method.

All participants will receive a comprehensive information sheet and informed consent form, ensuring that they have sufficient detail to make an informed decision regarding their participation. Enough time, as much as the participant desires,

will be given for consideration, and all participants will be informed that participation is entirely voluntary, with the option to withdraw at any time without consequences for their ongoing medical care.

As part of the screening process, all prospective participants will complete a brief, five-minute simulated drive to assess for motion sickness (kinetosis or dizziness), which constitutes an exclusion criterion. This session also ensures that participants fully understand the simulation task requirements before enrollment.

Consent and assent. Verbal and written informed consent will be obtained by the sponsor-investigator or his designee (certified by GCP) prior to any study-specific procedures, including the collection of biological samples. The informed consent process will be documented in the participant's file. Any deviations from the consent process as outlined in the protocol will be recorded and justified.

The participant will receive a copy of the signed consent form; the original will be retained in the study files. Consent includes permission for data collection, storage, analysis, and verification by study monitors or the ethics committee. For any ancillary studies not explicitly covered in the primary study's informed consent form, additional written informed consent will be obtained from each participant prior to their involvement.

Intervention and comparator

The intervention consists of a ≥ 7 -day withdrawal of CPAP therapy in patients with established adherence to treatment.

To reliably induce EDS for assessment, the study employs a validated intervention model that involves temporarily withdrawing CPAP therapy in patients with OSA [24]. The intervention arm involves discontinuation of CPAP therapy for a minimum of seven consecutive days prior to the assessment period. The assessments are conducted after either one or two weeks on the respective treatment, depending on the assigned crossover sequence.

Discontinuation or modification of the allocated interventions is not anticipated under normal study conditions. However, discontinuation may occur in case of adverse events or if participants fail to comply with protocol requirements. Additionally, if necessitated by organizational constraints or participant-related circumstances, a switch between treatment conditions may be permitted.

Concomitant care is managed with attention to participant well-being. If a participant experiences discomfort or requires assistance, a designated contact person may be called to escort them. If no such person is available, transportation to the participant's home will be arranged by taxi. In the case of adverse events (AEs) during or following the study, participants are instructed to contact the sponsor-investigator or study physician, who will initiate appropriate medical follow-up. At the discretion of the clinical team, additional follow-up visits may be scheduled as needed.

Participant timeline

The overview of the participant timeline is shown in the SPIRIT [28] schedule of enrollment, interventions, and assessment Fig 1.

The study comprises seven visits per participant, combining in-laboratory assessments with home-based PSG and interventional testing. Visit 1 involves participant onboarding, including informed consent, eligibility confirmation, collection of baseline vital signs (including blood pressure), a brief screening drive (to assess for motion sickness as an exclusion criterion), and randomization by lottery. This visit lasts approximately 1 hour.

Visits 2 and 5 are dedicated to mounting EEG electrodes for home-based PSG recordings. Prior to each PSG recording, participant anthropometric and physiological parameters, including weight, height, and heart rate, are documented. Each of these visits requires approximately 1 hour of participant presence. PSGs are recorded overnight on the nights prior to visits 3, 4, and 6, 7.

Visits 3, 4, 6, and 7 include the primary experimental procedures: MWT (M) and the DS-MWT (D). These assessments comprise four 40-minute-long sessions (also termed "runs") with EEG measurements including video, starting at 9 a.m.,

	TRIAL PERIOD							
	Enrollment		Post-randomization					
Visit		V1	V2	V3	V4	V5	V6	V7
Timepoint [days]	<-1	0	$t_1 \geq 7$	$t_1 + 1$	$t_1 + 2$	$t_2 \geq t_1 + 7$	$t_2 + 1$	$t_2 + 2$
Duration [h]	0.5	1	1	8	8	1	8	8
ENROLLMENT:								
First contact: with patients by physician investigator/ by flyer with HC	X							
Second contact: offering full information by study investigator	X							
Eligibility screen		X						
Informed consent		X						
Randomization		X						
INTERVENTION/ COMPARATOR:								
Intervention: CPAP withdrawal (Group 1)*			→					
Intervention: CPAP withdrawal (Group 2)*						→		
Comparator: CPAP continuation (Group 1)*						→		
Comparator: CPAP continuation (Group 2)*			→					
Healthy control			X	X	X	X	X	X
ASSESSMENTS:								
Questionnaire: general demographic data		X						
Questionnaire: driving behavior		X						
Questionnaire: medical history		X						
Questionnaire: Epworth Sleepiness Scale (ESS)		X						
Questionnaire: Karolinska Sleepiness Scale (KSS)			X	X		X	X	
Mounting EEG			X			X		
PSG				X	X		X	X
DS-MWT condition (Group A)^				X				X
DS-MWT condition (Group B)^					X		X	
MWT condition (Group A)^					X		X	
MWT condition (Group B)^				X				X
ET recording				X	X		X	X
EEG recording				X	X		X	X
Sampling (blood, saliva, exhaled breath, DBS)				X	X		X	X
Questionnaire: Feedback								X

Fig 1. SPIRIT schedule of enrollment, interventions, and assessments. V=visit; HC=healthy control group; PSG = at-home polysomnography prior to respective visit; DS-MWT=driving simulation-based maintenance of wakefulness test; MWT=maintenance of wakefulness test; CPAP=continuous positive airway pressure; EEG=electroencephalography; ET=eye tracking; DBS=dried blood spots. Visit 1 comprises inclusion and a brief screening. Visit 2 and Visit 5 serve merely to mount PSGs. *=prior to V2-V4 patients are randomly assigned to either a starting condition CPAP withdrawal (Group 1) or to a starting condition CPAP continuation (Group 2), and crossover to the respective other condition during V5-V7. ^=participants are randomly assigned to either DS-MWT (Group A) or MWT (Group B) starting condition during V3, V4 and then crossover their condition during V6, V7.

<https://doi.org/10.1371/journal.pone.0341997.g001>

11 a.m., 1 p.m. and 3 p.m. During each of these visits, blood pressure measurements are additionally performed as part of participant safety monitoring and to account for potential effects on collected bio-samples.

Each session is preceded by a subjective sleepiness rating using KSS questionnaire. During the MWT, participants are seated in a comfortable semi-recumbent position in a dim and quiet environment; in the DS-MWT, participants perform a

monotonous nighttime car-following task in a full-vehicle driving simulator while seated in a comfortable supported driving posture comparable to the MWT setup. After each session, participants answer the KSS questionnaire again.

All sessions include continuous EEG monitoring and video recording. ET is implemented in all DS-MWT sessions and in MWT sessions where instrumentation is available. Following each run, bio-samples — including blood, saliva, exhaled breath, and dried blood spots — are collected.

The presence time for visits 3, 4, 6, and 7 is estimated at 8 hours each, including approximately 4 hours of active testing (4 × 40-minute sessions) and 15-minute post-session sampling intervals. EEG electrodes are removed after visits 4 and 7, but are checked for signal quality during all relevant visits and adjusted if needed.

To reduce the potential influence of wake-promoting substances on study outcomes, participants are instructed to abstain from caffeine and alcohol during the study period, including the week prior to and throughout all study visits. Compliance is based on participant self-report and is not verified through blood or toxicological testing. In addition, participants are required to document all current medications in standardized questionnaires.

Participants are invited to provide feedback upon study completion, either via an online form or in a personal follow-up conversation. An optional post-study interview is available for participants who wish to discuss their experiences, perceived effects, or reflections on the intervention's impact on daily life.

Randomization

Participants will be randomized to one of two starting conditions—either CPAP withdrawal (W) or continuation (C)—and assigned to one of two test sequences ($D \rightarrow M$ or $M \rightarrow D$), which determine the measurement sequence. Randomization will occur following successful screening and is implemented via random lot drawing from an opaque lottery box containing sealed paper assignments. The randomization scheme can be seen on [Fig 2](#)

No stratification will be applied based on demographic or clinical characteristics (e.g., sex, age, medication, dosage, or treatment duration). However, exploratory post hoc analyses will assess whether these factors influence outcome measures.

To ensure allocation concealment, the randomization lots are prepared in advance by the sponsor-investigator, who is not involved in the enrollment or assignment process. This separation preserves the integrity of the randomization and minimizes selection bias.

Due to the nature of the intervention, full blinding of participants and investigators is not feasible. Therefore, no formal unblinding procedure is necessary. However, outcome assessors will remain blinded to group allocation, including post-test, in order to prevent bias in the evaluation of subsequent participants.

Outcomes

The primary outcome measure of this trial will be the difference in mean latency between the classical MWT and the DS-MWT. Sleep latency in MWT is determined according to the American Academy of Sleep Medicine (AASM) criteria [\[29\]](#), and DS-MWT latency is defined as the time to either the occurrence of sustained sleep according to AASM criteria or a SA with eyes closed, whichever occurs first [\[5\]](#). For both MWT and DS-MWT, the mean latency will be calculated as the average of all latencies across four standardized test sessions (each lasting a maximum of 40 minutes) conducted on the same test day.

For the secondary outcomes, the trial will include several measurement variables and assessments, designed to evaluate a broader range of physiological, behavioral, and subjective markers of sleepiness and driving performance.

1. Subjective sleepiness assessments:

- **Situational Sleepiness:** Participants will complete the Karolinska Sleepiness Scale (KSS) questionnaire immediately before and after each test run (at 9 a.m., 11 a.m., 1 p.m., and 3 p.m.) to assess momentary fluctuations in alertness [\[30,31\]](#).

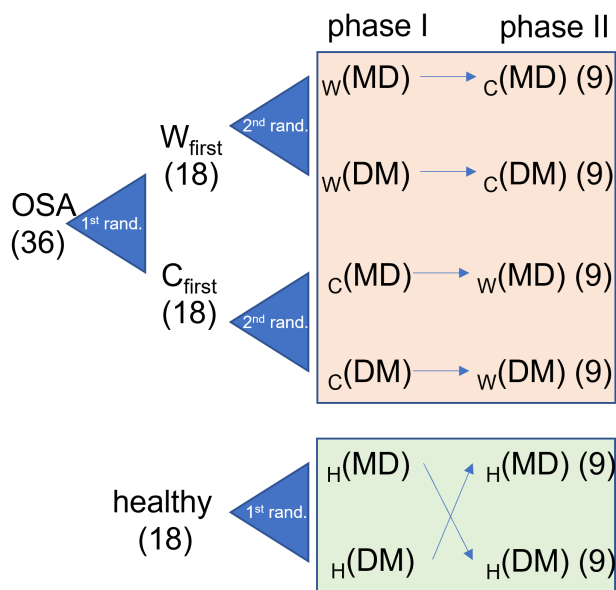


Fig 2. Overall randomization scheme for the intervention group and the comparison group. W_{first} = Group starting with the sequence in the condition “withdrawal”, i.e., the intervention. C_{first} = Group starting with the sequence in the condition “continuation”. D = driving simulation-based maintenance of wakefulness test, M = maintenance of wakefulness test.

<https://doi.org/10.1371/journal.pone.0341997.g002>

- General Daytime Sleepiness: At the screening visit, participants will complete the Epworth Sleepiness Scale (ESS) to establish a baseline measure of habitual EDS [32].
- 2. Driving performance metrics: The investigators will analyze whether driving behavior parameters, including SDLP, standard deviation of velocity (SDS), distance to the leading vehicle in the car-following task (SDD), and steering wheel angle (SDSW), correlate with CPAP treatment status and subjective sleepiness scores.
- 3. EEG-based markers of sleepiness: The investigators will examine electrophysiological correlates of EDS, including MSEs [33] and EEG features associated with sleep onset and wake–sleep transition states according to AASM criteria, including alterations in alpha activity, increased theta and delta activity, as well as exploratory EEG-derived markers such as changes in the θ/α ratio [34,35]. Their association with CPAP withdrawal and subjective sleepiness (ESS, KSS) will also be assessed.
- 4. Eye-tracking correlates: The investigators will examine whether eye-tracking features (e.g., blink rate, blink duration, eye openness, gaze direction, and attention to defined areas of interest) reflect treatment condition and align with subjective and objective markers of sleepiness
- 5. Night-to-night variability in sleep architecture: The investigators will assess sleep quality parameters and variability in OSA-related PSG parameters across repeated at-home PSG measurements. Sleep quality assessment will include classical PSG-derived metrics such as total sleep time (TST), sleep efficiency (SE), sleep onset latency (SOL), wake after sleep onset (WASO), sleep stage distribution, and arousal index. OSA-related parameters, including apnea–hypoapnea index (AHI), oxygen desaturation index (ODI), mean SpO_2 , time spent below 90% oxygen saturation ($t < 90$), and arousal index, will additionally be evaluated. Correlations with behavioral and subjective outcomes will be analyzed, and CPAP device data will be used to monitor treatment adherence. [26,29].

6. Biomarker profiling: The investigators will evaluate whether previously identified proteomic and metabolomic biomarkers of sleepiness in non-invasive samples—such as saliva, exhaled breath, and dried blood spots [18]—correlate with CPAP status and subjective sleepiness ratings.

Data collection and management

Data collection. All data is recorded semi-automatically (driving data, EEG, ET) and/or without interference from the investigator. To maximize both temporal accuracy and data reliability, the study employs a dual-recording strategy. First, all multimodal data ET, EEG, and driving simulation output (SILAB)—are recorded and synchronized in real time using the Lab Streaming Layer (LSL) framework [36]. This ensures precise temporal alignment across systems during active data collection. Concurrently, each subsystem independently records high-resolution raw data in its native format, providing a redundant backup to safeguard against potential streaming failures. This local storage approach guarantees data completeness.

To promote the quality and integrity of the collected data, the study team is trained in standardized procedures for assessment and documentation. The course of data collection is documented using standardized paper-based Case Report Forms (CRFs), completed by study investigators for each participant. These forms record all trial-relevant tests and assessments. Participants are identified by a unique participant ID (PID), which is assigned sequentially. A separate subject identification log links PIDs to participant identities and tracks screening outcomes, withdrawals, and dropouts.

All entries into the CRFs are verified against the source documents, which include medical records, laboratory results, psychological assessments, and other original documentation. In cases where data are directly extracted from instruments and not transcribed into the CRF—such as EEG or eye-tracking recordings—these are clearly identified as source data. The specific location of such source data is documented in the CRF in the designated data storage log.

Data management. All aspects of data management are coordinated by the Division of Traffic Medicine, which also serves as the central coordinating body for the study. Data are handled in compliance with Swiss data protection regulations. CRFs and other study documents are stored in locked areas with access limited to authorized personnel. Biological samples are identified solely by PID, stored under secure conditions, and transported personally by authorized staff to the designated analysis laboratory with authorized restricted access. No genetic data are collected, and no biobank will be established.

Data quality assurance procedures include double data entry, automated error-check algorithms, manual data audits after each experimental session, and automated post-acquisition audits after completion of all study visits to identify missing or inconsistent data.

In case of participant's discontinuation, data and samples collected up to the point of withdrawal may be retained and analyzed. If a participant leaves early, no additional final visit is planned, but support is provided for a safe departure if needed.

After study completion, all study documents and data are archived for 20 years at the Institute for Forensic Medicine, University of Zurich. This includes CRFs, consent forms, medical records, and laboratory results. Data analysis will be performed only after trial completion and independently of funding sources. Data will be coded, processed, and analyzed in accordance with the predefined statistical analysis plan.

Trial instruments and measures. For the primary outcome, EEG recordings are required to determine MWT-latencies, based on precise identification of sleep onset (SO). According to AASM classification, SO is defined as the start of the first epoch scored as any stage other than stage W (wake). In most individuals, this will usually be the first epoch of stage N1, characterized by SEM (slow eye movements) and LAMF (low-amplitude, mixed-frequency) in EEG activity [29]. For the calculation of DS-MWT-derived latencies, the classical EEG-based SO is complemented by the capture of SA in a driving simulator [5].

As for secondary outcome measures:

1. In addition to determining primary sleep latency outcomes, electrophysiological data collected during both MWT and DS-MWT sessions will provide insights into the dynamics of the sleep-wake transition. Of particular interest are MSEs—brief intrusions of sleep-like EEG activity during otherwise wakeful states—which may precede or occur independently of formal SO. Automatically detected MSEs will be analyzed in this study to characterize their distribution, duration, and relationship to subjective and behavioral sleepiness [33].

All EEG recordings, accompanied with electrooculogram (EOG), electromyogram (EMG), and electrocardiogram (ECG)—including those conducted during in-laboratory test sessions and overnight at-home PSG—will be performed using the Nox A1 PSG System (Nox Medical, Reykjavík, Iceland). This device has been validated for clinical sleep research and supports high-fidelity multi-channel EEG acquisition in both ambulatory and controlled lab environments [37]. Recordings will be analyzed using the dedicated Noxturnal software (Noxturnal, Nox Medical, Reykjavík, Iceland) and custom analysis scripts implemented in MATLAB (Version R2023a, The MathWorks, Inc., Natick, MA, USA).

2. Driving performance will be assessed using a custom-modified full-chassis BMW i3 simulator, engineered to deliver a high degree of ecological validity. The simulator is equipped with actuator systems replacing standard shock absorbers to simulate realistic motion cues, including acceleration, deceleration, road texture, and cornering forces. The vehicle is embedded within a 270-degree projection system offering 360-degree field-of-view coverage, thereby immersing participants in a virtual driving environment that closely mimics real-world conditions.

The simulated environment is powered by SILAB Driving Simulation Software (Scenario Package “Driver Fitness and Ability,” Version 7.0; Würzburg Institute for Traffic Sciences GmbH, Veitshöchheim, Germany), which delivers a suite of validated scenario modules for evaluating FTD [38]. SILAB records a broad set of vehicle dynamics and behavioral parameters, including SDLP, SDS, number of lane departures, etc.

Participants will perform a customized driving task, which consists of an approximately 40-minute, highly monotonous car-following task in a simulated nighttime setting, chosen to increase the likelihood of drowsiness-related events. The scenario involves continuous driving without stops, traffic lights, or turns on a two-lane motorway separated by a double solid centerline, with low oncoming traffic on the left. This low-stimulation environment requires sustained attention and minimal steering correction, closely mimicking real-world conditions where sleep-related driving incidents are most likely to occur.

3. Subjective sleepiness will be evaluated using computer-based questionnaires (KSS, ESS), administered and scored according to validated standard procedures.
4. Eye-tracking metrics will be continuously recorded during DS-MWT and MWT sessions using high-resolution eye-tracking systems Tobii Pro Glasses 3 (Tobii AB, Danderyd/Stockholm, Sweden). A range of ocular motor features, including blink rate, saccade latency, saccade peak velocity and smooth pursuit velocity gain, will be analyzed. These parameters will be examined to determine whether they reflect levels of sleepiness and whether they hold predictive value for SO or MSE during the driving simulation [39,40].
5. Unstimulated oral fluid will be collected using neutral Salivette devices (Sarstedt, Sevelen, Germany), following standardized procedures for non-invasive saliva sampling [41]. Samples will be immediately frozen and stored at -80°C until the conclusion of the trial. After completion of data collection, samples will undergo targeted metabolomic analysis using ultra-performance liquid chromatography coupled to high-resolution mass spectrometry (UPLC-HRMS) [19].
6. Dried blood spot (DBS) samples will be collected using MICROLET lancets (Bayer HealthCare LLC, Whippany, NJ, USA) applied onto DBS collection cards (CAMAG AG, Muttenz, Switzerland), which allow for precise, standardized

blood volume collection. All DBS samples will be stored under controlled conditions and subsequently processed and analyzed analogously to oral fluid samples.

7. Exhaled breath samples will be collected using MIR Spirobank Smart spirometer (MIR, Rome, Italy) and the Breath-Explor device (Munkplast AB, Uppsala, Sweden). These devices are designed for safe, efficient collection of aerosol particles from the exhaled air, suitable for subsequent biochemical and metabolomic analyses.

Statistical methods

There is no dedicated statistician other than the investigators involved. However, a statistician will be available throughout the study and subsequent analyses.

Sample size calculation. The total sample size is $N_{\text{total}} = 54$, consisting of the intervention group $N_i = 36$ and a healthy comparison group $N_h = 18$.

Sample size of $N_i = 36$ for the intervention group was estimated according Formula (1), taking into account the adapted randomized crossover RCT design [42]. The study involves a cohort of OSA patients randomized to both the CPAP starting condition and the measurement sequence, resulting in four groups (W-DM, W-MD, C-DM, C-MD) across two phases. A mixed-model ANOVA will be used for data analysis, as it appropriately accounts for the correlation of repeated measures within individuals, handles missing data more effectively, and incorporates both fixed and random effects.

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \cdot 2\sigma^2(1-\rho)}{\Delta^2} \quad (1)$$

Where,

- n = Required number of participants per group.
- $Z_{1-\alpha/2}$ = Z-value corresponding to the chosen significance level (e.g., 1.645 for $\alpha = 0.1$ two-sided) [43].
- $Z_{1-\beta}$ = Z-value corresponding to the desired power (e.g., 1.28 for 90% power) [43].
- σ = Standard deviation of the measurements = 13.5 minutes (from Bonnet & Arand [16]).
- ρ = Intra-class correlation (ICC) = 0.5, i.e., the correlation between repeated measures within the same participant.
- Δ = Equivalence margin or effect size = 1 pooled SD = 13.5 minutes (from Bonnet & Arand [16]).

For the sample size calculations, a two-sided $\alpha = 0.1$ was used, assuming that a potential difference between M and D would have a clear direction and a maximal value (bounded test range) of 40 min for the test latency. When conducting an equivalence test, using a 0.1 two-sided alpha-level effectively matches the 0.05 significance level for each side of the distribution (one for each tail), aligning the rigor of the test with the traditional hypothesis test for no difference. This assumption holds, as the reference standard deviation of 13.5 min was obtained with an equivalently bounded test range.

A graphically extracted and pooled standard deviation from Bonnet & Arand [16] (Fig 2, baseline and wake conditions, p. 388) was used, representing the only available data approximating repeated MWT test–retest measurements. One standard deviation was additionally defined as the equivalence margin. Sample sizes were calculated for both 80% and 90% power assuming a mixed-model ANOVA analysis, which allows pooling of variances in within-subject designs and thereby reduces required sample sizes. An interim analysis will be performed at the sample size corresponding to 80% power to evaluate study feasibility and reassess assumptions regarding the intra-class correlation (ρ). A conservative within-subject correlation of $\rho = 0.5$ was prespecified; however, a higher observed intra-class correlation would indicate lower within-subject variability across conditions and may reduce the required final sample size.

Similarly, a sample size of $N_H = 18$ for the comparison group (healthy) was estimated for the comparison group of healthy participants, based on a crossover design without retest, as applied in the test-retest framework. This corresponds to 9 participants per group to achieve 90% power, assuming a two-way ANOVA (see Formula (2), adapted from Chow & Liu [44]), resulting in a total of 18 participants across both groups. Although a smaller standard deviation ($\sigma = 5.9$ minutes) was used — derived from normative data on healthy individuals reported by Banks et al. [12] — the required sample size remains comparable to that of the OSA cohort. This is primarily due to the absence of additional variability-reducing conditions that were present in the OSA sample.

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \cdot 2\sigma^2}{\Delta^2} \quad (2)$$

Where,

- n = Required number of participants per group.
- $Z_{1-\alpha/2}$ = Z-value corresponding to the chosen significance level (e.g., 1.645 for $\alpha = 0.1$ two-sided) [43].
- $Z_{1-\beta}$ = Z-value corresponding to the desired power (e.g., 1.28 for 90% power) [43].
- σ = Standard deviation of the measurements = 13.5 minutes (from Bonnet & Arand [16]).
- ρ = Intra-class correlation (ICC), i.e., the correlation between repeated measures within the same participant.
- Δ = Equivalence margin or effect size = 1 pooled SD = 13.5 minutes (from Bonnet & Arand [16]).

Statistical analysis plan. Based on the design of this prospective, randomized, crossover trial, statistical analysis will be performed to detect both intra- and inter-individual (i.e., individual and group-level) differences. All analyses will be conducted in relation to two key contextual variables:

- Sleep quality of the night preceding each test session, as assessed by PSG, and
- Time of day at which each test run was performed (9 a.m., 11 a.m., 1 p.m., 3 p.m.).

The primary hypothesis will be tested solely based on the determined mean sleep latencies obtained from the MWT and DS-MWT, without incorporating additional secondary or exploratory parameters. For the primary outcome analysis, linear mixed-effects models (LMMs) will be applied, with test latency as the dependent variable and test condition (MWT vs. DS-MWT) as the main fixed effect. Interindividual variability will be accounted for by including participant as a random effect. Additional covariates, including previous-night PSG-derived sleep quality parameter, test sequence, and time of day, will also be included in the models. To account for potential non-linear variation across the day, time of day will be treated as a categorical factor. Possible interactions between condition and time of day will be explored, based on the assumption that time-of-day-related fluctuations in behavioral performance and sleepiness may differentially affect test latency depending on the testing modality. Results will be considered significant if $p < 0.05$.

All secondary variables will be analyzed post hoc, with the goal of generating new hypotheses rather than performing formal confirmatory tests. These analyses will include individual evaluations of each collected secondary parameter in reference to:

- The experimental condition (MWT vs. DS-MWT),
- The treatment state (CPAP continuation vs. withdrawal) or healthy comparison,
- The test sequence (DM vs. MD).

Where appropriate, group-level comparisons will be followed by multiple comparison tests to examine differences in greater detail. For example, time-course effects in driving behavior parameters—such as SDLP—will be examined using repeated measures ANOVA across ten 4-minute segments within each test session, or through paired t-tests comparing the first and final modules of each session.

For these secondary exploratory analyses, linear mixed models will be used as the initial framework, with stepwise model testing based on biologically plausible relationships rather than relying solely on statistical significance. Where applicable, Akaike Information Criterion (AIC) will be used to identify the most parsimonious model, and AICc will be applied in cases involving small sample sizes ($N = 18$). Cross-validation techniques, such as k-fold cross-validation, will be implemented to evaluate predictive performance and reduce overfitting.

Beyond aggregated latency measures, predictive models for SO and/or SA will be explored at finer temporal resolutions, including:

- Per 40-minute run ($N = 4 \text{ runs} \times 54 \text{ participants} = 216 \text{ observations}$),
- Per 4-minute driving module ($N = 10 \text{ modules} \times 216 \text{ runs} = 2,160 \text{ observations}$),
- Or via dynamic time-series analysis using shorter time units still to be defined.

These exploratory analyses will aim to maintain an events-per-variable (EPV) ratio between 10 and 15, depending on the number of predictors included. These models will evaluate whether secondary variables collected during each run can predict the occurrence of SO or MSE in real time.

While p-values will not be the primary decision metric for exploratory models, False Discovery Rate (FDR) corrections may be applied to account for multiple testing while preserving statistical power. Bonferroni corrections, due to their conservativeness, will be avoided to reduce the risk of Type II errors (false negatives). Significant findings from exploratory analyses will be interpreted cautiously and treated as hypothesis-generating, requiring independent replication in future studies.

Although our study is not powered to detect gender differences, sex/gender will be included as a covariate in all relevant models for control purposes.

All statistical analyses will be performed using R, Python, MatLab and GraphPad Prism. Any deviations from this pre-defined statistical analysis plan will be documented and justified in the final trial report.

The study is intended to be completed with $N = 54$ participants, unless recruitment proves infeasible within the projected time frame. The last participant's last visit (LPLV) is planned for 30 April 2027, though an extension of the recruitment period may be considered if necessary.

Missing datasets due to participants' withdrawal from the trial will not be multiply imputed due to the low sample size and inherently low statistical power. All available data will be used, if possible, for statistical analysis. Drop-outs during the study phase will be replaced by recruitment of new participants from the accessible participant pool, if possible.

Administrative information

Monitoring. Internal monitoring duties will be carried out by Dr. med. Kristina Keller, who is not actively taking part in recruitment, data acquisition or data analysis. There is no additional external monitoring foreseen or required. Internal monitoring will be conducted to ensure adherence to the study protocol, regulatory requirements, and data quality standards. All source data and relevant documents will be accessible to the internal monitor. Monitoring visits may take place without prior notice and will include a review of available documentation and clarification of any issues directly with the site team.

Given the exploratory nature of this study, the establishment of a steering committee, an endpoint adjudication committee, or an independent data monitoring committee is not required and has therefore not been implemented.

Discussion

The classical maintenance of wakefulness test (MWT) is a widely accepted method for assessing EDS across various clinical conditions, although its broader applicability remains debated [4]. The MWT evaluates the ability to remain awake in a quiet, non-stimulating environment [45]. It is often used to assess the presence of EDS, to monitor treatment response in sleep disorders, and to evaluate wakefulness in individuals whose sustained alertness is safety-critical [14]. While valuable in this context, the MWT may not fully capture the challenges of driving, which involves cognitive engagement and environmental stimulation. This protocol seeks to establish a more ecologically valid proxy for real-world driving risks than conventional MWT. It remains to be examined whether DS-MWT latencies differ significantly from MWT latencies. Such a finding would support the view that traditional MWT alone may under- or overestimate real-life driving impairment in sleepy individuals. This could inform clinical and regulatory practice regarding FTD assessments in patients with EDS or sleep disorders, potentially leading to a meaningful refinement in how sleepiness-related driving risk is evaluated in traffic medicine.

By embedding the wakefulness assessment in a realistic driving simulation — a monotonous, nighttime, car-following task — the driving simulation-based maintenance of wakefulness test (DS-MWT) may capture subtle impairments in vigilance, attention maintenance, and vehicle control that traditional MWT, lacking behavioral and environmental context, might miss.

Moreover, the extended set of secondary outcomes, including driving behavior metrics (e.g., SDLP, SDS, lane departures, etc.), eye-tracking and EEG parameters, subjective sleepiness ratings, polysomnographic sleep quality data, and biomarker analyzes (saliva, exhaled breath, and DBS), offers the possibility for a multimodal characterization of sleepiness and alertness. If associations emerge among these parameters in this naturalistic test setting, they may suggest objective, sensitive markers of sleepiness that better predict risk of impaired driving than classical latency measures alone. Ultimately, this could support the development of novel, non-invasive screening tools for real-world drowsy driving risk, benefiting public health and potentially reducing preventable traffic accidents. In addition, by using a within-subject crossover design with CPAP withdrawal and resumption, the study may yield insights into the dynamics of EDS and its reversibility under therapeutic intervention. This could reinforce the importance of CPAP adherence in OSA management.

Limitations

A major strength of the design is the within-subject crossover format, which controls for inter-individual variability in sleep physiology, OSA severity, medication status, and other confounders. Inclusion of a well-adherent OSA patient pool, pre-screened for comorbidities, ensures clinical safety and homogeneity. However, several limitations are clearly present. First, blinding is not feasible for participants or study personnel due to the nature of the intervention (CPAP withdrawal and simulation vs. classical MWT). Although outcome assessors will be blinded where possible and data acquisition is automated, some bias cannot be excluded — for example, participants' awareness of driving simulation might influence their subjective sleepiness ratings, motivation, or compensation behaviors. This may limit the extent to which DS-MWT findings can be generalized to unmonitored, real-life driving situations where motivation, distraction, or compensatory behaviors differ.

Second, the healthy control group is relatively small and was not recruited using strict matching criteria with the clinical OSA cohort, which limits direct comparability between groups despite planned statistical adjustments. However, the primary focus of the study is the comparison between the two testing modalities (MWT vs. DS-MWT) rather than direct between-group comparisons.

Third, the multiplicity of secondary and exploratory analyses increases the risk of false-positive findings or overfitting if included in one model. While the planned use of model-selection criteria (AIC/AICc), cross-validation, and FDR correction mitigates this risk, such methods do not fully eliminate it — and any promising associations will require independent

replication in larger, prospective cohorts. Therefore, the results of this study can provide a foundation for further studies evaluating the DS-MWT in clinically and demographically diverse populations — including CPAP-naïve individuals, patients with other sleep disorders, shift workers, and professional drivers. Moreover, if certain biomarkers or eye-tracking / EEG features prove to correlate robustly with simulated driving impairment, future work could aim to develop non-invasive screening tools for use in clinical, occupational, or roadside screening settings. This could be complemented by machine-learning approaches trained on multi-modal data to predict drowsiness or microsleep risk in individual drivers. In conclusion, this trial might bridge existent gaps in sleep and traffic medicine by evaluating a novel, ecologically valid simulation-based assessment of wakefulness and driving performance in OSA patients and healthy controls. In conclusion, by combining classical neurophysiological measures, behavioral simulation, subjective sleepiness scales, and multi-modal biomarker sampling, the study has the potential to significantly advance our understanding of how EDS translates into driving risks.

Supporting information

S1 File. SPIRIT 2025 checklist.

(DOCX)

S2 File. DS-MWT2 study protocol approved by the ethics committee.

(DOCX)

Acknowledgments

We thank Lana Brockbals and Michael Scholz (Forensic Pharmacology & Toxicology, Institute of Forensic Medicine, University of Zurich) for the discussion on sample taking and future analyses.

Author contributions

Conceptualization: Esther Irene Schwarz, Stefan Lakämper.

Data curation: Veronika Gambin, Nan Li.

Funding acquisition: Stefan Lakämper.

Investigation: Veronika Gambin, Nan Li, Stefan Lakämper.

Methodology: Veronika Gambin, Nan Li, Stefan Lakämper.

Project administration: Veronika Gambin, Nan Li, Esther Irene Schwarz, Stefan Lakämper.

Resources: Esther Irene Schwarz, Stefan Lakämper.

Software: Nan Li.

Supervision: Kristina Keller.

Writing – original draft: Veronika Gambin.

Writing – review & editing: Veronika Gambin, Nan Li, Esther Irene Schwarz, Kristina Keller, Stefan Lakämper.

References

1. Ward KL, Hillman DR, James A, Bremner AP, Simpson L, Cooper MN, et al. Excessive daytime sleepiness increases the risk of motor vehicle crash in obstructive sleep apnea. *J Clin Sleep Med*. 2013;9(10):1013–21. <https://doi.org/10.5664/jcsm.3072> PMID: [24127145](https://pubmed.ncbi.nlm.nih.gov/24127145/)
2. McNicholas WT, Rodenstein D. Sleep apnoea and driving risk: the need for regulation. *Eur Respir Rev*. 2015;24(138):602–6. <https://doi.org/10.1183/16000617.0049-2015> PMID: [26621974](https://pubmed.ncbi.nlm.nih.gov/26621974/)
3. Levine JD. A road to injustice paved with good intentions: Maggie's misguided crackdown on drowsy driving. *Hastings LJ*. 2004;56:1297.

4. McCall CA. Moving towards a more naturalistic approach to evaluating drowsy driving risk. *Sleep*. 2024;47(8):zsae139. <https://doi.org/10.1093/sleep/zsae139> PMID: [38908012](#)
5. Pisteljic M, Keller K, Lakämper S. Capturing sleep accidents in driving simulation as a promising tool to assess excessive daytime sleepiness with high ecological validity—a pilot study. *Sleep*. 2024;47(8):zsae110. <https://doi.org/10.1093/sleep/zsae110> PMID: [38758109](#)
6. Gandhi KD, Mansukhani MP, Silber MH, Kolla BP. Excessive daytime sleepiness: a clinical review. *Mayo Clin Proc*. 2021;96(5):1288–301. <https://doi.org/10.1016/j.mayocp.2020.08.033> PMID: [33840518](#)
7. Pagel JF. Excessive daytime sleepiness. *Am Fam Physician*. 2009;79(5):391–6. PMID: [19275068](#)
8. Patel SR, White DP, Malhotra A, Stanchina ML, Ayas NT. Continuous positive airway pressure therapy for treating sleepiness in a diverse population with obstructive sleep apnea: results of a meta-analysis. *Arch Intern Med*. 2003;163(5):565–71. <https://doi.org/10.1001/archinte.163.5.565> PMID: [12622603](#)
9. Lal C, Weaver TE, Bae CJ, Strohl KP. Excessive daytime sleepiness in obstructive sleep apnea. mechanisms and clinical management. *Ann Am Thorac Soc*. 2021;18(5):757–68. <https://doi.org/10.1513/AnnalsATS.202006-696FR> PMID: [33621163](#)
10. Garbarino S, Durando P, Guglielmi O, Dini G, Bersi F, Fornarino S, et al. Sleep apnea, sleep debt and daytime sleepiness are independently associated with road accidents. A cross-sectional study on truck drivers. *PLoS One*. 2016;11(11):e0166262. <https://doi.org/10.1371/journal.pone.0166262> PMID: [27902703](#)
11. Terán-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. *N Engl J Med*. 1999;340(11):847–51. <https://doi.org/10.1056/NEJM199903183401104>
12. Banks S, Barnes M, Tarquinio N, Pierce RJ, Lack LC, McEvoy RD. Factors associated with maintenance of wakefulness test mean sleep latency in patients with mild to moderate obstructive sleep apnoea and normal subjects. *J Sleep Res*. 2004;13(1):71–8. <https://doi.org/10.1111/j.1365-2869.2003.00383.x> PMID: [14996038](#)
13. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF, New Collective Author. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep*. 2003;26(2):117–26. <https://doi.org/10.1093/sleep/26.2.117> PMID: [12683469](#)
14. Krahn LE, Arand DL, Avidan AY, Davila DG, DeBassio WA, Ruoff CM, et al. Recommended protocols for the multiple sleep latency test and maintenance of wakefulness test in adults: guidance from the American Academy of Sleep Medicine. *J Clin Sleep Med*. 2021;17(12):2489–98. <https://doi.org/10.5664/jcsm.9620> PMID: [34423768](#)
15. Mayer G. The need for new instruments to measure the manifold aspects of excessive daytime sleepiness. *Sleep*. 2022;45(8):zsac144. <https://doi.org/10.1093/sleep/zsac144> PMID: [35750025](#)
16. Bonnet MH, Arand DL. Impact of motivation on multiple sleep latency test and maintenance of wakefulness test measurements. *J Clin Sleep Med*. 2005;1(4):386–90. <https://doi.org/10.5664/jcsm.26367> PMID: [17564407](#)
17. Garbarino S. Excessive daytime sleepiness in obstructive sleep apnea: implications for driving licenses. *Sleep Breath*. 2020;24(1):37–47. <https://doi.org/10.1007/s11325-019-01903-6> PMID: [31342234](#)
18. Scholz M, Lakaemper S, Keller K, Dobay A, Steuer AE, Landolt H-P, et al. Metabolomics-based Sleepiness Markers for Risk Prevention and Traffic Safety (ME-SMART): a monocentric, controlled, randomized, crossover trial. *Trials*. 2023;24(1):131. <https://doi.org/10.1186/s13063-023-07154-x> PMID: [36810100](#)
19. Bjørkum AA, Griebel L, Birkeland E. Human serum proteomics reveals a molecular signature after one night of sleep deprivation. *Sleep Adv*. 2024;5(1):zpae042. <https://doi.org/10.1093/sleepadvances/zpae042> PMID: [39131770](#)
20. Pizza F, Contardi S, Mondini S, Trentin L, Cirignotta F. Daytime sleepiness and driving performance in patients with obstructive sleep apnea: comparison of the MSLT, the MWT, and a simulated driving task. *Sleep*. 2009;32(3):382–91. <https://doi.org/10.1093/sleep/32.3.382> PMID: [19294958](#)
21. Philip P. Excessive daytime sleepiness versus sleepiness at the wheel, the need to differentiate global from situational sleepiness to better predict sleep-related accidents. *Sleep*. 2023;46(11):zsad231. <https://doi.org/10.1093/sleep/zsad231> PMID: [37668542](#)
22. Horne J, Reyner L. Vehicle accidents related to sleep: a review. *Occup Environ Med*. 1999;56(5):289–94. <https://doi.org/10.1136/oem.56.5.289>
23. Kohler M, Stoewhas A-C, Ayers L, Senn O, Bloch KE, Russi EW, et al. Effects of continuous positive airway pressure therapy withdrawal in patients with obstructive sleep apnea: a randomized controlled trial. *Am J Respir Crit Care Med*. 2011;184(10):1192–9. <https://doi.org/10.1164/rccm.201106-0964OC> PMID: [21836134](#)
24. Schwarz EI, Stradling JR, Kohler M. Physiological consequences of CPAP therapy withdrawal in patients with obstructive sleep apnoea—an opportunity for an efficient experimental model. *J Thorac Dis*. 2018;10(Suppl 1):S24–32. <https://doi.org/10.21037/jtd.2017.12.142> PMID: [29445525](#)
25. Roeder M, Sievi NA, Kohler M, Schwarz EI. Predictors of changes in subjective daytime sleepiness in response to CPAP therapy withdrawal in OSA: a post-hoc analysis. *J Sleep Res*. 2021;30(2):e13078. <https://doi.org/10.1111/jsr.13078> PMID: [32441873](#)
26. Roeder M, Bradicich M, Schwarz EI, Thiel S, Gaisl T, Held U, et al. Night-to-night variability of respiratory events in obstructive sleep apnoea: a systematic review and meta-analysis. *Thorax*. 2020;75(12):1095–102. <https://doi.org/10.1136/thoraxjnl-2020-214544> PMID: [32792405](#)
27. Schwarz EI, Martinez-Lozano Sinues P, Bregy L, Gaisl T, Garcia Gomez D, Gaugg MT, et al. Effects of CPAP therapy withdrawal on exhaled breath pattern in obstructive sleep apnoea. *Thorax*. 2016;71(2):110–7. <https://doi.org/10.1136/thoraxjnl-2015-207597> PMID: [26671307](#)

28. Chan A-W, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. *BMJ*. 2025;389:e081477. <https://doi.org/10.1136/bmj-2024-081477> PMID: [40294953](https://pubmed.ncbi.nlm.nih.gov/40294953/)
29. Berry RB, Brooks R, Gamaldo CE, Harding SM, Lloyd RM, Quan SF. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. 2017.
30. Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. *Int J Neurosci*. 1990;52(1–2):29–37. <https://doi.org/10.3109/00207459008994241> PMID: [2265922](https://pubmed.ncbi.nlm.nih.gov/2265922/)
31. Kaida K, Takahashi M, Akerstedt T, Nakata A, Otsuka Y, Haratani T, et al. Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clin Neurophysiol*. 2006;117(7):1574–81. <https://doi.org/10.1016/j.clinph.2006.03.011> PMID: [16679057](https://pubmed.ncbi.nlm.nih.gov/16679057/)
32. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540–5. <https://doi.org/10.1093/sleep/14.6.540> PMID: [1798888](https://pubmed.ncbi.nlm.nih.gov/1798888/)
33. Skorucak J, Hertig-Godeschalk A, Achermann P, Mathis J, Schreier DR. Automatically detected microsleep episodes in the fitness-to-drive assessment. *Front Neurosci*. 2020;14. <https://doi.org/10.3389/fnins.2020.00008>
34. Tracey B, Culp M, Fabregas S, Mignot E, Buhl DL, Volfson D. Novel biomarkers derived from the maintenance of wakefulness test as predictors of sleepiness and response to treatment. *Sleep*. 2024;47(12):zsae148. <https://doi.org/10.1093/sleep/zsae148> PMID: [38954525](https://pubmed.ncbi.nlm.nih.gov/38954525/)
35. Chatburn A, Lushington K, Cross ZR. Considerations towards a neurobiologically-informed EEG measurement of sleepiness. *Brain Res*. 2024;1841:149088. <https://doi.org/10.1016/j.brainres.2024.149088> PMID: [38879143](https://pubmed.ncbi.nlm.nih.gov/38879143/)
36. Stenner T, Boulay C, Grivich M, Medine D, Kothe C, Tobiascherzke, et al. Scnn/liblsl: V1.16.2. Zenodo; 2023.
37. Medical N. Nox A1s: Sleep Recorder — User Manual. 2025.
38. Krueger H, Grein M, Mark C. SILAB—A Task Oriented Driving Simulation. In: *Proceedings of the Driving Simulator Conference (DSC)*. Orlando, FL; 2005. pp. 323–31.
39. Aitken B, Downey LA, Rose S, Arkell TR, Shiferaw B, Hayley AC. Driving performance and ocular activity following acute administration of 10 mg methylphenidate: A randomised, double-blind, placebo-controlled study. *J Psychopharmacol*. 2024;38(11):998–1006. <https://doi.org/10.1177/02698811241286715> PMID: [39394668](https://pubmed.ncbi.nlm.nih.gov/39394668/)
40. Roy D, Nah FF. A review on eye-tracking metrics for sleepiness. In: Nah FF, Siau K, editors. *HCI International 2020—Late Breaking Papers: Interaction, Knowledge and Social Media*. Cham, Switzerland: Springer; 2020. pp. 575–89.
41. Kaufman E, Lamster IB. The diagnostic applications of saliva—a review. *Crit Rev Oral Biol Med*. 2002;13(2):197–212. <https://doi.org/10.1177/154411130201300209> PMID: [12097361](https://pubmed.ncbi.nlm.nih.gov/12097361/)
42. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge; 1988. Available from: <https://doi.org/10.4324/9780203771587>
43. Moore DS, McCabe GP, Craig BA. *Introduction to the Practice of Statistics*. W. H. Freeman; 2016. Available from: <https://books.google.ch/books?id=pp-jDgAAQBAJ>
44. Chow SC, Liu JP. *Design and analysis of bioavailability and bioequivalence studies*. CRC Press; 1999.
45. Arand D, Bonnet M, Hurwitz T, Mitler M, Rosa R, Sangal RB. The clinical use of the MSLT and MWT. *Sleep*. 2005;28(1):123–44. <https://doi.org/10.1093/sleep/28.1.123> PMID: [15700728](https://pubmed.ncbi.nlm.nih.gov/15700728/)