

Residual effects of medications for sleep disorders on driving performance: a systematic review and network meta-analysis of randomized controlled trials

"NMA driving and hypnotics"

Michele Fornaro^{1#*}, Claudio Caiazza^{1#}, Flavia Rossano¹, Flavia Cilmi¹, Michele De Prisco^{2,3,4}, Eduard Vieta^{2,3,4}, Trevor Thompson⁵, Marco Solmi^{6,7,8,9}, Andre F Carvalho¹⁰, Felice Iasevoli^{11,12}, Andrea de Bartolomeis^{11,12}

***Corresponding author:** Michele Fornaro, MD. Address: Via Sergio Pansini n.5, Università di Napoli Federico II, Edificio 18, Psichiatria, Napoli ZIP 80131, Italy. Email: dott.fornaro@gmail.com

#Joint co-first.

1 Clinical Section of Psychiatry and Psychology - Department of Neuroscience, Reproductive Sciences, and Odontostomatology, University School of Medicine Federico II, Naples, Italy

2 Bipolar and Depressive Disorders Unit, Hospital Clinic de Barcelona. C. Villarroya, 170, 08036 Barcelona, Spain.

3 Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), c. Villarroya, 170, 08036 Barcelona, Spain.

4 Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain

5 Centre for Chronic Illness and Ageing, University of Greenwich, London, England

6 Department of Psychiatry, University of Ottawa, Ontario, Canada.

7 Regional Centre for the Treatment of Eating Disorders and On Track: The Champlain First Episode Psychosis Program, Department of Mental Health, The Ottawa Hospital, Ontario, Canada.

8 Ottawa Hospital Research Institute (OHRI) Clinical Epidemiology Program University of Ottawa, Ottawa, Ontario.

9 Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany

10 Innovation in Mental and Physical Health and Clinical Treatment (IMPACT) Strategic Research Centre, School of Medicine, Barwon Health, Deakin University, Geelong, VIC, Australia.

11 Unit of Treatment-Resistant Psychosis, Section of Psychiatry, Department of Neuroscience, Reproductive Science and Odontostomatology, University School of Medicine of Naples Federico II, Naples, Italy.

12 Laboratory of Molecular and Translational Psychiatry, University School of Medicine of Naples Federico II, Naples, Italy.

Abstract

Sleep medications often carry residual effects potentially affecting driving safety, warranting network meta-analysis (NMA). PubMed/EMBASE/TRID/Clinicaltrials.gov/WHO-ICTRP/WebOfScience were inquired for randomized controlled trials of hypnotic driving studies in persons with insomnia and healthy subjects up to 05/28/2023, considering the vehicle's standard deviation of lateral position - SDLP (Standardized Mean Difference/SMD) and driving impairment rates on the first morning (co-primary outcomes) and endpoint. Risk-of-bias, global/local inconsistencies were measured, and CINeMA was used to assess the confidence in the evidence. Of 4,805 identified records, 26 cross-over RCTs were included in the systematic review, of which 22 entered the NMA, focusing on healthy subjects only. After a single administration, most molecules paralleled the placebo, outperforming zopiclone regarding SDLP. In contrast, ramelteon 8 mg, daridorexant 100 mg, zolpidem 10 mg bedtime, zolpidem middle-of-the-night 10 mg and 20 mg, mirtazapine 15–30 mg, and triazolam 0.5 mg performed significantly worse than placebo. Lemborexant 2.5–5 mg, suvorexant 15–20 mg, and zolpidem 3.5 mg middle-of-the-night associated with lower impairment than zopiclone. Repeated administration (maximum follow-up time of ten days) caused fewer residual effects than acute ones, except for flurazepam. Heterogeneity and inconsistency were negligible. Confidence in the evidence was low/very low. Sensitivity analyses confirmed the main analyses. Most FDA-approved hypnotics overlapped placebo at in-label doses, outperforming zopiclone. Repeated administration for 15 days or less reduced residual effects, warranting further research on the matter. Funding: None.

1. Introduction

Insomnia entails a significant public health burden, affecting approximately 10 % of the adult population worldwide, with an additional 20 % experiencing transient insomnia (Morin and Jarrin, 2022). The high prevalence of insomnia encouraged the widespread prescription of sleep medications. In 2020, the National Health Interview Survey data showed that about 8.4 % of U.S. adults took sleep medications every day or most days, while 10 % took them sporadically [2]; thus, 18.4 % of the population was exposed to these drugs and their potential risks. Sedatives and hypnotics potentially carry a "hangover" effect, namely, next-day residual effects (after bedtime administration), which may cause daytime sleepiness and cognitive and psychomotor impairments (Vermeeren, 2004). The public health concern stems from the higher rate of vehicle accidents (potentially fatal) in people with sleep disorders who take sleep medications (Gustavsen et al., 2008; Hansen et al., 2015; Verster et al., 2009), which may contribute to daytime sleepiness and reduce sustained attention, thus potentially impairing driving-related skills. Similarly, poor sleep quality may impair driving skills (McManus and Stavrinou, 2021), increasing the risk of daytime accidents due to driver fatigue and sleepiness beyond medications' effects. Therefore, insomnia should be treated to minimize daytime tiredness while avoiding the driving-impairing risk of sedatives/hypnotics. Driving performance impairment is a complex task that can be reliably measured by the Standard Deviation of Lateral Position (SDLP) which is the most consistently and reliably adopted proxy measure that estimates the vehicle's lateral swerving from an ideal straight-line (Vinckenbosch et al., 2020) and is often used to examine driving ability after the administration of a sleep medication (Verster and Roth, 2011). Other proxies (e.g., Brake Reaction Time (BRT), speed changes, number of driving errors, number of collisions, speed limit infractions) can be likewise considered to appraise driving performance, yet the latter are less consistent in corresponding operational definition and less reliable, while SDLP also reliably correlates to Blood Alcohol Concentrations (BAC) (Louwerens et al., 1987) and can estimate the risk of vehicle accidents (Ramaekers, 2017).

Network meta-analysis (NMA) is a powerful statistical approach that can simultaneously compare multiple interventions (Rouse et al., 2017) by combining direct and indirect evidence within a network of Randomized Controlled Trials (RCTs) and provides exploratory estimates of comparisons between interventions that have not been directly compared in single RCTs (Mavridis, 2019).

The present study aims to investigate the next-day potential driving impairment of acute and repeated bedtime administration of sedative/hypnotic drugs in persons with insomnia and healthy

subjects through an NMA within the frequentist framework, considering outcomes of SDLP and associated measures and placebo as a negative control. Such data can be critical to increase awareness about the potential driving influence of sleep medications and can assist clinicians in the prescription process.

2. Materials and methods

The present systematic review and NMA was performed according to the 2020 edition of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021), following an a priori designed protocol registered in PROSPERO (CRD42023430416). Subsequent protocol amendments are presented in Appendix S7.

2.1. Literature search

We systematically searched PubMed, EMBASE, Transport Research International Documentation (TRID), ClinicalTrials.gov, WHO-ICTRP, and Clarivate Web of Science databases from inception until May 28th, 2023. This search strategy was augmented by searching the reference lists of included studies and previous reviews (Roth et al., 2014; Verster et al., 2006). Appendix S1 provides full details on the search strategy. Two authors (CC, FC) independently screened the papers, and a third (MF) resolved eventual discrepancies.

2.2. Inclusion and exclusion criteria

Inclusion criteria were: i) randomized controlled studies (RCTs); ii) participants with a sleep disorder diagnosis (considering a valid assessment tool – e.g., DSM, any edition) or healthy participants; (iii) studies performing *next-morning* driving tests; both on-the-road and simulated driving (after bedtime or middle-of-the-night drug administration), iv) studies employing any FDA-approved hypnotic for sleep disorders (i.e., triazolam, estazolam, temazepam, flurazepam, quazepam, secobarbital, butabarbital, zolpidem, eszopiclone, zaleplon, ramelteon, tasimelteon, suvorexant, lemborexant, daridorexant, doxylamine, diphenhydramine, doxepin) (Food and Administration, 2021), v) drugs frequently employed in sleep disorders but not FDA-approved (i.e., melatonin, zopiclone, trazodone, mirtazapine, hydroxyzine, quetiapine), vi) reporting information on lateral position deviation.

Exclusion criteria were: i) subjects with significant daytime wakefulness-impairing diagnoses (i.e., narcolepsy, obstructive or central sleep apnea, idiopathic hypersomnia, alcohol use disorder, sedative, hypnotic, or anxiolytic use disorder, traumatic brain injury), ii) augmentation therapies, iii)

post-dose driving test, iv) quasi-randomized/non-randomized designs, v) retrospective studies, vi) case reports/series. No restrictions in sample size, trial duration, or language were applied.

Cross-over trials were included in the analyses only when period bias and carry-over effects were absent. Endpoint effects were considered. Only eligible arms were included when a study performed three or more arms; one did not match the inclusion criteria.

2.3. Data extraction

Two independent authors (CC, FC) identified and extracted data from eligible trials. Inconsistencies were resolved by consulting a third one (MF). Extracted information was: author, publication year, country, study design, intervention(s) and comparator(s) with doses and administration schedule, sample size and percentage of males, age range, mean age \pm standard deviation (SD), diagnosis (i.e., sleep disorders or healthy subjects) and diagnostic assessment, setting (on-the-road or simulated driving), trial duration, sponsorship, outcome measures (either SDLP or number of events) both at morning one after bedtime administration and at trial endpoint. When data was solely reported in figures, we used WebPlotDigitizer ([Rohatgi, 2017](#)). In addition, author contact was attempted for the missing record twice.

2.4. Outcomes

Co-primary outcomes were: i) next-morning driving impairment as per SDLP (cm) after the first bedtime drug administration to assess the acute residual effect of hypnotics on driving performance; ii) driving impairment rate (rates of drivers with >2.4 cm SDLP difference from placebo, corresponding to the 'swerving' of 0.5 g/L BAC ([Louwerens et al., 1987](#)), the legal limit in several countries).

We planned the following secondary outcomes: i) co-primary outcomes evaluated at the study endpoint to assess chronic (repeated administration) hypnotic driving influence after body tolerance building to the residual effects of the drugs. Only studies (or arms) that administered the medication every night until the endpoint were included in the pertinent analyses.

2.5. Risk-of-bias evaluation

The risk of bias (RoB) was assessed using version 2 of the Cochrane risk-of-bias tool for randomized cross-over trials (RoB2) ([Sterne et al., 2020](#)).

2.6. Data analysis

We performed a random-effects NMA for each outcome within the frequentist framework. The nodes included the specific molecules and different doses, accounting for potential dose-related influence on residual effects. Only mirtazapine antidepressant dose (15–30 mg) and suvorexant similar dosages (15–20 mg and 30–40 mg) were pooled for practical reasons. Estimates of the variability proportion and the true variance's absolute value (heterogeneity) were measured using I^2 statistical test and a τ^2 statistic, respectively. We also assessed the statistical difference between direct and indirect effect sizes by evaluating the global inconsistency through a Q-statistic test (Higgins et al., 2012). Local inconsistency was measured with a loop-specific approach to assess the agreement between direct and indirect estimates (Veroniki et al., 2013). A hierarchy of the evaluated drugs was calculated for each outcome based on the cumulative ranking curve (SUCRA) (Rücker and Schwarzer, 2015).

We planned independent analyses considering on-the-road and simulated driving studies, healthy subjects, and persons with insomnia, and evaluating elder (>65y) subjects who have their own benefit-risk profile (Louzada et al., 2021). In addition, we performed the following sensitivity analyses: i) removing studies with a high RoB, ii) removing studies (or arms) with FDA-unapproved molecules/dosages for sleep disorders.

We used the standardized mean difference (SMD) to measure SDLP. We used the risk ratio (RR) for the impairment rates. All outcomes were evaluated with the corresponding 95 % confidence interval (C.I.), setting statistical significance at $p=.05$. Following the Cochrane Handbook, the missing standard deviations (SD) were imputed by computing the weighted mean of the available SDs of similar-setting studies (Higgins et al., 2022).

We inquired about publication bias and small-study effect by comparison-adjusted funnel plot visual inspection and Egger's test p-value (Lin and Chu, 2018) for all analyses relying on more than ten studies. All analyses were conducted with R version 4.3.2 (Team, 2023) based on netmeta v2.8–2 (Balduzzi et al., 2023).

2.7. Confidence in evidence

The confidence in the co-primary outcomes' evidence was assessed within the Confidence in Network Meta-Analysis (CINeMA) framework (Nikolakopoulou et al., 2020), as detailed in Appendix S6.

3. Results

We identified 4805 records from various sources, and after semi-automatic duplicate removal, we screened 3036 records at a title-abstract level. We sought 194 full texts for retrieval. One could not be retrieved (Robbe et al., 1991), therefore, 193 full texts were assessed for eligibility. Finally, the review included 27 eligible records from 26 studies (Bocca et al., 1999, 2011; Iwamoto et al., 2022, 2013; Kay et al., 2016; Leufkens et al., 2009, 2014; Leufkens and Vermeeren, 2009; Mets et al., 2011; Muehlan et al., 2022; Otmani et al., 2008; Philip et al., 2011; Ramaekers et al., 1998; Sasada et al., 2013; Simen et al., 2015; Torres et al., 2022; Vermeeren et al., 1998a, 2019, 1998b; Vermeeren et al., 2002, 2015, 2016; Vermeeren et al., 2014; Verster et al., 2002; Volkerts et al., 1992; Wingen et al., 2005). Details on the study flow can be found in the PRISMA diagram in Fig. 1. All included studies were cross-over RCTs. Sixteen studies appraised the on-the-road driving (Leufkens et al., 2009, 2014; Leufkens and Vermeeren, 2009; Mets et al., 2011; Philip et al., 2011; Ramaekers et al., 1998; Vermeeren et al., 1998a, 2019, 1998b; Vermeeren et al., 2002, 2015, 2016; Verster et al., 2002; Volkerts et al., 1992; Wingen et al., 2005), while 10 (Bocca et al., 1999, 2011; Iwamoto et al., 2022, 2013; Kay et al., 2016; Muehlan et al., 2022; Otmani et al., 2008; Roth et al., 2011; Sasada et al., 2013; Simen et al., 2015; Torres et al., 2022) assessed simulated driving. Finally, 22 studies entered the NMAs. The remaining studies did not enter the analyses due to inadequate data or potential carry-over effects (Bocca et al., 1999; Otmani et al., 2008; Philip et al., 2011; Ramaekers et al., 1998) and were included in the sole SR. Characteristics of included studies are detailed in Table 1. Excluded full-text with the reason(s) are described in Appendix S8 and Fig. 1. All studies included healthy subjects, while only one (Leufkens et al., 2014) evaluated persons with insomnia; therefore, we could not conduct the analyses for persons with insomnia only, and all analyses were conducted on healthy subjects. Zopiclone 7.5 mg showed a significantly higher driving-impairing effect than placebo in all our analyses, like in previous evidence (Verster et al., 2011), and was a positive control. It was also included in analyses with FDA-approved molecules for sleep disorders only as a positive control. No eligible studies were found regarding estazolam, quazepam, eszopiclone, secobarbital, and butabarbital. The primary analyses are summarized in Figs 2 and 3. Major findings are summarized in Table 2.

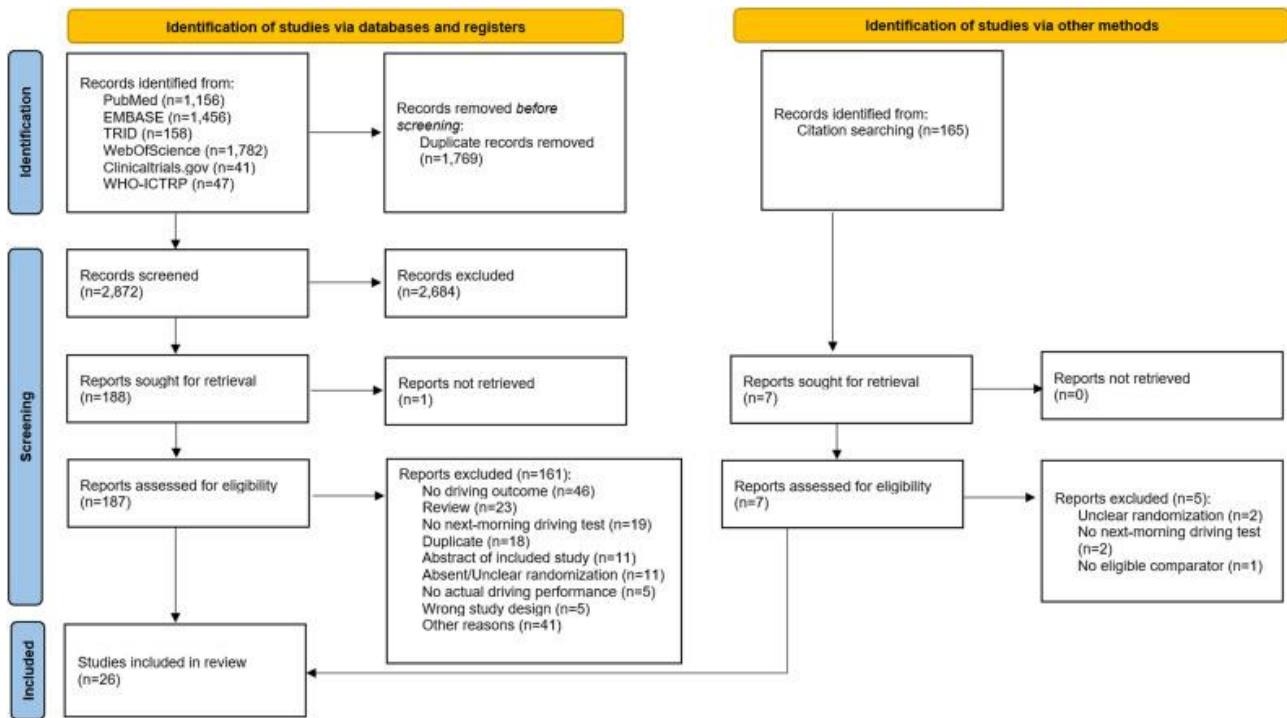


Fig. 1. PRISMA flow diagram.

Table 1. Characteristics of included studies.

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnosis	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
Muehlan et al., 2022	The Netherlands	Simulated	9h	60 (50)	50–80y (half 50–65y and half 65–80y) (64.6 + 7.3)	Healthy subjects	daridorexant 50 mg vs. daridorexant 100 mg vs. zopiclone 7.5 mg vs. placebo	5 days	SDLP	Idorsia Pharmaceuticals	Some concerns	Daridorexant showed less impaired subjects than zopiclone (both day 1 and day 5)
Iwamoto et al., 2022	Japan	Simulated	9h	30 (100)	22–59y (33.1 + 10)	Healthy subjects	zopiclone 7.5 mg vs. placebo	8 days	SDLP	Taisho Pharmaceutical	High	SDLP was significantly higher with zopiclone
Torres et al., 2022	US	Simulated	9h	48 (58.3)	21–55y (NI)	Healthy subjects	Tasimelteon 20 mg vs. zopiclone 7.5 mg vs. placebo	1 day	SDLP	Vanda Pharmaceuticals	Some concerns	Tasimelteon was comparable to placebo and significantly less impairing than zopiclone
Vermeeren et al., 2019	The Netherlands	On-the-road	9h	48 (45.8)	21–78y (half 21–65y and	Healthy subjects	Lemborexant 2.5 mg vs. Lemborexant 5	8 days	SDLP	Eisai, Inc.	Low	All lemborexant doses paralleled placebo effect

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnoses	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
					half 65–78y) (58.5 + 13.3)		mg vs. Lemborexant 10 mg vs. Zopiclone 7.5 mg vs. placebo					
Kay et al., 2016	US	Simulated	8.25h	59 (78)	25–55y (41+9)	Healthy subjects	Diphenhydramine citrate 76 mg vs. triazolam 0.5 mg vs. placebo	1 day	SDLP	Pfizer	Low	Diphenhydramine paralleled placebo, triazolam significantly increased SDLP more than placebo
Simen et al., 2015	Belgium	Simulated	10h	30 (53.3)	25–50y (39.7 + 6.76)	Healthy subjects	Zopiclone 7.5 mg vs. placebo	1 day	SDLP	Merck	Some concerns	Zopiclone significantly increased SDLP
Vermeeren et al., 2015	The Netherlands	On-the-road	unclear	28 (46.4)	23–64y (45.6 + 13.2)	Healthy subjects	Suvorexant 20 mg vs. suvorexant 40 mg vs. zopiclone 7.5 mg vs. placebo	8 days	SDLP	Merck	Some concerns	Suvorexant SDLP was higher than placebo at day 1, but not at day 9

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnoses	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
Vermeeren et al., 2014	The Netherlands	On-the-road	3–4 h (for zolpidem)) 9 h (for zopiclone)	40 (50)	21–64y (37.3 + 14.8)	Healthy subjects	Zolpidem 3.5 mg MOTN vs. zopiclone vs. placebo	1 day	SDLP	Transcept Pharmaceuticals, Inc. Purdue Pharma L.P.	Some concerns	Zolpidem slightly significantly increased SDLP vs. placebo, less than zopiclone, but the impairment rate was comparable to placebo
Leufkens et al., 2014	The Netherlands	On-the-road	9h	16 (56.3) 16 (56.3)	52–71y (62.9 + 4.3) 52–71y (62.6 + 5.25)	Healthy subjects Persons with insomnia (DSM-IV)	zopiclone 7.5 mg vs. placebo	1 day	SDLP	–	Low	SDLP was significantly higher with zopiclone
Iwamoto K. et al., 2013	Japan	Simulated	unclear	13 (100)	32–49y (39.2 + 6.2)	Healthy subjects	Mirtazapine 7.5 mg vs. mirtazapine 15 mg vs. placebo	8 days	SDLP	–	High	Mirtazapine 7.5 mg paralleled placebo, while 15 mg

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnosis	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
Sasada et al., 2013	Japan	Simulated	unclear	19 (100)	26–49y (38.88+6.8)	Healthy subjects	Mirtazapine 15 mg vs. Trazodone 25 mg vs. placebo	8 days	SDLP	–	Some concerns	significantly increased SDLP Mirtazapine (but not trazodone) significantly increased SDLP
Mets et al., 2011	The Netherlands	On-the-road	8.30h	30 (50)	21–55y (25.9 + 6.53)	Healthy subjects	Ramelteon 8 mg vs. zopiclone 7.5 mg vs. placebo	1 day	SDLP	Takeda Europe	Some concerns	Ramelteon significantly increased SDLP, paralleling zopiclone
Bocca M.L. et al., 2011	France	Simulated	10h	16 (50)	NI (60.3 + 2.9)	Healthy subjects	Zolpidem 10 mg vs. Flunitrazepam 1 mg vs. Zopiclone 7.5 mg vs. Placebo	1 day	SDLP	–	Some concerns	Flunitrazepam paralleled placebo, while Zolpidem paralleled zopiclone
Bocca et al., 1999	France	Simulated	9–11h	16 (56.2)	20–30y (24.5+NI)	Healthy subjects	Zolpidem 10 mg vs. flunitrazepam 1 mg vs.	1 day	Mean variance	–	High	Zolpidem paralleled placebo, flunitrazepam and

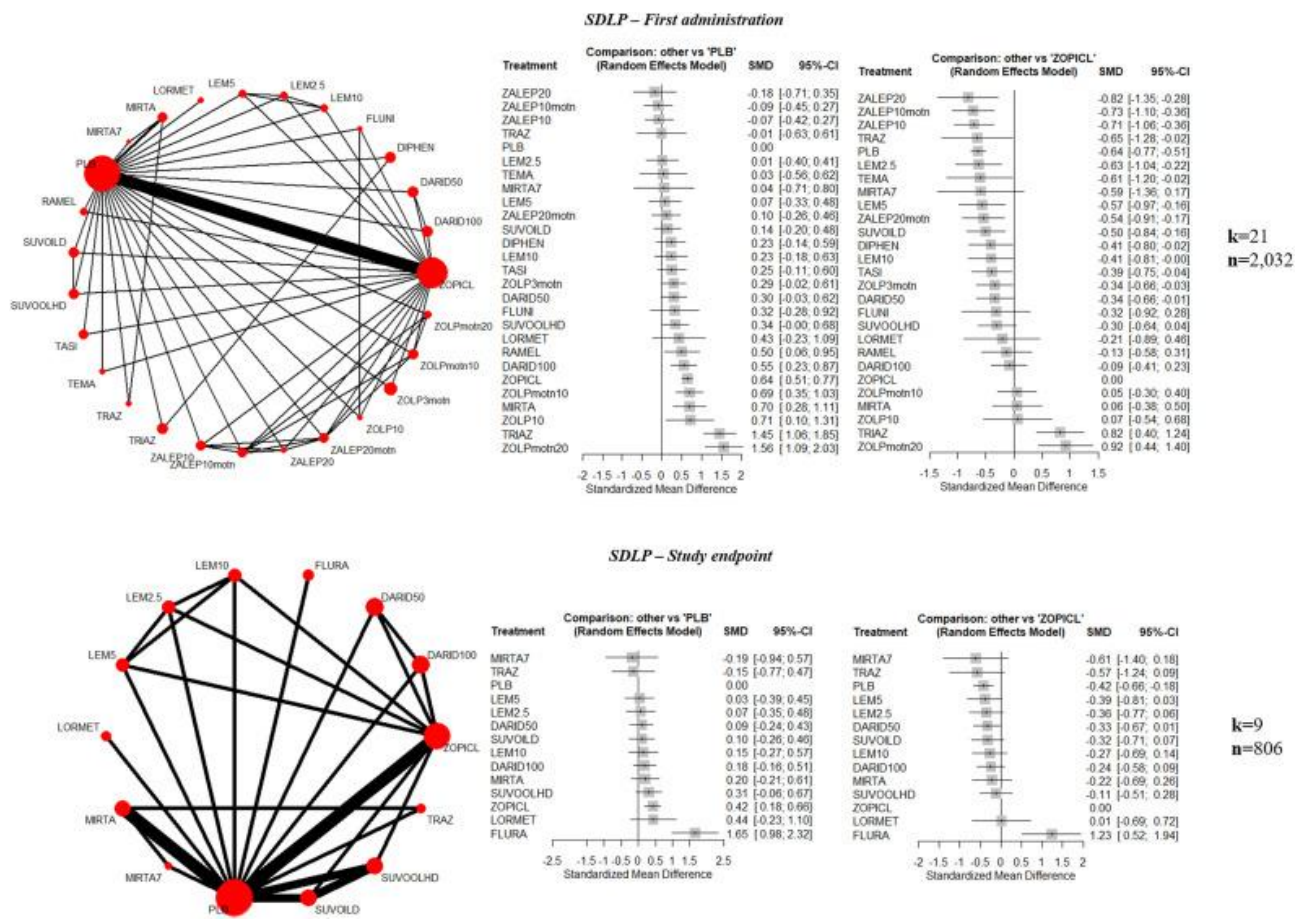
Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnosis	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
							zopiclone 7.5 mg vs. placebo		of lateral position			zopiclone showed higher impairment
Leufkens and Vermeeren, 2009	The Netherlands	On-the-road	10–11h	18 (44.4)	55–75y (64.3 + 4.24)	Healthy subjects	Temazepam 20 mg vs. Zopiclone 7.5 mg vs. placebo	1 day	SDLP	–	High	Temazepam paralleled placebo
Leufkens et al., 2009	The Netherlands	On-the-road	4 h (middle-of-the-night) 10–11 h (zopiclone)	28 (NI)	22–44y (31.4 + 7.5)	Healthy subjects	Zolpidem 10 mg MOTN vs. Zopiclone 7.5 mg vs. placebo	1 day	SDLP	Lundbeck	High	Zolpidem paralleled zopiclone
Otmami et al., 2008	France	Simulated	13h	16 (75)	55–65y (59.4 + 3.2)	Healthy subjects	Zolpidem 10 mg vs. Melatonin prolonged release 2 mg vs. Zolpidem + melatonin vs. placebo	1 day	SD of difference in ideal route	Neurim pharmaceuticals	Some concerns	zolpidem and combination arm had significantly higher SD of difference from ideal route

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnosis	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
Verster et al., 2002	The Netherlands	On-the-road	4h	30 (50)	NI (24+2.4)	Healthy subjects	Zolpidem 10 mg MOTN vs. Zolpidem 20 mg MOTN vs. Zaleplon 10 mg MOTN vs. zaleplon 20 mg MOTN vs. placebo	1 day	SDLP	Wyeth-Ayerst Research	Some concerns	Zolpidem formulation significantly increased SDLP, zaleplon paralleled placebo
Vermeeren et al., 1998a	The Netherlands	On-the-road	10h	24 (0)	22–45y (29+6.1)	Healthy subjects	Flurazepam 30 mg vs. placebo	3 days	SDLP	Schering Corporation	Low	Flurazepam significantly increased SDLP
Vermeeren et al., 2002	The Netherlands	On-the-road	10–11h	30 (50)	21–45y (31.6 + 6.9)	Healthy subjects	Zaleplon 10 mg vs. zopiclone 7.5 mg vs. placebo	1 day	SDLP	Wyeth-Ayerst Research	Low	Zaleplon paralleled placebo
Vermeeren et al., 2016	Belgium	On-the-road	9h	24 (58.3)	65–80y (68.8 + 2.7)	Healthy subjects	Suvorexant 15 mg vs. Suvorexant 30 mg vs. zopiclone	8 days	SDLP	Merck & Co. Inc.	Low	Suvorexant (both doses) paralleled placebo (both day 2 and 9)

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnosis	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
							7.5 mg vs. placebo					
Volkerts et al., 1992	The Netherlands	On-the-road	10h	18 (100)	25–31y (26.3)	Healthy subjects	Lormetazepam 1 mg vs. placebo	2 days	SDLP	–	Low	Lormetazepam slightly increased SDLP relative to placebo
Ramaekers et al., 1998	The Netherlands	On-the-road	15–18h	18 (50)	21–35y (NI)	Healthy subjects	Mirtazapine 15–30 mg vs. placebo	15 days	SDLP	NV Organon	High	Mirtazapine significantly increased SDLP (both day 2 and day 16)
Vermeeren et al., 1998b	The Netherlands	On-the-road	5–6 h 10–11h	28 (50)	23–40y (31+5.7)	Healthy subjects	Zaleplon 10 mg vs. Zaleplon 20 mg vs. Zaleplon 10 mg MOTN vs. Zaleplon 20 mg MOTN vs. zopiclone 7.5 mg vs. placebo	1 day	SDLP	Wyeth-Ayerst	Some concerns	All zaleplon formulations paralleled placebo

Study	Country	Setting	Time from bedtime dose	randomized sample size (male%)	Age range (mean + SD)	Diagnoses	Intervention and dosage	duration of trial	outcome	Sponsor	RoB2	Findings
Philip et al., 2011	France	On-the-road	11h	36 (NI)	21–44y (NI)	Healthy subjects	Doxylamine 7.5 mg vs. diphenhydramine 25 mg vs. zopiclone 7.5mg	1 day	SDLP	Sanofi-Aventis	Some concerns	Doxylamine and zopiclone SDLPs were higher than diphenhydramine
Wingen et al., 2005	The Netherlands	On-the-road	12h	18 (50)	NI (31.4 + 5.8)	Healthy subjects	Mirtazapine 15–30 mg vs. placebo	8 days	SDLP	H. Lundbeck A/S	Low	Mirtazapine significantly increased SDLP

DSM, Diagnostic and Statistical Manual of Mental Disorders (...edition/revision); **MOTN**, middle-of-the-night; **NI**, No Information; **SD**, Standard Deviation, **SDLP**, Standard Deviation of Lateral Position.



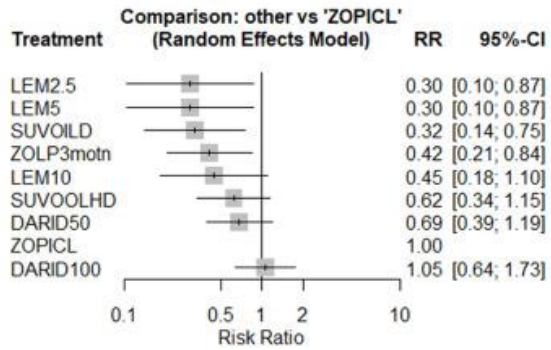
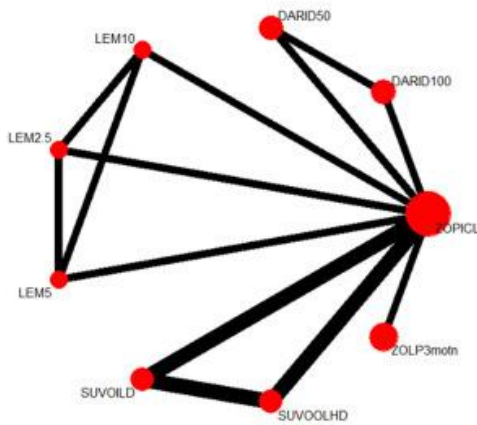
k=21
n=2,032

k=9
n=806

Fig. 2. Standard Deviation of Lateral Position (SDLP), main analyses.

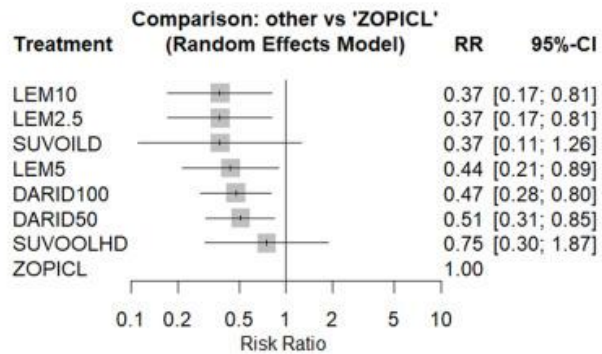
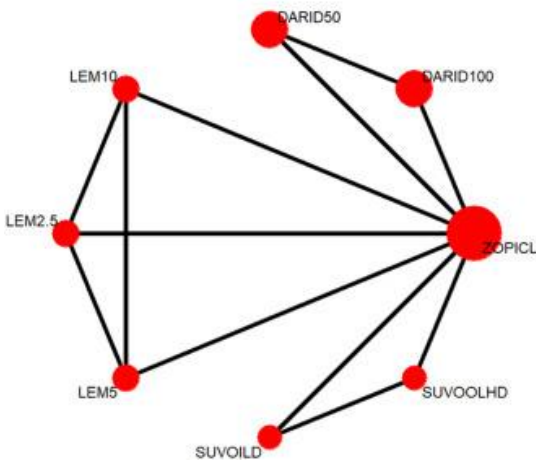
DARID50, daridorexant 50 mg; DARID100, daridorexant 100 mg; DIPHEN, Diphenhydramine 76 mg; FLUNI, Flunitrazepam 1 mg; FLURA, Flurazepam 20 mg; LEM2.5, Lemborexant 2.5 mg; LEM5, Lemborexant 5 mg; LEM10, Lemborexant 10 mg; LORMET, Lormetazepam 1 mg; MIRTA7, Mirtazapine 7 mg; MIRTA, Mirtazapine 15–30 mg; PLB, Placebo; RAMEL, Ramelteon 8 mg; SDLP, Standard Deviation of Lateral Position; SUVOILD, Suvorexant in-label dose (15–20 mg); SUVOOLHD, Suvorexant off-label high dose (30–40 mg); TASI, tasimelteon 20 mg; TEMA, Temazepam 20 mg; TRAZ, Trazodone 25 mg; TRIAZ, Triazolam 0.5 mg; ZALEP10, Zaleplon 10 mg; ZALEP20, Zaleplon 20 mg; ZALEP10motn, Zaleplon 10 mg middle-of-the-night; ZALEP20motn, Zaleplon 20 mg middle-of-the-night; ZOLP3motn, Zolpidem 3.5 mg middle-of-the-night; ZOLPmotn10, Zolpidem 10 mg middle-of-the-night; ZOLP10, Zolpidem 10 mg; ZOLPmotn20, Zolpidem 20 mg middle-of-the-night; ZOPICL, Zopiclone 7.5 mg.

**Driving impairment rate (>2.4cm SDLP vs. PLB)
First administration**



**k=5
n=593**

**Driving impairment rate (>2.4cm SDLP vs. PLB)
Study endpoint**



**k=3
n=399**

Fig. 3. Driving impairment rates (>2.4 cm Standard Deviation of Lateral position vs. placebo), main analyses. DARID50, daridorexant 50 mg; DARID100, daridorexant 100 mg; LEM2.5, Lemborexant 2.5 mg; LEM5, Lemborexant 5 mg; LEM10, Lemborexant 10 mg; SDLP, Standard Deviation of Lateral Position; SUVOILD, Suvorexant in-label dose (15–20 mg); SUVOOLHD, Suvorexant off-label high dose (30–40 mg); ZOPICL, Zopiclone 7.5 mg.

Table 2. Summary of major findings.

In-label doses FDA-approved hypnotics for sleep disorders, both bedtime and middle-of-the-night, paralleled placebo and performed better than zopiclone 7.5 mg both for single and repeated acute (<10 days) administration.

Zolpidem 10 mg and Ramelteon 8 mg performed slightly worse than placebo and paralleled zopiclone 7.5 mg, along with triazolam 0.5 mg which also performed worse than zopiclone, and caution should be applied.

Off-label doses and molecules performed significantly worse than placebo and paralleled zopiclone 7.5 mg and are discouraged.

Repeated acute (<10 days) administration reduced drug differences with placebo, except for flurazepam.

Drivers should not drive if they feel impaired in any way, regardless of the molecule and its dosage.

3.1. Risk of bias assessment

The risk of bias was low in seven, some concerns in 12, and high in six RCTs, respectively ([Table 1](#) & [e-table 1](#)).

3.2. Next-morning driving impairment – first administration (SDLP)

Regarding overall analysis ($k = 21$, $n = 2032$), among 27 interventions, several molecules paralleled placebo SDLP, including zaleplon 10 mg and 20 mg bedtime and middle-of-the-night, trazodone 25 mg, lemborexant 2.5–10 mg, temazepam 20 mg, mirtazapine 7.5 mg, diphenhydramine 76 mg, tasimelteon 20 mg, zolpidem 3.5 mg middle-of-the-night, daridorexant 50 mg, suvorexant 15–20 mg and 30–40 mg, flunitrazepam 1 mg, and lormetazepam 1 mg. Conversely, ramelteon 8 mg (SMD=0.50, 95 %C.I.=0.06;.95), daridorexant 100 mg (SMD=0.55, 95 %C.I.=0.23;.87), zolpidem 10 mg bedtime (SMD=0.71, 95 %C.I.=0.10;1.31), zolpidem middle-of-the-night 10 mg (SMD=0.69, 95 %C.I.=0.35;1.03) and 20 mg (SMD=1.56, 95 %C.I.=1.09;2.03), mirtazapine 15–30 mg (SMD=0.70, 95 %C.I.=0.28;1.11), and triazolam 0.5 mg (SMD=1.45, 95 %C.I.=0.06;1.85) performed significantly worse than placebo. Compared to *zopiclone*, the following medications showed a lower potential for driving impairment, considering SDLP (presented according to the SUCRA rankings): zaleplon 20 mg (SMD=-0.82, 95 %C.I.=-1.35;-0.28), zaleplon 10 mg middle-of-the-night (SMD=-0.73, 95 %C.I.=-1.10;-0.36), zaleplon 10 mg (SMD=-0.71, 95 %C.I.=-1.06;-0.36), trazodone 25 mg (SMD=-0.65, 95 %C.I.=-1.28;-0.02), lemborexant 2.5 mg (SMD=-0.63, 95 %C.I.=-1.04;-0.22), temazepam 20 mg (SMD=-0.61, 95 %C.I.=-1.20;-0.02),

lemborexant 5 mg (SMD=-0.57, 95 %C.I.=-0.97;-0.16), zaleplon 20 mg middle-of-the-night (SMD=-0.54, 95 %C.I.=-0.91;-0.17), suvorexant 15-20 mg (SMD=-0.5, 95 %C.I.=-0.84;-0.16), diphenhydramine 76 mg (SMD=-0.41, 95 %C.I.=-0.80;-0.02), lemborexant 10 mg (SMD=-0.41, 95 %C.I.=-0.81;-0.01), tasimelteon 20 mg (SMD=-0.39, 95 %C.I.=-0.75;-0.04), zolpidem 3 mg middle-of-the-night (SMD=-0.34, 95 %C.I.=-0.66;-0.03), and daridorexant 50 mg (SMD=-0.34, 95 %C.I.=-0.66;-0.01). Furthermore, triazolam 0.5 mg and zolpidem 20 mg middle-of-the-night were outperformed by all other drugs, also performing worse than zopiclone (SMD=1.45, 95 %C.I.=1.06;1.85 & SMD=1.56, 95 %C.I.=1.08;2.03) (e-figures 1-4 & e-table 2). Publication bias was found in zopiclone ($p=.0469$) but not in placebo ($p=.17$) comparisons (e-figures 5-6). Such findings were consistent with the single on-the-road subset ($k = 13$, $n = 1228$) (e-figures 7-12 & e-Table 3) except for zolpidem 3 mg middle-of-the-night and suvorexant 30-40 mg that showed slightly higher SDLP than placebo, but consistently outperformed zopiclone. Regarding simulated driving ($k = 8$, $n = 788$) subset, differences from the main analysis emerged for trazodone 25 mg, mirtazapine 7.5 mg, tasimelteon 20 mg, diphenhydramine 76 mg, and daridorexant 50 mg that showed non-significant lower SDLP compared to zopiclone (e-figures 13-17 & e-Table 4). Consistent findings emerged from sensitivity analysis removing high-risk-of-bias (e-figures 18-29 & e-Tables 5-7) and FDA-approved sleep medication molecules and dosages (e-figures 30-43 & e-Tables 8-10). All analyses showed 0 % heterogeneity and irrelevant inconsistency ($p>.50$) (Appendix S5).

Regarding the elderly ($k = 3$, $n = 243$, all on-the-road studies), lemborexant 2.5mg-10 mg, suvorexant 15-30 mg, and temazepam 20 mg paralleled placebo. All but suvorexant 30 mg and lemborexant 10 mg performed better than zopiclone (presented according to the SUCRA rankings): lemborexant 5 mg SMD=-0.80, 95 %C.I.=-1.39;-0.20; suvorexant 15 mg SMD=-0.71, 95 %C.I.=-1.23;-0.18, lemborexant 2.5 mg SMD=-0.69, 95 %C.I.=-1.28;-0.09; temazepam 20 mg SMD=-0.62, 95 %C.I.=-1.23;-0.01) (e-figures 44-47 & e-Table 11).

3.3. Next-morning driving impairment – study endpoint (SDLP)

The mean number of administration days of our study endpoint overall analysis is 6.56days±2.24 ($k = 9$, $n = 806$). After repeated administration, all evaluated compounds (mirtazapine 7.5 mg and 15-30 mg, trazodone 25 mg, lemborexant 2.5-10 mg, daridorexant 50-100 mg, suvorexant 15-40 mg, lormetazepam 1 mg) paralleled placebo, except for Flurazepam 30 mg (SMD=1.65, 95 %C.I.=0.98;2.32). While almost all compounds showed lower SDLP than zopiclone, no comparison reached statistical significance (e-figures 48-52). No publication bias was detected ($p=.92$) (e-figure 53). Also, flurazepam 30 mg was outperformed by all other compounds (e-Table 12). Similar trends

emerged when separately analyzing on-the-road studies ($k = 6, n = 490$), except for lemborexant 5 mg, which showed lower SDLP than zopiclone (e-figures 54–57 & e-Table 13). In simulated driving analysis ($k = 3, n = 316$), all compounds but flurazepam 30 mg paralleled placebo and zopiclone as in the principal analysis (e-figures 58–61 & e-Table 14). Sensitivity analyses without high-risk-of-bias studies consistently supported the main analysis findings (e-figures 62–73 & e-Tables 15–17), as well as the sensitivity analysis including FDA-approved sleep medication molecules and dosages (e-figures 74–81 & e-Tables 18–19). All analyses showed 0 % heterogeneity and irrelevant inconsistency ($p > .50$) (Appendix S5).

3.4. Next-morning driving impairment rate (SLDP >2.4 cm than placebo) – first administration

The following medications showed a lower impairment risk ratio (presented according to the SUCRA) compared to zopiclone ($k = 5, n = 593$): lemborexant 2.5 mg (RR=0.30, C.I.=0.10;.87), lemborexant 5 mg (RR=0.30, C.I.=0.10;.87), suvorexant 15–20 mg (RR=0.32, C.I.=0.14;.75), zolpidem 3.5 mg middle-of-the-night (RR=0.42, C.I.=0.21;.84). Conversely, lemborexant 10 mg, suvorexant 30–40 mg, and daridorexant 50–100 mg paralleled zopiclone (e-figures 82–84 & e-Table 20). The on-the-road subset analysis confirmed the main analysis (e-figures 85–87 & e-Table 21). Heterogeneity was low (11.2 %), and inconsistency was irrelevant ($p = .32$). All studies were low/unclear risk-of-bias; therefore, we did not perform the risk-of-bias sensitivity analysis. Studies were insufficient to perform simulated driving and FDA-approved sleep medication analyses.

3.5. Next-morning driving impairment rate (SLDP >2.4 cm than placebo) – study endpoint

The mean study endpoint of our overall analysis is 7days \pm 1.41 ($k = 3, n = 399$). The following drugs performed better than zopiclone: lemborexant 2.5 mg (RR=0.37, C.I.=0.17;.81), 5 mg (RR=0.44, C.I.=0.21;.89), 10 mg (RR=0.37, C.I.=0.17;.81), daridorexant 50 mg (RR=0.51, C.I.=0.31;.85), 100 mg (RR=0.47, C.I.=0.28;.80), while suvorexant 15–20 mg/30–40 mg paralleled zopiclone (e-figures 88–90 & e-Table 22). Heterogeneity and inconsistency could not be computed. On-the-road subset analysis (e-figures 91–93 & e-Table 23) was consistent with the main analysis. All studies were low/unclear risk-of-bias; therefore, the pertinent sensitivity analysis was not conducted. Studies were insufficient to perform simulated driving and FDA-approved sleep medication analyses.

3.6. Cinema confidence in evidence

CINeMA methods and domain rating reasoning are presented in Appendix S6. Regarding next-day driving impairment – first administration, considering SDLP (overall analysis), confidence in evidence was low for five mixed evidence comparisons and very low for all other mixed and indirect comparisons (e-Table 27). Concerning on-the-road studies, three low mixed evidence ratings emerged; all other mixed and indirect comparisons were rated very low (e-Table 28). Considering simulated studies, two moderate, three low, and 16 very low emerged in mixed comparisons, while three moderate, seven low, and 36 very low emerged in indirect comparisons (e-Table 29). All ratings were very low considering impairment rate analyses, likely due to the small number of included studies (e-Tables 30–31).

4. Discussion

The report is the first NMA to assess hypnotic residual effects on next-morning driving impairment. Compared to the most current meta-analyses on this topic ([Roth et al., 2014](#); [Verster et al., 2006](#)), the present report provides an up-to-date appraisal of several new drugs (both approved and unapproved) and considers different doses, and thus results provide optimal clinical relevance ([De Prisco and Oliva, 2023](#)). It also included simulated driving studies, accounting for acute and repeated administration of hypnotics. Moreover, the NMA tool allowed multiple simultaneous head-to-head comparisons, which was impossible for previous pairwise meta-analyses. All comparisons showed 0 % heterogeneity and non-significant global inconsistency ($p > .50$), suggesting high comparability of the included studies and the validity of the transitivity assumption, notwithstanding the heterogeneity in driving test settings of the included studies (i.e., on-the-road and simulated driving).

Overall, most appraised FDA-approved molecules and dosages for sleep disorders proved like a placebo for driving-impairing next-morning residual effects. The exceptions were ramelteon 8 mg, triazolam 0.5 mg, and zolpidem 10 mg bedtime which showed significant driving-impairing potential, like zopiclone. Regarding zolpidem 10 mg bedtime findings, caution is warranted, considering that our analysis relied on one study, while another study ([Bocca et al., 1999](#)) (included in the sole systematic review) showed placebo-like impairment, as well as other analysis-ineligible evidence ([Vermeeren et al., 1995](#)). Included study participants' mean age was higher than the latter, suggesting an age influence on zolpidem residual effects, to be investigated by further evidence. Also, most formulations outperformed zopiclone 7.5 mg, renowned for a significant driving-impairing potential. Zaleplon ranked highest as the least driving-impairing in both 10–20 mg

bedtime formulations and an off-label middle-of-the-night formulation, likely due to its very low half-life (≈ 1 h) and lack of active metabolites, which avoids next-morning residual effects. All Dual Orexin Antagonists (DORAs) showed placebo-comparable impairing effects, suggesting a low propensity of driving-related residual effects. Lemborexant 2.5–10 mg followed zaleplon in the SUCRA ranking, proving non-significant residual effects with respect to the driving performance. Also, daridorexant 50 mg and suvorexant 20 mg paralleled placebo.

Tasimelteon 20 mg, a new melatonin receptor agonist approved in 2014 for non-24-hours-sleep-wake-disorder, was associated with placebo-like impairing potential and outperformed zopiclone. Moreover, diphenhydramine 76 mg, temazepam 20 mg, and zolpidem 3.5 mg middle-of-the-night (Intermezzo) paralleled placebo impairment, proving less impairing than zopiclone. Regarding the latter, it proved safe for driving despite being thought for middle-of-the-night awakenings, likely due to the very low dose that can induce sleep by sublingual administration but seems not to retain residual next-morning effects.

Off-label sedating drugs that may be employed when insomnia is comorbid with psychiatric diseases also reduce polypharmacy (Fornaro et al., 2024). Low-dose trazodone (25 mg) and mirtazapine (7.5 mg) showed non-significant driving-related residual effects. However, antidepressant-dose mirtazapine (15–30 mg) caused significant impairment, comparable to zopiclone. Further studies with higher doses are needed to evaluate antidepressants' next-morning residual effect potential, especially at full doses.

As people age, they struggle to fall asleep and maintain their sleep (Gulia and Kumar, 2018). Older adults widely use hypnotics (Conti et al., 2017); however, they may be more sensitive to hypnotic side effects (especially benzodiazepines) due to age-related brain receptor differences and reduced metabolism of certain drugs (Bogunovic and Greenfield, 2004). Therefore, caution is needed when prescribing a hypnotic to such a population. Our analysis (despite relying on few comparisons) showed that, after a single administration, lemborexant 2.5–5 mg, temazepam 20 mg, and suvorexant 15 mg parallel placebo in terms of driving-impairing potential, indicating lower impairing potential than zopiclone.

Repeated administration of hypnotics seemed to reduce the driving-related residual effects, likely due to body tolerance to such effects, except for flurazepam 30 mg, which effects may need more time to build tolerance, given that it showed higher SDLP than placebo and zopiclone-like driving impairment after a three-days administration. This could be problematic, considering benzodiazepines are only approved for short-term management of insomnia, considering their abuse and addiction potential and dangerous interactions with other drugs (Food and Administration,

2020). Mirtazapine 15–30 mg and high doses of daridorexant and suvorexant (100 mg and 40 mg, respectively) paralleled placebo at the study endpoint, notwithstanding they caused significant impairment after the first administration.

Impairment rate analyses are particularly important because they are not affected by placebo absolute SDLP; conversely, they appraise an impairment definition with a threshold of 2.4 cm difference with placebo, corresponding to the 'swerving' of 0.5 g/L BAC (Louwerens et al., 1987), the legal limit in several countries. Lemborexant 2.5–5 mg, suvorexant 15–20 mg, and zolpidem 3.5 mg middle-of-the-night impairment rates were significantly lower than zopiclone at the first-morning driving test. In comparison, suvorexant 30–40 mg and daridorexant 50–100 mg showed zopiclone-like impairment rates. However, at the study endpoint, while daridorexant proved less impairing than zopiclone, both suvorexant in-label (20 mg) and off-label (40 mg) doses showed similar impairment rates vs. zopiclone, suggesting a less tolerance-building potential for its driving-related residual effects.

4.1. Limitations of the study

The study's main limitation is represented by the few reports that most comparisons relied on, mainly due to the recent introduction of several appraised drugs and the small sample of most studies, warranting further evidence to confirm our findings, especially regarding repeated administration of hypnotics. Furthermore, all appraised studies included healthy subjects, while hypnotics are thought for people suffering from insomnia, prompting caution in the critical interpretation and applicability of the findings. Hypnotic residual effects likely depend on the pharmacodynamics, pharmacokinetics, and dosage of the employed drugs (Vermeeren, 2004). However, persons with insomnia may already experience daytime sleepiness due to the underlying sleep disorder and may experience fewer residual effects due to an overall improvement in daytime vigilance, performance, and driving-related skills, if insomnia is successfully treated. Furthermore, persons with insomnia could have already used hypnotics for prolonged periods, showing increased tolerance to the residual effects. Therefore, driving studies performed on healthy subjects may overestimate the residual effect. Despite previous evidence showing that such overestimation is minor (Leufkens et al., 2014), further person-with-insomnia-based evidence is warranted to confirm our findings. Furthermore, while SDLP may not fully capture the complex aspects of driving performance, we did not appraise other proxy measures since the latter relied on inconsistent operational definitions and lower reliability. Finally, the low/very low confidence in the evidence of most comparisons elicits further caution in interpreting the results. We employed a conservative

reasoning approach, especially regarding including simulated driving studies, as detailed in Appendix S6.

4.2. Conclusions

In-label doses of most FDA-approved hypnotics proved comparable to placebo and better than zopiclone regarding residual drive-impairing effects, both in single and repeated administration (appraised in less than ten continuous days of administration, always). Caution should be applied with ramelteon, triazolam 0.5 mg, flurazepam 30 mg, zolpidem 10 mg due to the emerging zopiclone-like impairment. Also, we discourage using off-label medications (unless approved for comorbid conditions), and doses emerge as potentially highly driving-impairing. Also, the driving-related residual effect of most hypnotics seems to be affected by tolerance after repeated administration. This is particularly important because insufficient sleep can affect daytime alertness and impair driving by itself. However, further studies appraising longer follow-up time are warranted, to evaluate the effects of a chronic hypnotic administration. Sleep-improving medications without driving-impairing effects are crucial to improve the daytime performance of persons with insomnia without increasing the traffic accident risk. Insomnia therapy should always be patient-tailored, considering insomnia subtypes, comorbidities, and side effects. Also, drivers should not drive if they feel impaired in any way, regardless of the molecule and its dosage.

Role of the funding source

None.

CRedit authorship contribution statement

Michele Fornaro: Data curation, writing - original draft, writing - review & editing; **Claudio Caiazza**: Conceptualization, data curation, formal analysis, Writing - original draft, writing - review & editing; **Flavia Rossano, Andrea de Bartolomeis, Felice Iasevoli**: Supervision; **Eduard Vieta, Trevor Thompson, Marco Solmi, Andre Ferrer Carvalho**: Writing - original draft, writing - review & editing, methodology; **Flavia Cilmi**: data extraction; **Michele De Prisco**: Statistical analysis and manuscript revision.

Declaration of competing interest

E.V. has received grants and served as consultant, advisor or CME speaker for the following entities: AB-Biotics, AbbVie, Adamed, Angelini, Biogen, Biohaven, Boehringer-Ingelheim, Celon Pharma, Compass, Dainippon Sumitomo Pharma, Ethypharm, Ferrer, Gedeon Richter, GH Research, Glaxo-Smith Kline, HMNC, Idorsia, Janssen, Lundbeck, Medincell, Merck, Novartis, Orion Corporation, Organon, Otsuka, Roche, Rovi, Sage, Sanofi-Aventis, Sunovion, Takeda, and Viatrix, outside the submitted work. **M.S.** has received honoraria/has been a consultant for Angelini, Lundbeck, and Otsuka. **M.F.** served as a rater for Massachusetts General Hospital Clinical Trials Network and Institute, Boston, MA, USA, and its subsidiaries for the following entities. **M.F.** received honoraria from the American Society of Clinical Psychopharmacology (ASCP) for his speaker activities, and from Angelini, Lundbeck, Bristol Meyer Squibb, and Boehringer-Ingelheim. **Ad.B.** has received research support from Janssen, Lundbeck, and Otsuka and lecture fees for educational meeting from Chiesi, Lundbeck, Roche, Sunovion, Viatrix, Recordati, Angelini and Takeda; he has served on advisory boards for Eli Lilly, Jansen, Lundbeck, Otsuka, Roche, and Takeda, Chiesi, Recordati, Angelini, Viatrix. **All other authors** declare that they have no conflicts of interest.